Statistics and Its Interface Volume 14 (2021) 229–241

Bayesian confidence intervals for variance of delta-lognormal distribution with an application to rainfall dispersion

Patcharee Maneerat*, Suparat Niwitpong†, and Sa-Aat Niwitpong‡,§

For climate studies in agriculture, rainfall records often involve data which contain zeros and highly non-zero skewness. This is mostly used in models for prediction or that use the mean for approximation. Rainfall dispersion is also important in evaluations as it can vary enormously, and it is a natural phenomenon which can lead to drought or flood. Herein, the goal of this paper is to propose a variational approximation computed with interval estimator based on Bayesian approach for delta-lognormal variance consisting of the highest posterior density interval based on vague prior (HPD-V) and the method of variance estimates recovery (MOVER). By way of comparison, the performances of these intervals were evaluated in terms of coverage probability and relative average length via a Monte Carlo simulation. The numerical results show that HPD-V was much more likely to outperform the other methods in many situations even large variance, although MOVER became the recommended method when both of variance and the probability of having zero were small. Our methods were then be utilized to analyze the variability in Nan province's daily rainfall dataset in a comparison with the other methods.

Keywords and phrases: Agriculture, Bayesian approach, MOVER, Natural rainfall, Vague prior, Variance.

1. INTRODUCTION AND MOTIVATION

Thailand is a highly agricultural country and natural rainfall is a major factor for crop farming. Rice is the country's most important crop, with some 60% of Thailand's 13 million farmers growing it on cultivated land [3], and it is expected that rice production will increased to approximately 21.2 million metric tons in the market year 2018–2019 [30]. Northern Thailand, one of the six geographical regions in the country and produces a great deal of the main agricultural crops, such as rice, fruit, vegetables, and so on. It is mostly covered by mountains and forest, and is the origin of many streams and rivers in Thailand.

In 2010, northern Thailand was one of the regions faced with the worst drought in the past 20 years and its impact significantly led to a severe reduction in the rainfall amount around the Mekong river (a principal river system in Thailand) [28], with researchers believing that this could well be due to climate change [48]. Conversely, Thailand is also prone to heavy rains and flooding due to the influence of the southwest and northeast monsoon winds [11]. Indeed, 12.8 million Thai people were subjected to widespread flooding in 2011, which also damaged around 16,668.55 km^2 of agricultural land [34]. The worst impact of this was in northern Thailand where flash flooding often occurs in the early rainy season [37]. In 2018, landslides during the rainy season (mid-May to mid-October) that occurred in Nan province (located in northern region) was due to heavy rainfall for around two weeks; there were even victims living in the mountainous areas [22].

Duangdai and Likasiri [16] argued that natural disasters such as droughts and floods are the result of severe rainfall oscillation and ineffectual or unsuccessful water management. These phenomena can have profound effects on agriculture and can lead to significant loss of life, damage to buildings and other infrastructure. Needless to say, if variation resulting in too little or too much rainfall are known in advance, this information could be advantageous to the Thai government to manage the occurrence of natural disasters and mitigate their effects. With this in mind, we consider rainfall oscillation in Nan province measured in terms of the variance to estimate changes in rainfall amount based on historical data. A histogram, Q-Q plot and Akaike information criteria (AIC), see an empirical application, results revealed that the daily rainfall data for Nan in northern Thailand during rainy season fit a delta-lognormal distribution

It is widely known that lognormal is the distribution of positive skewness depending on the variance in the probability theory e.g. [4, 7, 23, 26]. Additionally, a few situations also contain zero observations indicating one or more of them are empty, which leads to a delta-lognormal distribution first introduced by [1]. In the real world, this distribution is utilized in several fields, including fisheries

*ORCID: 0000-0003-2145-9657.
†ORCID: 0000-0003-3059-1131.
‡ORCID: 0000-0001-8269-3397.
§Corresponding author.

[19, 33, 35, 36, 47], environment [32, 38, 39] and medicine [40, 49].

Variance is regarded as a measure of dispersion in applied statistics and is one of the most popular parameters of interest in probability and statistical inference. It is defined as the second central moment, and the positive square root of the variance is the standard deviation [10].

The application of confidence intervals (CIs) for the variance has been utilized by various authors in a number of fields. For example, Mathew and Webb [29] estimated the difference in tube-to-tube dispersion between a new gun tube and a control tube to investigate the gun tube accuracy of an M1 Series tank, while Krishnamoorthy et al. [26] evaluated the levels of lead in air as a health hazard impacting the well-being of personnel at the site in Alma American Labs. Furthermore, Cojbasic and Loncar [13] assessed the annual revenue levels of the industry players' market shares to measure market concentration indices.

Several researchers have investigated and developed CIs for variance with different methods and distributions. For example, Burdick and Graybill [8] presented a modification of the Graybill-Wang procedure to establish them for a linear combination of the among- and within-group variances in an unbalanced one-way classification model; their results led them to recommend the Graybill-Wang shortest interval instead of either the Graybill-Wang equal-tail or Satterthwaite equal-tail intervals. Bebu and Mathew [4] studied a modified signed log-likelihood ratio test with generalized confidence intervals (GCIs) in a bivariate lognormal distribution; they found that GCIs are recommended in cases where the sample size was small. Cohen [12] adjusted CIs for the variance in a normal distribution using the two questions posed by Tate and Klett. Mathew and Webb [29] presented the generalized p-value and GCI for variance components. Harvey and Merwe [23] proposed Bayesian CIs for variance of a lognormal distribution with additional zero observations. Note that Bayesian intervals (including fiducial ones) are fundamentally different to frequentist intervals: the former are interpreted as probability statements a posteriori while the latter are procedures that have certain properties a prior. Niwitpong [31] recommended GCIs for the function of variance in a lognormal distribution because it performed well in terms of coverage.

The expected value or mean is a measure of central tendency [10], although the variance of rainfall data can be indicated a measure of rainfall dispersion. Importantly, its estimated variance can be pointed out unnatural rainfall fluctuation which can result in such natural disasters as drought and floods. Determining the CIs for variance is well-known and can be applied in many research areas, as mentioned previously. Casella and Berger [10] guaranteed that a CI provides information on the parameter of interest more than the point estimates. As such, CIs for the variance of delta-lognormal distribution are needed to predict extreme rainfall events. Herein, we present a developed method to construct Bayesian CIs for the variance in a delta-lognormal distribution called the highest posterior density (HPD) credible interval based on vague prior (HPD-V) and the method of variance estimates recovery (MOVER). The reason is because the results of the existing HPD interval based on the independence Jeffreys' prior (HPD-J) of Harvey and Merwe [23] is a better performance for variance of lognormal with additional zero observations than the equal-tailed CI. Furthermore, there are the existing methods: the GCI of Wu and Hsieh [47], and FGCI of Hasan and Krishnamoorthy [24] (FGCI-HK). These were obtained from the CIs for delta-lognormal mean. In Section 2, the theories and concepts of all of the methods are elaborated. In Section 3, the design and results of the simulation procedure are shown. In Section 4, the dispersion of daily rainfall in Nan province is estimated by all of the methods. Finally, a summary and brief discussion are contained in Section 5.

2. NOTATION AND INTERVALS FORMED ON THE SINGLE VARIANCE

Consider the population data including many zeroes with the probability $0 < \delta < 1$ and the non-zeroes with the remain probability $1-\delta$. The number of zero $n_{(0)}$ has a binomial distribution with parameters n and δ, while the non-zero values have lognormal with the mean μ and the variance σ^2. These characteristics are called the delta-lognormal distribution. Define $X = (X_1, X_2, \ldots, X_n)$ be a non-negative random samples of delta-lognormal distribution presented by [2], denoted as $\Delta(\mu, \sigma^2, \delta)$. The distribution function of X is given by

$$(1) \qquad G(x;\mu,\sigma^2,\delta) = \begin{cases} \delta; & x = 0 \\ (1-\delta)F(x;\mu,\sigma^2); & x > 0 \end{cases}$$

where $F(x;\mu,\sigma^2)$ is the distribution function of lognormal so that $\ln X \sim N(\mu,\sigma^2)$ where μ and σ^2 are the mean and variance of $\ln X$. The maximum likelihood estimates (MLEs) of μ, σ^2 and δ are $\hat{\mu} = \frac{1}{n_{(1)}} \sum_{i=1}^{n_{(1)}} \ln X_i$, $\hat{\sigma}^2_{mle} = \frac{1}{n_{(1)}} \sum_{i=1}^{n_{(1)}} [\ln X_i - \hat{\mu}]^2$ and $\hat{\delta} = \frac{n_{(0)}}{n}$; $n = n_{(0)} + n_{(1)}$ where $n_{(1)}$ are the number of non-zero observed values. The population variance of X is

$$(2) \qquad \psi = (1-\delta)\exp(2\mu+\sigma^2)\left[\exp(\sigma^2) - (1-\delta)\right]$$

which is log-transformed as $\varphi = \ln(1-\delta) + (2\mu+\sigma^2) + \ln\left[\exp(\sigma^2) - (1-\delta)\right]$. The MLEs $\hat{\mu}$, $\hat{\sigma}^2$ and $\hat{\delta}$ are replaced and led to obtain the estimate $\hat{\varphi} = \ln(1-\hat{\delta}) + (2\hat{\mu}+\hat{\sigma}^2) + \ln\left[\exp(\hat{\sigma}^2) - (1-\hat{\delta})\right]$. The CIs for ψ are established using the following methods.

2.1 HPD credible interval

According to Casella and Berger [10], Bayesian approach is defined as a parameter ψ regarded as a quantity whose variation described by a probability distribution, called the prior distribution. This is formulated based on subjective belief before the data were collected. A sample is then obtained from a population with parameter ψ. The posterior distribution is derived from the updated prior distribution with this sample information. Apply Bayes' Rule, this information about updating a distribution is obtained.

The HPD credible interval is constructed to estimate parameter based on Bayesian approach. Given the posterior density function, the HPD is regarded as the narrowest possible interval for a parameter of interest at a probability $100(1-\alpha)\%$. Note that this interval contains the required mass such that all points within the interval has a higher probability density than any other point outside, defined by [5]. This study investigates $X \sim \Delta(\mu, \sigma^2, \delta)$ where $Y = \ln X \sim N(\mu, \sigma^2)$ and $n_{(0)} \sim B(n, \delta)$. Let $\psi = \begin{pmatrix}\mu & \sigma^2 & \delta\end{pmatrix}'$ be a unknown parameter. Then, the likelihood distribution of ψ is

$$L(\psi|x) = \binom{n}{n_{(0)}} \delta^{n_{(0)}} (1-\delta)^{n_{(1)}} (2\pi\sigma^2)^{-n_{(1)}/2} \times \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n_{(1)}} (\ln x_i - \mu)^2\right\} \tag{3}$$

which is taken by logarithm

$$\log L(\psi|x) = constant + n_{(0)} \log\delta + n_{(1)} \log(1-\delta) - \frac{n_{(1)}}{2}\log 2\pi\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n_{(1)}} (\ln x_i - \mu)^2$$

The first and second derivatives are

$$\frac{\partial}{\partial\mu}\log L(\psi|x) = \frac{1}{\sigma^2}\sum_{i=1}^{n_{(1)}} (\ln x_i - \mu)$$

$$\frac{\partial}{\partial\sigma^2}\log L(\psi|x) = -\frac{n_{(1)}}{2\sigma^2} + \frac{1}{2\sigma^4}\sum_{i=1}^{n_{(1)}} (\ln x_i - \mu)^2$$

$$\frac{\partial}{\partial\delta}\log L(\psi|x) = \frac{n_{(0)}}{\delta} - \frac{n_{(1)}}{1-\delta}$$

and

$$\frac{\partial^2}{\partial\mu^2}\log L(\psi|x) = -\frac{n_{(1)}}{\sigma^2}$$

$$\frac{\partial^2}{\partial(\sigma^2)^2}\log L(\psi|x) = \frac{n_{(1)}}{2\sigma^4} - \frac{1}{\sigma^6}\sum_{i=1}^{n_{(1)}} (\ln x_i - \mu)^2$$

$$\frac{\partial^2}{\partial\delta^2}\log L(\psi|x) = -\frac{n_{(0)}}{\delta^2} - \frac{n_{(1)}}{(1-\delta)^2}$$

where $n_{(0)} \sim Binomial(n, \delta)$ and $n = n_{(0)} + n_{(1)}$. The Fisher information matrix was proved from the likelihood functions (3) as

$$I(\psi) = \operatorname{diag}\begin{bmatrix}\frac{n(1-\delta)}{\sigma^2} & \frac{n(1-\delta)}{2\sigma^4} & \frac{n}{\delta(1-\delta)}\end{bmatrix} \tag{4}$$

The following priors of ψ are detailed as follows.

2.1.1 Independence Jeffrey's prior

Harvey and Merwe [23] proposed the independence Jeffreys' prior to construct the equal-tailed CI and HPD interval for the variance of lognormal distribution with excess zeros. This prior is obtained to be proportional to the square root of the determinant of the Fisher information matrix in Eq. (4) for ψ so that the independence Jeffrey's prior for ψ is

$$P(\psi)_J = \sqrt{I(\sigma^2)I(\delta)} \propto \sigma^{-2}[\delta(1-\delta)]^{-1/2} \tag{5}$$

2.1.2 Vague prior

The prior distribution of this study based on our believe is proposed as $P(\psi)_V = \sigma^{-2}\delta^{c-1}(1-\delta)^{d-1}/B(c,d)$; $\delta \sim Beta(c = 1, d = 1)$ and $B(c,d) = \Gamma(c)\Gamma(d)/\Gamma(c+d)$. Since $c = d$ such that the beta distribution is symmetric. Moreover, a special case of the $Beta(1,1)$ can be called the uniform prior presented by [9] applied in astronomy.

Theorem 1. *Let $X \sim \Delta(\mu, \sigma^2, \delta)$, and the variance of X is $\phi = (1-\delta)\exp(2\mu+\sigma^2)[\exp(\sigma^2) - (1-\delta)]$. The vague prior for ψ is $P(\psi) = \sigma^{-2}\delta^{c-1}(1-\delta)^{d-1}/B(c,d)$. Given observed observations x, it is combined with its likelihood so that the posterior distribution of ψ becomes*

$$p(\psi|x) = \binom{n}{n_{(0)}} \frac{\delta^{n_{(0)}+c-1}(1-\delta)^{n_{(1)}+d-1}}{B(c,d)} (2\pi)^{-\frac{n_{(1)}}{2}} (\sigma^2)^{-\frac{n_{(1)}+2}{2}} \times \exp\left\{-\frac{1}{2\sigma^2}\left[(n_{(1)}-1)\hat{\sigma}^2 + n_{(1)}(\hat{\mu}-\mu)^2\right]\right\} \tag{6}$$

where the marginal posterior densities of σ^2, μ, and δ are $\sigma^2|x \sim InverseGamma((n_{(1)}-1)/2, (n_{(1)}-1)\hat{\sigma}^2/2)$; $\mu|\sigma^2, x \sim N(\hat{\mu}, \sigma^2/n_{(1)})$, and $\delta|x \sim Beta(n_{(0)}+c, n_{(1)}+d)$; $c = d = 1$, respectively.

Theorem 1 was proved; see the Appendix A. Algorithm 1 describes the steps to compute both HPD intervals.

Algorithm 1

1. Generate $\sigma^{2*} \sim IG([n_{i(1)}-1]/2, [n_{i(1)}-1]\hat{\sigma}^2/2)$
2. Given σ^{2*}, generate $\mu^* \sim N(\hat{\mu}, \sigma^{2*}/n_{(1)})$
3. Generate δ^*, denoted as the posterior of δ based on priors:
   - Jeffreys' prior: $\delta_J \sim Beta(n_{(0)}+0.5, n_{(1)}+0.5)$
   - Vague prior: $\delta_V \sim Beta(n_{(0)}+1, n_{(1)}+1)$
4. Compute ψ^* in each prior
5. Repeat 1–4 a number of times (say, $m = 10000$)
6. Compute $(100-\alpha)\%$HPDs for ψ^* in each prior

2.2 GCI

Weerahandi [44] proposed the GCI method which has crucial conditions under the ideas of pivotal quantity (GPQ), defined as

Definition 1. *Suppose that* $X = (X_1, X_2, \ldots, X_n)$ *be random variables with* $f_X(x; \varphi, \varsigma)$ *where* φ *is the interesting parameter and* ς *is the vector of nuisance parameter. The percentile of GPQ* $R(X; x, \varphi, \varsigma)$ *is the interval estimator for* φ *based on the GCI if the following conditions are satisfied as follows:*

(GPQ1) Given X*, the* $R(X; x, \varphi, \varsigma)$ *distribution is free from all unknown parameters.*
(GPQ2) The observed value of $R(X; x, \varphi, \varsigma)$ *does not depend on the nuisance parameters.*

Then, the $100(1-\alpha)\%$GCI for φ is $CI_{gci} = [R(\alpha/2), R(1-\alpha/2)]$; $R(\alpha)$ be the α-th percentile of $R(X; x, \varphi, \varsigma)$. In our study, the three pivot quantities of μ, σ^2 and δ are used to construct CIs for ψ. Krishnamoorthy and Mathew [25] presented the GPQs for μ and σ^2 as

$$R_\mu = \hat{\mu} - W\sqrt{\frac{(n_{(1)}-1)\hat{\sigma}^2}{n_{(1)}U}} \tag{7}$$
$$R_{\sigma^2} = \frac{(n_{(1)}-1)\hat{\sigma}^2}{U}$$

where $W = (\hat{\mu}-\mu)/\sqrt{\frac{\sigma^2}{n_{(1)}}} \sim N(0,1)$ and $U = \frac{(n_{(1)}-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n_{(1)}-1}$ are mutually independent random variables. For considering the GPQ of δ, the coverage properties of variance stabilizing transformation (VST) were better than Wald interval for binomial proportion, see more details [6, 14]. Furthermore, the works of [14], and [47] were investigated the GPQ for δ, given by

$$R_{\delta.vst} = \sin^2\left[\arcsin\sqrt{\hat{\delta}} - \frac{T}{2\sqrt{n}}\right] \tag{8}$$

where $T = 2\sqrt{n}(\arcsin\sqrt{\hat{\delta}} - \arcsin\sqrt{\delta}) \sim N(0,1)$. Hence, the GPQ of ψ is

$$R_{\psi.vst} = (1-R_{\delta.vst})\exp(2R_\mu + R_{\sigma^2}) \times \left[\exp(R_{\sigma^2}) - (1-R_{\delta.vst})\right] \tag{9}$$

which satisfies the two conditions for being a GPQ. The $100(1-\alpha)\%$GCI for ψ is

$$CI_{\psi.vst} = [L_{\psi.vst}, U_{\psi.vst}] = [R_{\psi.vst}(\alpha/2), R_{\psi.vst}(1-\alpha/2)] \tag{10}$$

where $R_{\psi.vst}(\alpha)$ denotes the α-th percentile of $R_{\psi.vst}$. The GCI is computed by the steps in Algorithm 2.

Algorithm 2

1. Generate $W_i \sim N(0,1)$, $U_i \sim \chi^2_{n_{i(1)}-1}$ and $W_i \sim N(0,1)$ are independent
2. Compute R_μ, R_{σ^2}, and $R_{\delta.vst}$
3. Compute $R_{\psi.vst}$
4. Repeat 1–4 a number of times (say, $m = 10000$)
5. Compute $(100-\zeta)\%$GCI for ψ

2.3 FGCI

The fiducial probability is an inverse probability, presented by [17, 18]. Hannig et al. [21] investigated to construct CI by the direct association between GCI and FGCI, while the fiducial GPQ (FGPQ) is a subclass of GPQ. Importantly, FGPQ had a good theory and performance when practiced in an application, see more details e.g. [20, 27, 42, 43]. The definition of FGPQ is elaborated by [21] indicating the condition (FGPQ1) is similar to the condition (GPQ1), although the condition (FGPQ2) is more accurate than (GPQ2).

Here FGCI for ψ is considered. Recall that $\hat{\mu} \sim N(\mu, \sigma^2/n_{(1)})$ and $(n_{(1)}-1)\hat{\sigma}^2/\sigma^2 \sim \chi^2_{n_{(1)}-1}$ are independent random variables. The structure functions of $\hat{\mu}$ and $\hat{\sigma}^2$ are $\left(\hat{\mu} = \mu + W\sqrt{\sigma^2/n_{(1)}}, \sigma^2 U/(n_{(1)}-1)\right)$ being the functions of W and U, respectively. Note that $W \sim N(0,1)$ and $U \sim \chi^2_{n_{(1)}-1}$ be independent random variables. Given observed values, the FGPQs for μ and σ^2 are defined as

$$G_\mu = \hat{\mu}_0 - W_0\frac{\hat{\sigma}}{\sqrt{n_{(1)}}}\sqrt{\frac{n_{(1)}-1}{U_0}} \tag{11}$$

$$G_{\sigma^2} = \frac{(n_{(1)}-1)\hat{\sigma}^2}{U_0} \tag{12}$$

where W_0 and U_0 are independent copy of U and W, respectively. For focusing on the FGPQ of δ, Hasan and Krishnamoorthy [24] proposed the GFPQ of $1-\delta$ as $G_{1-\delta} \sim Beta(n_{(1)}+0.5, n_{(0)}+0.5)$. Their FGPQ of ψ is then

$$G_\psi^{HK} = G_{1-\delta}\exp(2G_\mu + G_{\sigma^2})\left[\exp(G_{\sigma^2}) - G_{1-\delta}\right] \tag{13}$$

The $[G_\psi^{HK}(\alpha/2), G_\psi^{HK}(1-\alpha/2)]$ becomes the $100(1-\alpha)\%$FGCI-HK for ψ. The steps of FGCI computations is detailed as Algorithm 3.

Algorithm 3

1. Generate $W, U \sim N(0,1)$ are independent
2. Compute G_μ, G_{σ^2} and $G_{1-\delta}^{HK}$ to obtain G_ψ^{HK}
3. Repeat 1–2 a number of times (say, $m = 10000$)
4. Compute $(100-\zeta)\%$FGCI-HK for ψ

2.4 MOVER

Zou et al. [50] first presented a MOVER to construct CIs for the difference of effect measures. The concept of MOVER are described. Let θ_i be the parameter of interest; $i = 1, 2$. Suppose that the point estimates $\hat{\theta}_i$ are independently distributed. Traditionally, the approximately $100(1-\alpha)\%$CI for $\theta_1 - \theta_2$ under the central limit theorem (CLT), defined as

$$CI = [L, U] = \left[\hat{\theta}_1 - \hat{\theta}_2 \mp z_{\alpha/2}\sqrt{v\hat{a}r(\hat{\theta}_1) + v\hat{a}r(\hat{\theta}_2)}\right] \tag{14}$$

where z_α stands for the α^{th} percentile of the standard normal and $\hat{var}(\hat{\theta}_i)$ be the estimated variance of $\hat{\theta}_i$. Assume that (l_1, u_1) and (l_2, u_2) are the $100(1-\alpha)\%$CI for θ_1 and θ_2, respectively. The $l_1 - u_2$ is closer to L than $\hat{\theta}_1 - \hat{\theta}_2$. The $u_1 - l_2$ is closer to U than $\hat{\theta}_1 - \hat{\theta}_2$. Hence, we estimate $\hat{var}(\hat{\theta}_1)$ under $\theta_1 = l_1$, and $\hat{var}(\hat{\theta}_2)$ under $\theta_2 = u_2$ to obtain L. Likewise, we estimate $\hat{var}(\hat{\theta}_1)$ under $\theta_1 = u_1$ and $\hat{var}(\hat{\theta}_2)$ under $\theta_2 = l_2$ to obtain U. Then,

$$\hat{var}(\hat{\theta}_i) = \frac{(\hat{\theta}_i - l_i)^2}{z^2_{\alpha/2}}; \quad \theta_i = l_i \qquad (15)$$
$$\hat{var}(\hat{\theta}_i) = \frac{(u_i - \hat{\theta}_i)^2}{z^2_{\alpha/2}}; \quad \theta_i = u_i$$

which are replaced in Eq. (14). Therefore, the $100(1-\alpha)\%$MOVER for $\theta_1 - \theta_2$ is given by

$$CI_{\theta_1-\theta_2} = \left[\hat{\theta}_d - z_{\alpha/2}\sqrt{(\hat{\theta}_1 - l_1)^2 + (u_2 - \hat{\theta}_2)^2},\ \hat{\theta}_d + z_{\alpha/2}\sqrt{(u_1 - \hat{\theta}_1)^2 + (\hat{\theta}_2 - l_2)^2}\right] \qquad (16)$$

where $\hat{\theta}_d = \hat{\theta}_1 - \hat{\theta}_2$. Similarly, the $100(1-\alpha)\%$CI for $\theta_1 + \theta_2$ was extended by [51], then

$$CI_{\theta_1+\theta_2} = \left[\hat{\theta}_s - z_{\alpha/2}\sqrt{(\hat{\theta}_1 - l_1)^2 + (\hat{\theta}_2 - l_2)^2},\ \hat{\theta}_s + z_{\alpha/2}\sqrt{(u_1 - \hat{\theta}_1)^2 + (u_2 - \hat{\theta}_2)^2}\right] \qquad (17)$$

where $\hat{\theta}_s = \hat{\theta}_1 + \hat{\theta}_2$. This method is used to establish CI for φ. Recall that $\varphi = \ln(1-\delta) + (2\mu + \sigma^2) + \ln\left[\exp(\sigma^2) - (1-\delta)\right]$. Using $\hat{\mu}$, $\hat{\sigma}^2$ and $\hat{\delta}$ from a sample, the estimate $\hat{\varphi}$ is obtained. The CIs for ψ will be completely constructed if the CIs for μ, σ^2 and δ are also established as follows.

Firstly, the ideas are practicable to establish the CI for σ^2. The unbiased estimate for σ^2 is $\hat{\sigma}^2 = \frac{1}{n_{(1)}-1}\sum_{i=1}^{n_{(1)}}(\ln X_i - \hat{\mu})^2$ and obtain that

$$U = \frac{(n_{(1)}-1)\hat{\sigma}^2}{\sigma^2} \qquad (18)$$

where U is chi-square distribution with $n_{(1)} - 1$ degrees of freedom. The coverage probability for $\chi^2_{n_{(1)}-1}$ is $P\left(\chi^2_{\frac{\alpha}{2}, n_{(1)}-1} \le \chi^2_{n_{(1)}-1} \le \chi^2_{1-\frac{\alpha}{2}, n_{(1)}-1}\right) = 1 - \alpha$ at the significant level α. Hence, the $100(1-\alpha)\%$CI for σ^2 is

$$CI_{\sigma^2} = [l_{\sigma^2}, u_{\sigma^2}] = \left[\frac{(n_{(1)}-1)\hat{\sigma}^2}{\chi^2_{1-\frac{\alpha}{2}, n_{(1)}-1}}, \frac{(n_{(1)}-1)\hat{\sigma}^2}{\chi^2_{\frac{\alpha}{2}, n_{(1)}-1}}\right] \qquad (19)$$

Next, $\hat{\mu} \sim N(\mu, \frac{\sigma^2}{n_{(1)}})$ so that the CI for μ is constructed. By the CLT, the random variable W is

$$W = \frac{\hat{\mu} - \mu}{\sqrt{\frac{(n_{(1)}-1)\hat{\sigma}^2}{n_{(1)}U}}} \qquad (20)$$

where $W \sim N(0,1)$ as $n_{(1)} \to \infty$. At the significant level α, the coverage probability for W is $P\left(W_{\frac{\alpha}{2}} \le W \le W_{1-\frac{\alpha}{2}}\right) = 1 - \alpha$. Therefore, the $100(1-\alpha)\%$CI for μ is

$$CI_\mu = [l_\mu, u_\mu] = \left[\hat{\mu} - W_{1-\frac{\alpha}{2}}\sqrt{\frac{(n_{(1)}-1)\hat{\sigma}^2}{n_{(1)}U}},\ \hat{\mu} + W_{1-\frac{\alpha}{2}}\sqrt{\frac{(n_{(1)}-1)\hat{\sigma}^2}{n_{(1)}U}}\right] \qquad (21)$$

Wilks [45] and Wilson [46] presented a score method to construct CIs for δ. Later, Donner and Zou [15] indicated that Wilson score interval performed well for small to moderate sample sizes. Hence, the $100(1-\alpha)\%$ Wilson interval for δ is

$$CI_{\delta.w} = [l_{\delta.w}, u_{\delta.w}] = \left[\hat{\delta}_w - \left(\frac{Z_{1-\frac{\alpha}{2}}}{n + Z^2_{\frac{\alpha}{2}}}\sqrt{\frac{n_{(0)}n_{(1)}}{n} + \frac{Z^2_{\frac{\alpha}{2}}}{4}}\right),\ \hat{\delta}_w + \left(\frac{Z_{1-\frac{\alpha}{2}}}{n + Z^2_{\frac{\alpha}{2}}}\sqrt{\frac{n_{(0)}n_{(1)}}{n} + \frac{Z^2_{\frac{\alpha}{2}}}{4}}\right)\right] \qquad (22)$$

where $\hat{\delta}_w = (n_{(0)} + Z^2_{\frac{\alpha}{2}}/2)/(n + Z^2_{\frac{\alpha}{2}})$ and $Z \sim N(0,1)$. From Eqs. (16) and (17) proposed by [50] and [51], the MOVER for φ is constructed. Let $\varphi_1 = \ln(1-\delta)$, $\varphi_2 = 2\mu + \sigma^2$ and $\varphi_3 = \ln\left[\exp(\sigma^2) - (1-\delta)\right]$. For focusing on φ_2, the $100(1-\alpha)\%$CI for φ_2 is

$$CI_{\varphi_2} = [l_{\varphi_2}, u_{\varphi_2}] = \left[\hat{\varphi}_2 - \sqrt{4(\hat{\mu} - l_\mu)^2 + (\hat{\sigma}^2 - l_{\sigma^2})^2},\ \hat{\varphi}_2 + \sqrt{4(u_\mu - \hat{\mu})^2 + (u_{\sigma^2} - \hat{\sigma}^2)^2}\right] \qquad (23)$$

where $\hat{\varphi}_2 = 2\hat{\mu} + \hat{\sigma}^2$. For focusing on the next term φ_3, the $100(1-\alpha)\%$CI for φ_3 is

$$CI_{\varphi_3} = [l_{\varphi_3}, u_{\varphi_3}] \qquad (24)$$

where

$$l_{\varphi_3} = \ln\left[\exp(\hat{\sigma}^2) - (1-\hat{\delta}) - \sqrt{A^2 + B^2}\right]$$
$$u_{\varphi_3} = \ln\left[\exp(\hat{\sigma}^2) - (1-\hat{\delta}) + \sqrt{C^2 + D^2}\right];$$

$A = \exp(\hat{\sigma}^2) - \exp(l_{\sigma^2})$, $B = (1 - l_{\delta.w}) - (1 - \hat{\delta})$, $C = \exp(u_{\sigma^2}) - \exp(\hat{\sigma}^2)$ and $D = (1 - \hat{\delta}) - (1 - u_{\delta.w})$. Next, the $100(1-\alpha)\%$CI for $\varphi_2 + \varphi_3$ is obtained as

$$\text{(25)} \quad CI_{\varphi_{2.3}} = [l_{\varphi_{2.3}}, u_{\varphi_{2.3}}] = \left[\hat{\varphi}_{2.3} - \sqrt{(\hat{\varphi}_2 - l_{\varphi_2})^2 + (\hat{\varphi}_3 - l_{\varphi_3})^2}, \; \hat{\varphi}_{2.3} + \sqrt{(u_{\varphi_2} - \hat{\varphi}_2)^2 + (u_{\varphi_3} - \hat{\varphi}_3)^2}\right]$$

where $\hat{\varphi}_{2.3} = \hat{\varphi}_2 + \hat{\varphi}_3$. Finally, the previous steps are merged such that the $100(1-\alpha)\%$MOVER for ψ is

$$\text{(26)} \quad CI_{\psi} = [\exp(L_{\varphi}), \exp(U_{\varphi})]$$

where $\hat{\varphi}_{1.2.3} = \hat{\varphi}_1 + \hat{\varphi}_{2.3}$ and

$$L_{\varphi} = \hat{\varphi}_{1.2.3} - \sqrt{[\hat{\varphi}_1 - \ln(1 - u_{\delta.w})]^2 + [(\hat{\varphi}_2 + \hat{\varphi}_3) - (l_{\varphi_{2.3}})]^2}$$

$$U_{\varphi} = \hat{\varphi}_{1.2.3} + \sqrt{[\ln(1 - l_{\delta.w}) - \hat{\varphi}_1]^2 + [u_{\varphi_{2.3}} - (\hat{\varphi}_2 + \hat{\varphi}_3)]^2}$$

Algorithm 4 shows the computational steps for MOVER interval.

Algorithm 4

1. Generate $U \sim \chi^2_{n_{(1)}-1}$, $W \sim N(0,1)$, and $Z \sim N(0,1)$ are independent
2. Compute CI_{σ^2}, CI_{μ} and $CI_{\delta.w}$
3. Compute CI_{φ_2} and CI_{φ_3} leading to obtain $CI_{\varphi_{2.3}}$
4. Compute CI_{φ}
5. Compute CI_{ψ} being the $(100-\alpha)\%$MOVER for ψ

3. SIMULATION STUDIES

The focus of this article is on approximating the delta-lognormal distribution variance via CI, so it was necessary to assess the proposed CI performances, including

- Coverage probability (CP): the percentage of the interval such that the true parameter falls within it.
- Average length (AL): the average of the widths of simulated intervals.
- Relative average length (RAL): the ratio between the AL of proposed CI and the AL of HPD-J.

By way of comparison, two performance conditions were used to ascertain the performance of the methods: the CI with the CP close or greater than the nominal confidence level $1-\alpha = 0.95$ and the smallest AL (a RAL close to zero and minimal). From previous studies on the delta-lognormal distribution mean, Harvey and Merve [23] showed that HPD-J outperformed MOVER based on Wilson in terms of coverage performance in many situations. Wu and Hsieh [47] indicated that GCI was satisfactory in terms of coverage, expected interval lengths, and relative bias. Hasan and Krishnamoorthy [24] reported that FGCI-HK was a good performance in terms coverage and tail error rate even small sample size. Thus, our numerical results compared the performances of proposed HPD-V and MOVER with the existing HPD-J, GCI and FGCI-HK. Throughout the simulation study, the following considered combinations was designed as

The probabilities of being zero: $\delta = 0.2, 0.5$
The sample sizes: $n = 10, 15, 20, 50, 100$
The mean: $\mu = 0, 1, 2$
The variance: $\sigma^2 = 0.25, 1.00, 2.25$

For each situation, we used 10,000 replications. Note that the all AL of methods are quite widespread in cases the number of positive observations is very small, then these combinations corresponding to $\delta = 0.5$, $n = 10$, $\mu = 0, 1, 2$ and $\sigma^2 = 0.25, 1.00, 2.25$ were excluded from the results of this simulation. From Tables 1–2, the numerical results were obtained in Algorithm 5.

Algorithm 5

1. Generate $X \sim \Delta(\mu, \sigma^2, \delta)$
2. Compute $n_{(0)}$, $n_{(1)}$, $\hat{\mu}$, $\hat{\sigma}^2$ and $\hat{\delta}$
3. Compute the 95% CIs based the different methods:
   3.1 HPD-J and HPD-V from Algorithm 1
   3.2 GCI, FGCI-HK and MOVER from Algorithms 2, 3 and 4, respectively
4. Repeat 1–3 10000 times, the CPs, ALs and RALs of all methods are obtained

The simulation results can be summarized as follows. For the probability of additional zero $\delta = 0.2$, the CP performances of HPD-J, HPD-V, GCI and FGCI-HK were stable and close to the desired confidence levels in almost all cases. However, when considering their lengths, the results indicate that HPD-V achieved the narrowest AL in most of considered cases even large σ^2. Moreover, MOVER provided good CPs and ALs when σ^2 was small for all sample sizes. For $\delta = 0.5$, both of the HPDs provided the best performances in terms of coverage probabilities, but HPD-V provided the shortest AL in some situations. Meanwhile, coverage performances of MOVER performed the same as for the $\delta = 0.2$, whereas its lengths gave wider than HPD-V and HPD-J for moderate ($n = 50$) to large sample sizes ($n = 100$). Furthermore, the CP and AL properties of GCI and FGCI-HK maintained the target similar to HPD-V and HPD-J, although these ALs tended to be quite broad for all sample sizes.

4. AN EMPIRICAL APPLICATION

The motivation for this study is to estimate and predict the variability of daily rainfall in Nan province from histor-

Table 1. The CP and RAL performances of 95% CIs for ψ: $\delta = 0.2$

$\delta = 0.2$			CP					RAL				
n	μ	σ^2	HPD-J	HPD-V	GCI	FGCI-HK	MOVER	HPD-J	HPD-V	GCI	FGCI-HK	MOVER
10	0	0.25	94.96	95.33	95.71	95.74	95.74	(**)	0.987	2.817	2.805	**0.825**
		1.00	95.56	95.42	95.10	95.15	90.95	(**)	1.621	2.01e3	1.79e3	–
		2.25	95.55	95.39	94.80	94.96	89.17	(**)	**0.761**	1.41e7	1.47e7	–
	1	0.25	95.20	95.51	95.73	95.81	95.69	(**)	0.979	2.907	2.857	**0.821**
		1.00	95.63	95.50	95.12	95.14	91.23	(**)	**0.627**	1.17e3	1.18e3	–
		2.25	95.79	95.70	95.21	95.28	89.63	(**)	**0.535**	2.94e7	1.09e7	–
	2	0.25	94.90	95.22	95.65	95.60	95.37	(**)	0.978	2.466	2.512	**0.837**
		1.00	95.21	95.11	94.98	95.01	90.69	(**)	**0.739**	1.27e3	1.57e3	–
		2.25	95.01	94.99	94.49	94.63	88.69	(**)	**0.041**	–	3.20e7	–
20	0	0.25	94.51	94.83	95.22	95.28	94.90	(**)	0.995	1.311	1.318	**0.941**
		1.00	95.32	95.21	94.63	94.63	89.62	(**)	**0.800**	5.853	5.521	–
		2.25	95.57	95.45	95.23	95.23	88.12	(**)	**0.881**	46.707	83.313	–
	1	0.25	94.71	95.07	95.52	95.56	95.22	(**)	0.995	1.311	1.317	**0.939**
		1.00	95.38	95.35	94.65	94.64	89.86	(**)	**0.996**	3.891	3.699	–
		2.25	95.70	95.62	94.94	94.94	88.36	(**)	1.010	56.210	65.968	–
	2	0.25	95.00	95.32	95.65	95.60	95.33	(**)	0.994	1.314	1.320	**0.937**
		1.00	95.51	95.46	94.99	94.97	89.96	(**)	**0.977**	3.584	3.664	–
		2.25	95.13	95.12	94.53	94.50	87.87	(**)	**0.951**	67.694	52.151	–
50	0	0.25	94.38	94.54	94.97	95.04	94.57	–	0.998	1.099	1.100	**0.974**
		1.00	95.72	95.62	95.22	95.29	89.31	(**)	**0.994**	1.396	1.403	–
		2.25	95.34	95.24	94.97	95.03	87.38	(**)	**0.989**	2.435	2.415	–
	1	0.25	94.50	94.77	95.50	95.52	95.08	(**)	0.998	1.098	1.100	**0.975**
		1.00	95.64	95.55	94.89	94.92	88.96	(**)	**0.994**	1.400	1.404	–
		2.25	95.17	95.14	94.34	94.36	86.73	(**)	**0.993**	–	–	–
	2	0.25	94.87	94.92	95.33	95.41	95.06	(**)	0.998	1.098	1.100	**0.973**
		1.00	95.51	95.41	94.83	94.82	88.92	(**)	**0.994**	1.398	1.404	–
		2.25	95.23	95.18	94.60	94.60	86.75	(**)	**0.995**	2.401	2.365	–
100	0	0.25	94.70	94.74	95.19	95.27	94.83	(**)	0.999	1.047	1.048	**0.985**
		1.00	95.38	95.42	94.74	94.73	88.28	(**)	**0.997**	1.182	1.184	–
		2.25	95.80	95.67	95.53	95.51	87.47	(**)	**0.997**	1.495	1.497	–
	1	0.25	94.44	94.61	95.14	95.21	94.76	–	0.999	1.047	1.048	**0.984**
		1.00	95.38	95.29	95.04	95.11	88.51	(**)	**0.997**	1.180	1.182	–
		2.25	95.90	95.86	94.96	94.97	86.84	(**)	**0.994**	1.497	1.501	–
	2	0.25	94.82	94.83	95.11	95.01	94.86	(**)	0.999	1.048	1.049	**0.984**
		1.00	95.65	95.59	94.94	94.91	88.65	(**)	**0.997**	1.182	1.184	–
		2.25	95.32	95.30	94.65	94.68	86.87	(**)	**0.996**	1.496	1.499	–

Note: (**)HPD-J performance satisfies the criteria, and –The method does not follow the given target.

ical rainfall records. Meanwhile, this is one of the most significant factors in agriculture. Farmers are often confronted with the uncertainty of rainfall, which can affect the productivity index of food in Thailand. Thus, information on rainfall dispersion can be utilized to conduct agricultural water management well. The daily rainfall amounts had been recorded and collected by the Thailand Meteorological Department at the following stations: Nan, Nan Agromet, Tha Wang Pha, Thung Chang, Wiang Sa and Chiang Klang. Thailands' climate is tropical with three different seasons: the hot season from March to mid-May, the rainy season from mid-May to October, and the cool season from November to February. Particularly, the rainy season is caused by the southwest monsoon. In this study, we are interested in the daily rainfall records during June-September 2018, which is a rice farming period during the middle of the rainy season. Importantly, this can be predicted occurring the drought or flooding during the same period in the next year. Figure 3 shows histogram plot of 610 records including 258 zero and 352 nonzero observations, which are also detailed in Table 3. A Q-Q plot of the logarithms of the positive rainfall amounts is shown in Figure 4. The AIC results show the fitting distribution of positive rainfall observations in Table 4. The zero values were also included, so this fits with the delta-lognormal model.

To approximate and compare the dispersion of daily rainfall in Nan province, the estimated variance of this dataset was $\hat{\psi} = 1927.328$, where $\hat{\mu} = 1.667$, $\hat{\sigma}^2 = 2.416$ and $\hat{\delta} = 0.423$. The 95% CIs for the daily rainfall variance ψ were calculated, as shown in Table 5. These results indicate

Table 2. The CP and RAL performances of 95% CIs for ψ: $\delta = 0.5$

$\delta = 0.5$			CP					RAL				
n	μ	σ^2	HPD-J	HPD-V	GCI	FGCI-HK	MOVER	HPD-J	HPD-V	GCI	FGCI-HK	MOVER
15	0	0.25	95.97	95.94	95.33	95.21	96.40	(**)	0.993	2.499	2.476	**0.901**
		1.00	95.99	95.87	94.40	94.40	90.89	(**)	1.489	1.20e4	–	–
		2.25	95.61	95.60	94.49	94.49	88.49	(**)	**0.544**	7.91e6	–	–
	1	0.25	96.21	96.18	95.58	95.48	96.63	(**)	0.995	2.340	2.354	**0.913**
		1.00	95.70	95.74	94.36	94.32	91.06	(**)	1.212	7.11e3	–	–
		2.25	95.61	95.54	94.50	94.50	89.19	(**)	1.441	2.40e7	4.80e8	–
	2	0.25	95.83	95.81	95.05	95.06	96.38	(**)	0.994	2.451	2.456	**0.897**
		1.00	96.05	96.05	94.65	94.59	91.15	(**)	1.650	7.48e3	4.99e3	–
		2.25	96.00	95.98	94.48	94.49	88.88	(**)	**0.965**	–	–	–
20	0	0.25	95.65	95.60	95.12	95.07	96.24	(**)	0.996	1.605	1.602	**0.956**
		1.00	95.77	95.73	94.70	94.64	90.77	(**)	**0.986**	24.671	22.217	–
		2.25	95.31	95.26	94.59	94.54	88.49	(**)	1.958	2.91e4	2.62e4	–
	1	0.25	95.70	95.68	95.08	95.05	96.30	(**)	0.997	1.640	1.640	**0.948**
		1.00	95.67	95.67	94.57	94.61	90.37	(**)	1.080	98.847	163.637	–
		2.25	95.66	95.65	95.01	95.08	88.72	(**)	**0.431**	1.60e4	4.50e3	–
	2	0.25	95.75	95.75	95.02	94.91	96.21	(**)	0.996	1.644	1.653	**0.952**
		1.00	95.92	95.81	95.13	95.11	91.15	(**)	1.128	115.894	63.249	–
		2.25	95.61	95.60	94.06	94.06	87.88	(**)	1.432	–	–	–
50	0	0.25	96.12	96.11	94.89	94.76	96.26	(**)	**0.999**	1.145	1.145	1.024
		1.00	96.19	96.14	94.82	94.77	89.81	(**)	**0.999**	1.749	1.755	–
		2.25	95.92	95.95	94.95	94.92	87.99	(**)	1.081	7.238	6.548	–
	1	0.25	95.83	95.87	94.90	94.87	96.23	(**)	**0.999**	1.144	1.145	1.024
		1.00	96.23	96.23	94.66	94.69	89.78	(**)	**1.001**	1.753	1.749	–
		2.25	95.69	95.67	94.79	94.82	87.84	(**)	**0.986**	5.010	4.842	–
	2	0.25	95.57	95.60	94.92	94.91	96.23	(**)	**0.999**	1.142	1.143	1.026
		1.00	95.84	95.87	94.47	94.58	89.72	(**)	**1.000**	–	1.755	–
		2.25	95.83	95.83	94.90	94.91	87.46	(**)	1.072	6.970	6.809	–
100	0	0.25	95.91	95.76	95.06	95.12	96.45	(**)	**1.000**	1.066	1.067	1.045
		1.00	96.41	96.41	94.99	95.03	89.91	(**)	**0.999**	1.282	1.283	–
		2.25	95.28	95.26	94.61	94.52	87.14	(**)	**0.995**	1.898	1.893	–
	1	0.25	95.63	95.71	94.94	94.98	96.36	(**)	**0.999**	1.066	1.067	1.044
		1.00	96.13	96.06	94.80	94.85	89.66	(**)	**1.000**	1.281	1.282	–
		2.25	95.97	95.91	95.38	95.41	87.44	(**)	**0.999**	1.886	1.872	–
	2	0.25	95.75	95.66	95.13	95.12	96.32	(**)	**0.999**	1.066	1.067	1.045
		1.00	96.36	96.36	95.11	95.13	89.99	(**)	**1.000**	1.283	1.284	–
		2.25	95.63	95.65	94.64	94.74	86.82	(**)	**1.000**	1.912	1.904	–

Note: (**)HPD-J performance satisfies the criteria, and –The method does not follow the given target.

that the rainfall variation largely oscillated during rice farming (June-September 2018), meanwhile it might be affected to occur the flood and flashflood in the next year. Likewise, it is in line with the occurrence's natural disasters (flooding and landslide) during July 2018 in Nan area reported by [22]. The simulation study results were clearly confirmed by these results in that HPD-V, HPD-J and MOVER had relatively smaller ALs compared with the other interval methods. Conversely, GCI and FGCI-HK had quite broad lengths.

5. DISCUSSION AND CONCLUSIONS

The calculation and comparison of all of the intervals of interest are presented in this article. The proposed method HPD-V and MOVER for the CIs of delta-lognormal distribution variance were compared with the existing methods: HPD-J, GCI, and FGCI-HK. Our findings show that the proposed methods were effective in the following situations. For $\delta = 0.2$ (Table 1 and Figure 1), it can be pointed toward there being two potential methods in various cases: MOVER for small σ^2 (all sample sizes) and HPD-V for large σ^2 (almost all sample sizes). For $\delta = 0.5$ (Table 2 and Figure 2), both of HPD-V and HPD-J are the recommended methods in almost all situations because their performances are not different, and MOVER is considered as an alternative method (similar to $\delta = 0.2$) when the moderate to large sample sizes are excluded.

As mentioned previously, MOVER outperformed the others because the CIs for δ based on Wilson worked well [15]. Moreover, many researchers have investigated and used Wilson interval in their studies e.g. [6, 23, 24, 27, 51]. Notice

Table 3. Data on the Nan's daily rainfall amounts measured in millimeter

Stations	Times	Daily natural rainfall records															
Nan	Jun	1.5	0	0	0	0.7	0	6.7	0	14.5	6.3	1.4	13.5	1.1	7.7	0	0
		0	1.3	0.7	0	0	15.2	14.1	0.5	13.1	51.4	10.5	1.4	0	0		
	Jul	0	0	0	0	0	0	0	7.1	0.2	32.6	18.9	2.4	0	0	0	21.2
		15.1	12.8	16.5	11.4	56.1	0	40.8	20.0	2.6	0.4	1.8	17.5	6.9	10.8	0.9	
	Aug	9.1	0	0	0	0	0	0	0.5	0	0	0	0.6	0	0.4	0	25.4
		71.5	1.2	0	24.3	29.5	27.6	3.4	3.2	5.8	0.3	8.7	1.3	1.4	0	0	
	Sep	0	62.7	6.3	0.8	0	0	0	7.4	0.7	0	0	0.2	0	0	0	0
		15.3	1.0	0.8	0	0	0.4	0	0	0	0	0	0	0	0		
Nan Agromet	June	0	0	0	0	0	0	0	0	13.0	2.3	5.6	8.0	0.5	0.4	0	0
		0	0.8	19.0	0	0	0	16.8	6.0	24.7	23.7	34.8	2.4	0	0		
	Jul	0	0	0	0	0	0	0	5.3	2.7	53.3	26.6	1.8	0	0	0	18.6
		15.9	2.7	12.4	9.2	62.8	0	32.8	24.0	1.0	1.2	2.0	26.6	9.5	5.0	0	
	Aug	0	0	0	0	0	0	0	0	15.6	0	0	0.6	0	0.4	1.5	35.7
		87.0	1	2.4	10	7.7	6.4	14.8	2.3	2.2	0	7.1	0	2.9	0	4.1	
	Sep	0	47.6	6.5	3.4	0	2.0	0	8.0	0	0	0	0.5	0	0	0	0
		118.5	2.4	0.5	0	0	0	0	0	0	0	0	0	0.2	0		
Tha Wang Pha	Jun	0.2	0	10.2	0	3.4	3.4	3.2	0	13.2	2.7	15.1	2.4	6.3	0	1.8	0
		0.2	22.9	20.8	0.9	2.9	0.2	48.8	0.8	28.0	14	13.2	10.9	4.6	0		
	Jul	10.4	0	0.4	0	10.7	0	3.6	1.5	0.9	3.6	68.9	14.7	0	0	9.0	19.3
		30.9	4.8	35.0	44.7	28.6	0	6.7	51.0	4.9	17.1	12.4	21.6	7	9.7	3.9	
	Aug	1.0	8.7	59	0	0.7	0	0	0	0	0	0	2.1	10.6	1.3	0	126.1
		126.1	3.5	0	2.4	4.4	9.8	2.6	4.5	0	0.8	0.1	0.1	4.8	1.8	0.3	
	Sep	1.0	28	34.1	0.8	0	0	0	38.8	0.8	0	0	1.5	0	5.1	1.5	0
		65	12.6	0.4	0	2.0	0	0	0	0	0	0	0	0	0.2		
Thung Chang	Jun	4.5	0	0.4	4.4	6	0.3	1.7	0		3.7	5.6	4.6	5.0	8.0	26.4	4.5
		16.8	73.8	12.4	4.9	0	6.2	94.6	3.8	100	15.5	8.5	3.2	0.6	0		
	Jul	2.2	7.9	0.9	0.2	0	0	0	5.5	0.1	9.0	1.0	11.4	0	0	0	11.4
		12.9	0.1	46.2	45.6	34.3	0	4.4	44.6	10.3	29.2	12.8	3.9	16.2	1.4	0	
	Aug	0.1	0	3.2	0	0	0	5.4	1.6	19.3	0	4.2	0	9.8	0.4	0	8
		106.1	22.5	0.5	25.6	4.8	29.6	9.1	3.6	1.3	33.8	5.4	0.7	0	0.5	3.9	
	Sep	0	18.9	24.2	0	0	0	0	53.1	20.1	0	0	28.6	1.4	0	0	0
		42.0	3.7	2.2	0	4.5	0	0	0	0	0	0	37.1	34.9	6.4		
Wiang Sa	Jun	0	0	0	0	43.4		1.5	0	8.4	4.6	2.5	11.6	0.8	0	15.8	0
		0	2.1	0.8	0	0	0	10.6	0	10	23.1	49.8	1.3	0	0		
	Jul	0	0	0	0	0	0	0	2.4	0	47.4	15.5	0	0	0	0	42.3
		36.3	9.1	13.3	9.2	33.8	0	12.5	15.5	2.3	0	0	24.2	1.9	13.5	2.1	
	Aug	1.3	0	0	0	0	4.6	0	4.3	18.3	0	0	10.7	0.6	0	0	11.7
		36.1	0	0	1.6	2.3	25.6	2.3	14.0	4.2	23.3	17.4	1.4	2.0	0	0	
	Sep	0	27.5	2.9	2.9	1.8	0	1.9	3.3	0	0	0	8.2	8.2	0	0	0
		4.8	0	0	0	0	0	1.2	0	0	0	0	0	0	0		
Chiang Klang	Jun	0	0	0	12.0	0.9	26.9	2.1	0	0	2.1	9.4	5.3	7.1	2.0	8.9	1.2
		1.5	42.1	6.4	0	1.1	5.1	16.6	0	24.4	20.1	23.5	14.1	0	0		
	Jul	2.5	6.3	8.8	0	0	0	5.9	9.0	3.8	17	25.5	50.0	0	0	5.8	28.1
		23.3	2.4	73.3	35.0	38.2	0	2.4	21.9	15.3	25.9	16.0	32.3	5.2	3.3	1.1	
	Aug	0	0	6.4	0	0	0	0	2.0	7.0	0	0	1.9	8.4	12.1	2.7	2.1
		153.0	3.7	0	31.6	2	36.1	0	4.6	1.6	14	1.3	0	4.7	0	3.3	
	Sep	0	20.4	18.0	0	0	0	0	24.2	1.4	0	0	5.1	5.1	0	0	0
		75.8	7.1	2.2	0	0	0	0	0	0	0	0	0.3	2.8	9.7		

Source: Meteorological Department, Thailand; url: https://www.tmd.go.th/services/weekly_report.php.

that for delta-lognormal distribution variance, MOVER relied on the CIs for μ, σ^2 and δ, which were developed to perform well. However, the CIs for σ^2 were narrow when σ^2 is very large, which affects the CP. This implies that the coverage by MOVER is difficult to control for the target in situations with large σ^2. Subsequently, the MOVER performance was only dependent on σ^2. The FGCI-HK cannot work well for delta-lognomal variance compared with HPD-V. The numerical results showed that the CP and AL performances of HPD-J [23] and HPD-V were not different

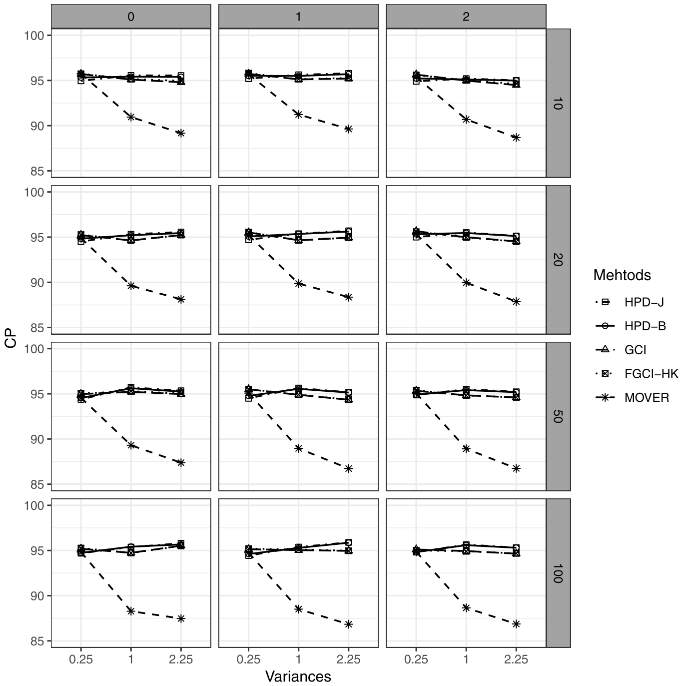

Figure 1. The CPs of 95%*CIs for* ψ*:* $\delta = 0.2$*.*

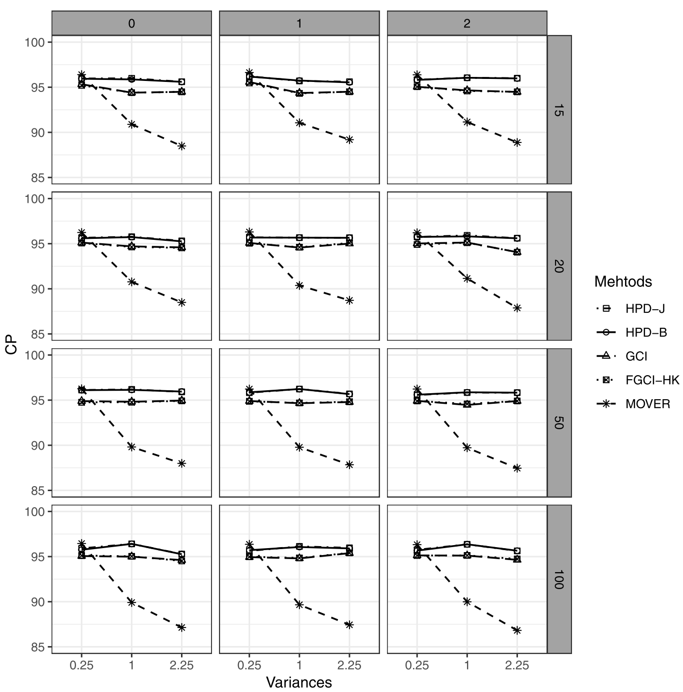

Figure 2. The CPs of 95%*CIs for* ψ*:* $\delta = 0.5$*.*

for $\delta = 0.2, 0.5$ and large sample sizes (e.g., $n \geq 50$) especially. For $\delta = 0.2$, small sample sizes e.g., $n \leq 20$ and large variance, HPD-V was better than HPD-J for some cases. However, the HPDs trended to result in under coverage in

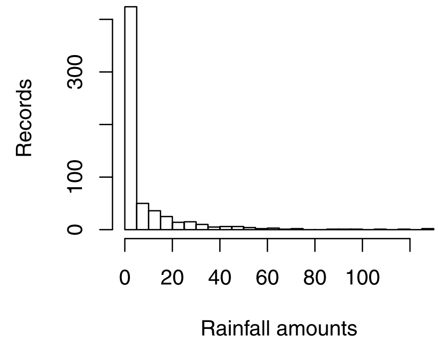

Figure 3. Histogram plot of rainfall records in Nan province, northern Thailand.

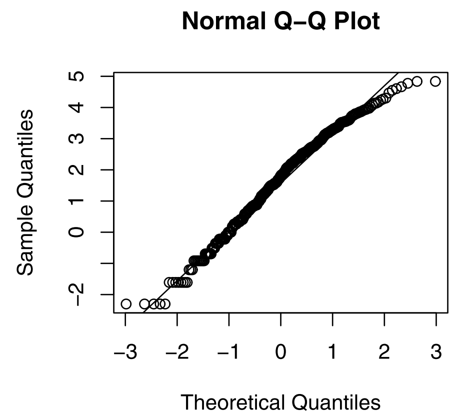

Figure 4. A Q-Q plot of the logarithm of positive rainfall amounts.

Table 4. The AIC results for checking the fitting distribution of nonzero rainfall records in Nan province, northern Thailand

Distributions	AIC
Cauchy	2831.567
Exponential	2554.028
Gamma	2499.899
Logistic	2983.933
Lognormal	**2486.244**
Normal	3109.472
t-distriution	2833.019

Table 5. 95%*CIs for the Nan's natural rainfall variation during June-September 2018*

Methods	CIs		Lengths
	Lower limit	Upper limit	
HPD-J	735.567	4090.961	3355.394
HPD-V	766.301	4095.341	3329.040
GCI	907.762	4662.233	3754.471
FGCI-HK	918.713	4709.598	3790.885
MOVER	1092.664	3730.706	2638.042

many cases, then their ALs were not compared. The reason is that the Jeffreys' prior of δ has the beta distribution with the parameters $c = 1/2$ and $d = 1/2$ leading to be a conjugated prior density, which means that its posterior is a same family with its prior. Meanwhile, the vague prior of δ is called the conjugate family for a binomial distribution, although the HPD-V performance based on the conjugated beta prior was better than that of HPD-J in some situations, even a sample size was small. This result was consistent with [9]. The HPD-J and HPD-V are based on different theories concerning the prior of δ based on Jeffreys (proportional to the square root of the determinant of the Fisher information matrix) and beta prior (a conjugated prior), respectively, while both posteriors of δ are derived as beta distributions with different parameters.

Overall, there are two problems concerning the extensive length. First, the sample size of non-zero observations was very small and insufficient for collecting information about the population, and this restriction directly affected the AL. Thus, Fletcher [19], Wu and Hsieh [47] kept the cases of positive values to only four in their studies. Next, our results show that the performance of some of the interval estimates gave wide ALs for a large σ^2 and small sample size, a limitation that was similarly discovered by [23]. These problems could be the goal for future research.

APPENDIX A

Proof of Theorem 1. The posterior density of ψ is obtained from the combination between the mentioned prior $P(\psi)_V$ and the likelihood $P(x|\psi)$, expressed as

$$
\begin{aligned}
&P(\psi|x)\\
&=\binom{n}{n_{(0)}}\frac{\delta^{n_{(0)}+c-1}(1-\delta)^{n_{(1)}+d-1}}{B(c,d)}(\sigma)^{-2}\prod_{i=1}^{n_{(1)}}(2\pi\sigma^2)^{-\frac{1}{2}}\\
&\quad\times\exp\left\{-\frac{1}{2\sigma^2}(\ln x_i-\mu)^2\right\}\\
&=\binom{n}{n_{(0)}}\frac{\delta^{n_{(0)}+c-1}(1-\delta)^{n_{(1)}+d-1}}{B(c,d)}(2\pi)^{-\frac{n_{(1)}}{2}}(\sigma)^{-(n_{(1)}+2)}\\
&\quad\times\exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n_{(1)}}(\ln x_i-\mu)^2\right\}\\
&=\binom{n}{n_{(0)}}\frac{\delta^{n_{(0)}+c-1}(1-\delta)^{n_{(1)}+d-1}}{B(c,d)}(2\pi)^{-\frac{n_{(1)}}{2}}(\sigma^2)^{-\frac{n_{(1)}+2}{2}}\\
&\quad\times\exp\left\{-\frac{1}{2\sigma^2}\left[(n_{(1)}-1)\hat{\sigma}^2+n_{(1)}(\hat{\mu}-\mu)^2\right]\right\}
\end{aligned}
\tag{27}
$$

where $\sum_{i=1}^{n_{(1)}}(\ln x_i-\mu)^2 = \sum_{i=1}^{n_{(1)}}\left[(\ln x_i-\hat{\mu})+(\hat{\mu}-\mu)\right]^2 = (n_{(1)}-1)\hat{\sigma}^2 + n_{(1)}(\hat{\mu}-\mu)^2$, then the integration $\int\exp\left\{-\frac{n_{(1)}}{2\sigma^2}(\hat{\mu}-\mu)^2\right\}d\mu = \left(\frac{2\pi\sigma^2}{n_{(1)}}\right)^{1/2}$. From Eq. (27), the posterior of σ^2 is

$$
\begin{aligned}
P(\sigma^2|x) &\propto (\sigma^2)^{-\left(\frac{n_{(1)}+2}{2}\right)}\int\exp\left\{-\frac{1}{2\sigma^2}(n_{(1)}-1)\hat{\sigma}^2\right.\\
&\qquad\left.+n_{(1)}(\hat{\mu}-\mu)^2\right\}d\mu\\
&\propto(\sigma^2)^{-\left(\frac{n_{(1)}+2}{2}\right)}\exp\left\{-\frac{(n_{(1)}-1)\hat{\sigma}^2}{2\sigma^2}\right\}\\
&\quad\times\int\exp\left\{-\frac{n_{(1)}}{2\sigma^2}(\hat{\mu}-\mu)^2\right\}d\mu\\
&\propto(\sigma^2)^{-\left(\frac{n_{(1)}-1}{2}+1\right)}\exp\left\{-\frac{(n_{(1)}-1)\hat{\sigma}^2}{2\sigma^2}\right\}
\end{aligned}
\tag{28}
$$

which is a inverted-gamma density with parameter $a = (n_{(1)}-1)/2$ and $b = (n_{(1)}-1)\hat{\sigma}^2/2$, denoted as $\sigma^2 \sim IG(a,b)$. Given σ^2 and data, the posterior of μ can be expressed as

$$P(\mu|\sigma^2,x)\propto\exp\left\{-\frac{n_{(1)}}{2\sigma^2}(\mu-\hat{\mu})^2\right\}\tag{29}$$

which is normal distribution, denoted as $\mu|\sigma^2,x \sim N(\hat{\mu},\sigma^2/n_{(1)})$. The posterior of δ is

$$P(\delta|x)\propto\delta^{n_{(0)}+c-1}(1-\delta)^{n_{(1)}+d-1}\tag{30}$$

Then, $\delta|x \sim Beta(n_{(0)}+c, n_{(1)}+d)$. The posterior $P(\sigma^2|x)$, $P(\mu|\sigma^2,x)$ and $P(\delta|x)$ are independently distributed, then the posterior $P(\psi)_V$ is obtained based on vague prior. □

ACKNOWLEDGEMENTS

The authors are grateful the associate editor and the referees for their valuable comments and suggestion. The first author would like to thank Uttaradit Rajabhat University for their supports. The third author is appreciated the grant number KMUTNB-63-KNOW-032 from King Mongkut's University of Technology North Bangkok.

Received 24 January 2020

REFERENCES

[1] Aitchison, J. (1955). On the distribution of a positive random variable having a discrete probability mass at the origin. *Journal of the American Statistical Association* **50** 901–908. MR0071685

[2] Aitchison, J. and Brown, J. A. (1963). *The Lognormal Distribution: With Special Reference to Its Uses in Economics*. Cambridge University Press, Cambridge. MR0084966

[3] Bangkok Post (2017). Thai organic foods have healthy growth potential. Asia Focus. SCB Economic Intelligence Center. Bangkok Post. Available at: https://www.bangkokpost.com/business/news/1193633/thai-organic-foods-have-healthy-growth-potential (accessed 12 July 2018).

[4] Bebu, I. and Mathew, T. (2008). Comparing the means and variances of a bivariate log-normal distribution. *Statistics in Medicine* **27** 2684–2696. MR2440059

[5] Box, G. E. P. and Tiao, G. C. (1973). *Bayesian Inference in Statistical Analysis*. Wiley, New York. MR0418321

[6] Brown, L. D., Tony Cai, T. T. and DasGupta, A. (2001). Interval estimation for a binomial proportion. *Statistical Science* **16** 101–117. MR1861069

[7] Buntao, N., Niwitpong, S.-A. and Kreinovich, V. (2012). Estimating statistical characteristics of lognormal and delta-lognormal distributions under Interval uncertainty: algorithms and computational complexity. Departmental Technical Report, Department of Computer Science, University of Texas at El Paso, pages 1–38.

[8] Burdick, R. K. and Graybill, F. A. (1984). Confidence intervals on linear combinations of variance components in the unbalanced one-way classification. *Technometrics* **26** 131–136. MR1192783

[9] Cameron, E. (2011). On the estimation of confidence intervals for binomial population proportions in astronomy: the simplicity and superiority of the Bayesian approach. *Publications of the Astronomical Society of Australia* **28** 128–139.

[10] Casella, G. and Berger, R. L. (2002). *Statistical Inference (2nd ed.)*. Thomson Learning, California.

[11] Climatological, G. (2015). The Climate of Thailand. Meteorological Development Bureau Meteorological Department. Available at https://www.tmd.go.th/en/downloads.php (accessed 14 March 2019).

[12] Cohen, A. (1972). Improved confidence intervals for the variance of a normal distribution. *Journal of the American Statistical Association* **67** 382–387. MR0312636

[13] Cojbasic, V. and Loncar, D. (2011). One-sided confidence intervals for population variances of skewed distributions. *Journal of Statistical Planning and Inference* **141** 1667–1672. MR2763198

[14] DasGupta, A. (2008). *Asymptotic Theory of Statistics and Probability*. Springer, New York. MR2664452

[15] Donner, A. and Zou, G. Y. (2011). Estimating simultaneous confidence intervals for multiple contrasts of proportions by the method of variance estimates recovery. *Statistics in Biopharmaceutical Research* **3** 320–335.

[16] Duangdai, E. and Likasiri, C. (2017). Rainfall model investigation and scenario analyses of the effect of government reforestation policy on seasonal rainfalls: A case study from Northern Thailand. *Atmospheric Research* **185** 1–12.

[17] Fisher, R. A. (1930). Inverse probability. *Mathematical Proceedings of the Cambridge Philosophical Society* **26** 528–535.

[18] Fisher, R. A. (1935). The fiducial argument in statistical inference. *Annals of Eugenics* **6** 391–398.

[19] Fletcher, D. (2008). Confidence intervals for the mean of the delta-lognormal distribution. *Environmental and Ecological Statistics* **15** 175–189. MR2399079

[20] Hannig, J. (2009). On generalized fiducial inference. *Statistica Sinica* **19** 491–544. MR2514173

[21] Hannig, J., Iyer, H. and Patterson, P. (2006). Fiducial generalized confidence intervals. *Journal of the American Statistical Association* **101** 254–269. MR2268043

[22] Hariraksapitak, P., Sriring, O. and Navaratnam, S. (2018). Landslides in northern Thailand kill seven; houses, roads damaged. Reuters. Available at: https://www.reuters.com/article/us-thailand-landslide/ (accessed 19 October 2019).

[23] Harvey, J. and Merwe, A. J. (2012). Bayesian confidence intervals for means and variances of lognormal and bivariate lognormal distributions. *Journal of Statistical Planning and Inference* **142** 1294–1309. MR2891482

[24] Hasan, M. S. and Krishnamoorthy, K. (2018). Confidence intervals for the mean and a percentile based on zero-inflated lognormal data. Journal of Statistical Computation and Simulation **88** 1499–1514. MR3779162

[25] Krishnamoorthy, K. and Mathew, T. (2003). Inferences on the means of lognormal distributions using generalized p-values and generalized confidence intervals. *Journal of Statistical Planning and Inference* **115** 103–121. MR1972942

[26] Krishnamoorthy, K., Mathew, T. and Ramachandran, G. (2006). Generalized P-values and confidence intervals: a novel approach for analyzing lognormally distributed exposure data. *Journal of Occupational and Environmental Hygiene* **3** 642–650.

[27] Li, X., Zhou, X. and Tian, L. (2013). Interval estimation for the mean of lognormal data with excess zeros. *Statistics & Probability Letters* **83** 2447–2453. MR3144024

[28] Marks, D. (2011). Climate change and Thailand: impact and response. *Contemporary Southeast Asia* **33** 229–258.

[29] Mathew, T. and Webb, D. W. (2005). Generalized p values and confidence intervals for variance components: applications to army test and evaluation. *Technometrics* **47** 312–322. MR2164703

[30] Nicely, R. and Counselor, A. (2018). Grain and feed update May 2018. Technical report. Available at https://www.fas.usda.gov/data/thailand-grain-and-feed-update-30 (accessed 5 March 2019).

[31] Niwitpong, S. (2017). Generalized confidence intervals for function of variances of lognormal distribution. *Advances and Applications in Statistics* **51** 151–163.

[32] Owen, W. J. and DeRouen, T. A. (1980). Estimation of the mean for lognormal data containing zeroes and left- censored values, with applications to the measurement of worker exposure to air contaminants. *Biometrics* 36 707–719.

[33] Pennington, M. (1983). Efficient estimators of abundance, for fish and plankton surveys. *Biometrics 39* 281–286.

[34] Poaponsakorn, N. and Meethom, P. (2013). Impact of the 2011 floods, and flood management in Thailand. ERIA Discussion Paper Series, pages 1–72.

[35] Smith, S. J. (1988). Evaluating the efficiency of the Δ-distribution mean estimator. *Biometrics* **44** 485–493.

[36] Smith, S. J. (1990). Use of statistical models for the estimation of abundance from groundfish trawl survey data. *Canadian Journal of Fisheries and Aquatic Sciences* **47** 894–903.

[37] Suppaisarn, C. (2011). Medium and heavy flood management in Chao Phraya river basin following the Royal initiative. Technical report.

[38] Tian, L. (2005). Inferences on the mean of zero-inflated lognormal data: the generalized variable approach. *Statistics in Medicine* **24** 3223–3232. MR2209054

[39] Tian, L. and Wu, J. (2006). Confidence intervals for the mean of lognormal data with excess zeros. *Biometrical Journal* **48** 149–156. MR2225140

[40] Tierney, W. M., Overhage, J. M., Murray, M. D., Harris, L. E., Zhou, X. H., Eckert, G. J., Smith, F. E., Nienaber, N., McDonald, C. J. and Wolinsky, F. D. (2003). Effects of computerized guidelines for managing heart disease in primary care. *Journal of General Internal Medicine* **18** 967–976.

[41] Zhou, X.-H., Eckert, G. J., Smith, F. E., Nienaber, N., McDonald, C. J. and Wolinsky, F. D. (2003). Effects of computerized guidelines for managing heart disease in primary care. *Journal of General Internal Medicine* **18** 967–976.

[42] Wandler, D. V. and Hannig, J. (2011). Fiducial inference on the largest mean of a multivariate normal distribution. *Journal of Multivariate Analysis* **102** 87–104. MR2729422

[43] Wandler, D. V. and Hannig, J. (2012). Generalized fiducial confidence intervals for extremes. *Extremes* **15** 67–87. MR2891310

[44] Weerahandi, S. (1993). Generalized confidence intervals. *Journal of the American Statistical Association* **88** 899–905. MR1242940

[45] Wilks, S. S. (1938). Shortest average confidence intervals from large samples. *The Annals of Mathematical Statistics* **9** 166–175.

[46] Wilson, E. B. (1927). Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association* **22** 209–212.

[47] Wu, W.-H., and Hsieh, H.-N. (2014). Generalized confidence interval estimation for the mean of delta-lognormal distribution: an application to New Zealand trawl survey data. *Journal of Applied Statistics* **41** 1471–1485. MR3292027

[48] Zhang, B., Zhang, L., Guo, H., Leinenkugel, P., Zhou, Y., Li,

L. and SHEN, Q. (2014). Drought impact on vegetation productivity in the Lower Mekong Basin. *International Journal of Remote Sensing* **35** 2835–2856.

[49] ZHOU, X. H. and TU, W. (2000). Confidence intervals for the mean of diagnostic test charge data containing zeros. *Biometrics* **56** 1118–1125. MR1815591

[50] ZOU, G. Y. and DONNER, A. (2008). Construction of confidence limits about effect measures: A general approach. *Statistics in Medicine* **27** 1693–1702. MR2420335

[51] ZOU, G. Y., HUANG, W. and ZHANG, X. (2009). A note on confidence interval estimation for a linear function of binomial proportions. *Computational Statistics & Data Analysis* **53** 1080–1085. MR2657072

Patcharee Maneerat
Department of Mathematics
Uttaradit Rajabhat University
Uttaradit
Thailand
E-mail address: m.patcharee@uru.ac.th

Suparat Niwitpong
Department of Applied Statistics
King Mongkut's University of Technology North Bangkok
Bangkok
Thailand
E-mail address: suparat.n@sci.kmutnb.ac.th

Sa-Aat Niwitpong
Department of Applied Statistics
King Mongkut's University of Technology North Bangkok
Bangkok
Thailand
E-mail address: sa-aat.n@sci.kmutnb.ac.th