Approximate Statistical Limits for a Gamma Distribution

PIAO CHEN and ZHI-SHENG YE

Department of Industrial & Systems Engineering
National University of Singapore, Singapore 117576

This study develops methods for constructing some important statistical limits of a gamma distribution. First, we construct upper prediction limits and tolerance limits for a gamma distribution. In addition, upper prediction limits for at least p of m measurements from a gamma distribution at each of r locations are constructed. This problem often arises in the field of environmental monitoring, as well as quality control. For each problem discussed in this study, the inferential procedure based on the generalized fiducial method is outlined and a simulation is conducted to assess its performance. Real data are used to demonstrate the proposed methods. In addition to the prediction and tolerance limits, our proposed methods can also be applied to the stress-strength problem involving two independent gamma random variables. Our investigation shows that the proposed methods perform uniformly well and they outperform existing methods.

Key Words: Coverage Probability; Fiducial Inference; Prediction Limits; Stress-Strength; Tolerance Limits.

1. Introduction

THE DETECTION of possible contamination based on a background sample of historical measurements is a fundamental problem in environmental monitoring (Bhaumik and Gibbons (2006)). As an example, consider the groundwater detection-monitoring programs at some waste disposal facilities. In some cases, the background sample may consist of repeated measurements from a single monitoring well. The next one or more future measurements from the same well will be compared with the upper prediction limit (UPL) computed from these historical measurements to check whether there is possible contamination. The upper prediction limit provides a $1-\alpha$ confidence of not being exceeded by the future measurements with a known number. For example, Gibbons et al. (2009) constructed a 95% prediction limit for the next single measurement of total organic carbon levels based on $n = 8$ historical measurements from the same well. In other cases, however, the background sample may consist of repeated measurements from some wells located upgradient of the waste-disposal facilities while the monitoring wells are located downgradient of the facilities. In such cases, we would expect that the contamination may impact all the monitoring wells and hence we want to determine whether this situation arises. In practice, a series of m future measurements from each of r monitoring wells are often used to compare with the prediction limit based on n historical measurements (Bhaumik and Gibbons (2006), Gibbons et al. (2009)). One common requirement in the groundwater monitoring program is that at least $p = 1$ of $m = 2$ future measurements at each of r wells are below the UPL with $1-\alpha$ confidence (Bhaumik and Gibbons (2006)). Alternative plans include $(m,p) = (3,1)$ and $(m,p) = (3,2)$. In addition to the environmental monitoring problem, this kind of UPL, which is called simultaneous UPL in Bhaumik and Gibbons (2006), is also found to be useful in many other important areas. For example, simultaneous prediction limits are used to determine whether a process is in control in some quality-control problems. In the field of molecular genetics, simultaneous evaluation of the expression levels of thousands of genes in one or more tissue samples is required (Chee et al. (1996), Gibbons et al. (2005)).

By using the prediction limits, we are able to de-

Mr. Chen is PhD Candidate in the Department of Industrial & Systems Engineering. His email address is cp@u.nus.edu.

Dr. Ye is an Assistant Professor in the Department of Industrial & Systems Engineering. He is the corresponding author. His email is yez@nus.edu.sg.

tect possible contamination in certain cases. In fact, the prediction limits are useful in the process of detection when we focus on a single or a small number of future measurements. In some cases, however, the number of future measurements is either large or unknown. Hence, it may be unreasonable to attempt to include all the measurements with $1 - \alpha$ confidence. In such cases, we may turn our attention from including a finite number of measurements to including a proportion β of all the measurements with $1 - \alpha$ confidence. An appropriate statistical limit for this application is known as a tolerance limit, which has also received a lot of attention in the environmental-monitoring problem. For example, the upper tolerance limits are often used to determine whether potential contamination has impacted the environment adversely (Bhaumik and Gibbons (2006)). In addition to the environmental-monitoring problem, the tolerance intervals are also widely used for the purpose of controlling product quality. See Fernández (2010) and Lee and Liao (2012) for examples.

As introduced before, because a lot of contamination like volatile organic substances may damage the environment and pose serious health risks, it is critically important to determine the evidence of its release. In other words, it is important to construct prediction and tolerance limits based on historical measurements. In the literature for the environmental-monitoring problem, the normal distribution for the monitored constituents, e.g., total organic carbon, is often assumed (Bhaumik and Gibbons (2006)). With the normal distribution assumption, there have been well-developed methods for both the prediction limits and tolerance limits (Gibbons et al. (2009)). However, according to Bhaumik and Gibbons (2006), the normal distribution fails to provide a good fit to many environmental data. Empirical distributions of some constituents are far from Gaussian (Gibbons et al. (2009)). Gibbons et al. (2009) found that the gamma distribution could characterize these environmental data quite well. This argument is also supported by Krishnamoorthy et al. (2008). Aside from its application in environmental studies, applications of the gamma distribution can be found in many other areas. For example, the gamma distribution offers a good fit to many lifetime data sets (e.g., Fan and Yu (2013), Chen and Ye (2015)). It also plays an important role in call-center queuing models (Avramidis et al. (2004)). In addition, the gamma distribution is used extensively to model the amount of daily rainfall (e.g., Stephenson et al. (1999)) and hydrological data (Aksoy (2000)). More recently, it is found to be an appropriate model for the signal-to-noise ratio of wireless channels (Atapattu et al. (2011)).

There are two methods to construct the prediction and tolerance limits for a gamma distribution. The first method is based on the assumption that one (or some) parameter(s) of the gamma distribution is known or that the true parameter values can be substituted by the estimators (Bain et al. (1984), Ashkar and Ouarda (1998), Bhaumik and Gibbons (2006)). However, the true parameter values are never known in practice and the estimators may have large mean-square errors when the sample size is small. On the other hand, the second method is based on a normal approximation to the gamma distribution (Aryal et al. (2007), Krishnamoorthy et al. (2008)). They assumed that the cube root of a gamma random variable is approximately normally distributed. Based on this approximation, the prediction and tolerance limits for the gamma distribution can be constructed by using the corresponding limits for a normal distribution, which has been well received in the literature (Gibbons et al. (2009)). However, according to our simulation as well as the simulation results in Krishnamoorthy et al. (2008), the normal-based method does not work well for either the upper prediction limits or the upper tolerance limits when the gamma shape parameter is small. This is undesirable because the gamma distribution with a small shape parameter is a useful model in various practical applications. For example, the gamma distribution is a useful model for fitting rainfall at certain locations. In a dry area with a relatively high variance, a small shape parameter and a small rate parameter are expected as a gamma distribution as such parameters have small mean but large variance (Husak et al. (2007), Liu et al. (2013)). In addition, the gamma distribution with $k < 1$ has a monotone decreasing hazard function and it is a useful lifetime model for certain types of electronic devices (Coit and Jin (2000), Lawless (2002)).

In this article, we propose new methods for constructing upper prediction and tolerance limits of a gamma distribution based on the generalized fiducial method (Hannig (2009)). In addition, the upper prediction limits for at least p of m observations from a gamma distribution at each of r locations are also constructed. The key idea of the generalized fiducial method is to switch the role of observed data and parameters. This method is found to provide correct coverage asymptotically under some mild conditions

(Hannig et al. (2006)). Many extensive simulations (e.g, Wang et al. (2012)) also show the good performance of the generalized fiducial method in various cases. In this study, we first derive the fiducial distribution of the parameters of a gamma distribution. By using the realizations of the parameters obtained from the fiducial distribution, we can then construct both the prediction and tolerance limits easily.

The proposed generalized fiducial method can be naturally extended to the stress-strength reliability problem, which was introduced by Birnbaum and McCarty (1958) in the context of reliability engineering. It is not uncommon to see that failure of a product unit, e.g., a structural component, is caused by the event that the external stress X_2 exceeds the inherent strength X_1 of the unit, where X_1 and X_2 are random. The stress-strength reliability of the unit is defined as $R = P(X_1 > X_2)$. In addition to the application in reliability engineering, this model arises in many other application areas. For example, the area above the receiver operating characteristic curve can be computed using the stress-strength reliability by properly defining variables X_1 and X_2 (Reiser (2000), Wang and Ye (2015)). There has been a bulk of literature focusing on constructing confidence limits for R by assuming that X_1 and X_2 are two independent normal random variables (e.g., Reiser and Guttman (1986), Weerahandi and Johnson (1992)). Unfortunately, there is still a lack of methods under the gamma-distribution assumption. To the best of our knowledge, the normal-based method seems to be the most accurate one. Similar to the prediction and tolerance limits case, the normal-based method does not work well when the gamma shape parameter is small. In addition, this method requires solving an equation to obtain the noncentrality parameter of a t distribution and hence it is not easy to implement. On the other hand, the proposed generalized fiducial method can be extended to the stress-strength reliability problem in a straightforward manner. The performance will be assessed by a simulation study.

The remainder of the paper is organized as follows. Section 2 gives a brief introduction to the generalized fiducial inference and explains how it can be applied to a gamma distribution. In Section 3, we provide the main algorithms for constructing prediction limits for a single future measurement and for at least p of m measurements at each of r locations. In addition, inferential methods for the tolerance limits as well as the confidence limits for the stress-strength reliability are also developed. Simulations are conducted to assess the performance of the proposed methods in Section 4 and real data are used for illustration in Section 5. Section 6 concludes the paper.

2. Generalized Fiducial Inference

2.1. Introduction of the Generalized Fiducial Inference

The fiducial inference was first developed by Fisher (1930) in an attempt to overcome what he perceived as a deficiency in the Bayesian framework, i.e., the use of prior distribution when little or no prior information about the parameters is available. Because some properties that Fisher (1930) originally claimed were discovered incorrect (Zabell (1992)), the fiducial inference did not gain much acceptance until the introduction of the generalized inference by Weerahandi (1993), where a close relationship between the fiducial and generalized inference was revealed. Hannig et al. (2006) further extended the fiducial inference to generalized fiducial inference. Generally speaking, the generalized fiducial method can provide quite accurate coverage under some mild conditions. More detailed discussion about the asymptotic and empirical properties of the generalized fiducial inference can be found in Hannig (2009).

The principal idea of the generalized fiducial inference is similar to that in the use of a likelihood function, i.e., to switch the role of the parameters and the observed data. To formally describe the fiducial inference, suppose that there is a data-generation process

$$Y = G(\xi, U), \tag{1}$$

where ξ is a vector of parameters and U is a random vector with a completely known distribution independent of any parameters. Equation (1) is also called a structural equation, which can be seen as a detailed description of the noise process U that is combined with the signal ξ to yield the observed data Y. Thus, for any fixed value of the parameter ξ, the distribution of U and the structural equation determine the distribution of the data Y. After Y is observed, the role of the data and the parameters ξ can be switched and one can infer a distribution for ξ. If Equation (1) is invertible, denote the inverse function as G^{-1}. Given Y and U, we can solve the structural Equation (1) to obtain

$$\xi = G^{-1}(Y, U). \tag{2}$$

Therefore, ξ is treated as random because of U. The distribution associated with ξ is called the fiducial distribution of ξ and it can be obtained based on the

observed data Y and U. If the fiducial distribution is hard to derive, then one can get a random realization from the fiducial distribution of ξ by first generating U and then plugging it into Equation (2).

2.2. Generalized Fiducial Inference for the Gamma Distribution

Consider a random variable X following a gamma distribution Gam(k, θ) with probability density function (PDF)

$$f_X(x) = \frac{\theta^k}{\Gamma(k)} x^{k-1} e^{-\theta x}, \quad x > 0,$$

where k is the shape parameter and θ is the rate parameter. Let $x_1, x_2, \ldots, x_n$ be a random sample from Gam(k, θ). Bain and Engelhardt (1975) assumed that

$$W(k) = 2nk \log(\bar{x}/\tilde{x}) \sim c\chi^2_{(v)},$$

where $\bar{x} = \sum_i x_i/n$ and $\tilde{x} = \prod_i x_i^{1/n}$; $\chi^2_{(v)}$ is the χ^2 distribution with v degrees of freedom. They found that, when $k \to 0$, W follows a χ^2 distribution exactly with $2(n-1)$ degrees of freedom. This finding gives a valid support for the χ^2 approximation. In the following, we will use W to represent $W(k)$ if there is no confusion. The values of c and v can be obtained by the moment matching method, i.e., by solving

$$E(W) = cv \quad \text{and} \quad \text{Var}(W) = 2c^2 v.$$

Define $S_0 = \bar{x}/\tilde{x}$ and $S_1 = \log S_0$. After a little manipulation, it is not difficult to show that

$$E(S_1) = -\log n + \psi(kn) - \psi(k)$$

and

$$\text{Var}(S_1) = -\psi_1(kn) + \psi_1(k)/n,$$

where $\psi(x)$ is the digamma function given by $\psi(x) = d\log\Gamma(x)/dx$ and $\psi_1(x)$ is the trigamma function given by $\psi_1(x) = d^2 \log\Gamma(x) dx^2$. Because $W = 2nkS_1$, we have $E(W) = 2nkE(S_1)$ and $\text{Var}(W) = 4n^2k^2\text{Var}(S_1)$. On the other hand, the mean and variance of $c\chi^2_{(v)}$ are cv and $2c^2v$, respectively. Consequently, c and v can be calculated as

$$\begin{aligned} v(k) &= 2E^2(W)/Var(W), \\ c(k) &= E(W)/v. \end{aligned} \tag{3}$$

We express c and v in the above way to reveal that they are functions of the unknown parameter k. The maximum likelihood (ML) estimator of k can be substituted to get the values of c and v. Because the ML estimator of k does not have a closed form, numerical algorithms, such as the quasi-Newton algorithm, may fail to converge when both k and n are very small. In order to overcome this problem, we could approximate the ML estimator of k as

$$\hat{k} = \frac{(n-1)\sum_i x_i}{n\sum_i x_i \log(x_i) - \sum_i \log(x_i) \sum_i x_i}. \tag{4}$$

This estimator of k is proposed by Ye and Chen (2016) and is shown to be unbiased and strongly consistent.

After obtaining c and v, define $U_0 \sim c\chi^2_{(v)}$. Because $W(k)$ is a monotone function of k, we can obtain the exact form of Equation (2) for k as

$$k = W^{-1}(U_0). \tag{5}$$

On the other hand, let $Z(\theta) = 2\theta\sum_i x_i$ and $U_1 \sim \chi^2_{(2nk)}$. Then $Z(\theta)$ and U_1 have the same distribution. Because $Z(\theta)$ is a monotone function of θ, we can obtain the exact form of Equation (2) for θ as

$$\theta = Z^{-1}(U_1). \tag{6}$$

The following Monte Carlo simulation procedure can be used to obtain the realizations of (k, θ), and the R code is provided in the Appendix.

Algorithm 1: Obtain Realizations of (k, θ)

(1) For a given data $x_1, x_2, \ldots, x_n$, compute the ML estimate of k by Equation (4), denoted as $\hat{k}$.

(2) Substitute $\hat{k}$ in Equation (3) to obtain the values of c and v, denoted as $\hat{c}$ and $\hat{v}$.

(3) Generate $U_{0:1}, \ldots, U_{0:B}$ from $\hat{c}\chi^2_{(\hat{v})}$ and substitute them into Equation (5) to get $k_1, \ldots, k_B$.

(4) Generate $U_{1:b}$ from $\chi^2_{(2nk_b)}$, $b = 1, \ldots, B$ and substitute $U_{1:1}, \ldots, U_{1:B}$ into Equation (6) to get $\theta_1, \ldots, \theta_B$.

Remark: Wang et al. (2012) also considered the generalized fiducial inference for a gamma distribution. The structural equation they used is $\xi = F^{-1}(Y, U)$, where $F^{-1}(\cdot)$ is the inverse of the cumulative distribution function of Y and U follows a standard uniform distribution. Using this standard method, the joint fiducial distribution of (k, θ) does not have a closed form and importance sampling is needed to generate realizations of (k, θ). On the other hand, by constructing the structural equations for k and θ sequentially, our algorithm is obviously more straightforward to follow and much less computationally demanding.

3. The Proposed Methods

In this section, we articulate the main algorithms for constructing upper prediction limits for a single future measurement and for at least p of m measurements at each of r locations. In addition, inferential procedures for tolerance limits and confidence limits of the stress-strength reliability are also provided.

3.1. Upper Prediction Limits for Single Future Measurement

Assume that n historical data, i.e., $x_1, \ldots, x_n$, from a gamma distribution are available. Consider a statistic $PL = PL(x_1, \ldots, x_n)$ and a future measurement x_{n+1} from the same gamma distribution. If PL and x_{n+1} have the following relationship,

$$P(x_{n+1} < PL) = 1 - \alpha,$$

where $0 < \alpha < 1$ is a given constant, then PL is called a $(1-\alpha)$ upper prediction limit for the single future measurement. In other words, we have $1-\alpha$ confidence that the next future measurement is below PL. In the literature, the normal-based method is often used for constructing the prediction limits of a gamma distribution (Krishnamoorthy et al. (2008), Gibbons et al. (2009)). Let $y_1, \ldots, y_n$ be the cube-root transformed sample, i.e, $y_i = x_i^{1/3}, i = 1, \ldots, n$. Then y_i's are approximately normally distributed and the $(1-\alpha)$ UPL for the transformed sample is $PL_n = \bar{y} + t_{n-1,1-\alpha} s_y \sqrt{1 + 1/n}$, where $\bar{y}$ and s_y are the mean and standard error of y_i's, and $t_{a,b}$ is the bth quantile of the t distribution with degrees of freedom a. The $(1-\alpha)$ UPL for the original gamma sample is then approximated by $(PL_n)^3$. The procedure of the normal-based method in constructing the upper prediction limit for a single future measurement is straightforward. However, according to our simulation, the normal-based method does not work well when the shape parameter k is small (e.g., $k < 1$), as the coverage probabilities so obtained may deviate significantly from the nominal values. On the other hand, we can use the generalized fiducial method to construct prediction limits of a gamma distribution. In the previous section, we have shown how to obtain realizations of k and θ. By utilizing these realizations, the prediction limits can be easily constructed, as shown in the following algorithm. In addition, we also provide R codes for implementing this algorithm in the Appendix.

Algorithm 2: UPLs for a Single Future Measurement

(1) Use algorithm 1 to obtain $k_1, \ldots, k_B$ and $\theta_1, \ldots, \theta_B$ based on n historical data.

(2) For each (k_b, θ_b), $b = 1, \ldots, B$, generate a random realization from $\text{Gam}(k_b, \theta_b)$, denoted as PL_b.

(3) Sort the values $PL_1, \ldots, PL_B$ in ascending order and use the $100(1-\alpha)$th percentile as the $(1-\alpha)$ upper prediction limit for the next single future measurement.

3.2. Upper Prediction Limits for at Least p of m Measurements at Each of r Locations

As introduced in Section 1, simultaneous prediction limits are required in environmental monitoring as well as quality-control problems. For example, in a groundwater-monitoring problem, a series of m sample measurements from each of r monitoring wells located hydraulically downgradient of the hazardous waste management facilities are often compared with the prediction limits based on n measurements from one or more upgradient sampling locations (Krishnamoorthy et al. (2008)). The statistical problem is to construct a prediction limit such that at least p of m measurements at each of r locations are below the limit with confidence $1-\alpha$. Previous methods for this problem are often based on the normal assumption for the monitoring constituents (e.g., Davis and McNichols (1987), Gibbons et al. (2009)). However, Bhaumik and Gibbons (2006) revealed that the gamma distribution is more appropriate for many constituents. The authors then developed a method to construct the upper prediction limits for at least p of m measurements at each of r locations by assuming the gamma distribution. They restricted the prediction limit in the form $\bar{x} + cS$, where $\bar{x}$ is the sample mean and $S = \sqrt{\hat{k}/\hat{\theta}}$ is an estimate of the population standard deviation. Therefore, the problem is simplified to determine the value of c. According to their simulation results, this method cannot guarantee the closeness of the coverage probabilities to the nominal values in many parameter settings. On the other hand, perhaps a more accurate method is the normal-based method. Based on the cube-root transformed data $y_1, \ldots, y_n$, the simultaneous prediction limits have the form $\bar{y} + cs_y$, where c is obtained by solving an equation that has a complicated form and involves numerical integration (Krishnamoorthy (2008)). This requirement poses difficulty in implementing the normal-based method. In addition, the normal-based method does not perform well when the shape parameter k is small. On the contrary, our generalized fiducial method again bridges this gap. The procedure is shown in the following algorithm.

Algorithm 3: UPLs for at Least p of m Measurements at Each of r Locations

(1) Use algorithm 1 to obtain $k_1, \ldots, k_B$ and $\theta_1, \ldots, \theta_B$ based on n historical data.

(2) For each (k_b, θ_b), $b = 1, \ldots, B$, generate m random realizations for each of r locations from $\text{Gam}(k_b, \theta_b)$ and find the smallest value, below which there are at least p of m values at each of r locations; denote the value as SPL_b.

(3) Sort the values $SPL_1, \ldots, SPL_B$ in ascending order and use the $100(1-\alpha)$th percentile as the $(1-\alpha)$ upper prediction limit for at least p of m measurements at each of r locations.

3.3. Tolerance Limits

When the number of future measurements is either large or unknown, the tolerance limit is of more interest than the prediction limit. Assume that n historical data $x_1, \ldots, x_n$ from a gamma distribution are available. Consider two statistics $L_1 = L_1(x_1, \ldots, x_n)$ and $L_2 = L_2(x_1, \ldots, x_n)$. Let β and α be two constants and $0 < \beta, \alpha < 1$. If L_1 and L_2 are determined such that

$$P[P(L_1 \leq X \leq L_2) \geq \beta] = 1 - \alpha,$$

then (L_1, L_2) is called a $(\beta, 1-\alpha)$ two-sided tolerance interval, which contains at least β proportion of the population with $1-\alpha$ confidence. If we set $L_1 = -\infty$, then the interval $(-\infty, L_2)$ is called a $(\beta, 1-\alpha)$ upper tolerance limit. Similarly, $L_2 = \infty$ corresponds to a $(\beta, 1-\alpha)$ lower tolerance limit.

To the best of our knowledge, the most accurate method to construct the tolerance limits for the gamma distribution is again the normal-based method. Based on the cube-root transformed data $y_1, \ldots, y_n$, a $(\beta, 1-\alpha)$ upper tolerance limit is approximated by $(\bar{y} + \frac{1}{\sqrt{n}} t_{n-1,1-\alpha}(z_\beta \sqrt{n}) s_y)^3$, where $\bar{y}$ and s_y are the mean and standard error of y_i's, $t_{a,b}(c)$ is the bth quantile of the t distribution with degrees of freedom a and noncentrality parameter c, and z_p is the pth quantile of the standard normal distribution. The normal-based lower tolerance limit can be constructed in a similar vein. Although the normal-based method for the one-sided tolerance limits is easy to follow, it does not work well when $k < 1$ because of the error of the normal approximation. Krishnamoorthy et al. (2008) stated that the coverage probabilities for the normal-based upper (lower) tolerance limit are expected to be smaller (larger) than the nominal values when the shape parameter $k < 1$.

Alternatively, we can again use the generalized fiducial method to construct the one-sided tolerance limits for the gamma distribution. The one-sided tolerance limit is directly related to the confidence limit for the quantiles. On one hand, the $(\beta, 1-\alpha)$ upper tolerance limit is equal to the $(1-\alpha)$ upper confidence limit for the βth quantile τ_β. On the other hand, the $(\beta, 1-\alpha)$ lower tolerance limit is equal to the $(1-\alpha)$ lower confidence limit for the $(1-\beta)$th quantile $\tau_{1-\beta}$. Because the quantile of a gamma distribution is a function of k and θ, we can easily obtain the one-sided tolerance limit by utilizing the realizations of k and θ. Algorithm 4 below shows the procedure to obtain the $(\beta, 1-\alpha)$ upper tolerance limit, and the R codes are given in the Appendix. The lower tolerance limits can be constructed in a similar vein.

Algorithm 4: Upper Tolerance Limits

(1) Use algorithm 1 to obtain $k_1, \ldots, k_B$ and $\theta_1, \ldots, \theta_B$ based on n historical data.

(2) Compute the βth quantile τ_β of a gamma distribution with parameters k_b and θ_b, denoted as τ_{β_b}.

(3) Sort the values $\tau_{\beta_1}, \ldots, \tau_{\beta_B}$ in ascending order and use the $100(1-\alpha)$th percentile as the $(\beta, 1-\alpha)$ upper tolerance limit L_2.

Remark: In most environmental monitoring and quality-control problems, the one-sided tolerance limits are of interest (e.g., Krishnamoorthy and Mathew (2004), Bhaumik and Gibbons (2006), Fernández (2010), Ryan (2011)). Therefore, we did not discuss construction of the two-sided tolerance intervals in this study. In fact, the generalized fiducial inference method could also be used for constructing the two-sided tolerance interval with some effort. The detailed procedure is given in the Appendix. According to our preliminary simulations (not reported), the performance of the generalized fiducial inference method and the normal-based method is quite similar. Although the normal-based method performs poorly for constructing the one-sided tolerance limits when k is small, its performance seems to be uniformly good for two-sided tolerance intervals regardless of the values of k (Krishnamoorthy et al. (2008)). From a practical point of view, we would recommend the normal-based method for constructing the two-sided tolerance intervals as its procedure is more straightforward.

3.4. Stress-Strength Reliability

As introduced before, the stress-strength reliability problem arises in many important areas, especially in the context of reliability engineering. Consider an inherent strength variable X_1 and an external stress variable X_2. A failure is defined as the event that X_2 exceeds X_1. The stress-strength reliability is then given by

$$R = P(X_1 > X_2).$$

The stress-strength reliability problem has been well addressed when both the stress and strength are normal (e.g., Reiser and Guttman (1986), Guo and Krishnamoorthy (2004)). Based on these well-developed methods, the normal-based method can again be used for constructing confidence limits for R when both X_1 and X_2 are gamma distributed (Krishnamoorthy et al. (2008)). Similar to the prediction and tolerance limits, the normal-based method does not perform well when either of the gamma shape parameters is small. In addition, the noncentrality parameter of a t distribution is needed by solving an equation during the procedure of the normal-based method. Therefore, it is not so straightforward to implement the normal-based method. For the detailed procedure of the normal-based method, we refer readers to Krishnamoorthy et al. (2008). On the contrary, our generalized fiducial method can be naturally extended to this problem.

Assume that $X_1 \sim \text{Gam}(k_1, \theta_1)$ and $X_2 \sim \text{Gam}(k_2, \theta_2)$. Based on the relationship between the gamma distribution and the F distribution, the stress-strength reliability parameter R can then be expressed as

$$R = P\left(F_{2k_1,2k_2} > \frac{k_2\theta_1}{k_1\theta_2}\right), \tag{7}$$

where $F_{2k_1,2k_2}$ represents an F-distributed random variable with degrees of freedom $2k_1$ and $2k_2$. By generating a series of realizations of (k_1, θ_1) and (k_2, θ_2) using algorithm 1, confidence limits for R can be easily constructed. The following procedure shows how to construct the lower confidence limit for R. Both the upper confidence limit and the confidence interval can be constructed in a similar way.

Algorithm 5: Lower Confidence Limits for Stress-Strength Reliability

(1) Use algorithm 1 to obtain $k_{1:1}, \ldots, k_{1:B}$, $\theta_{1:1}, \ldots, \theta_{1:B}$ based on the n_1 inherent strength data and $k_{2:1}, \ldots, k_{2:B}, \theta_{2:1}, \ldots, \theta_{2:B}$ based on the n_2 external strength data.

(2) For each $(k_{1:b}, \theta_{1:b})$ and $(k_{2:b}, \theta_{2:b})$, $b = 1, \ldots, B$, compute R_b by using Equation (7).

(3) Sort the values $R_1, \ldots, R_B$ in ascending order and use the 100αth percentile as the $(1-\alpha)$ lower confidence limit for R.

4. Simulation Study

In this section, we conduct simulation studies to assess the performance of algorithms 2–5. We generate $B = 10{,}000$ realizations of k and θ in each simulation. Coverage probabilities of each statistical limit discussed in algorithms 2–5 are obtained based on 10,000 Monte Carlo replications.

4.1. Upper Prediction Limits for Single Future Measurement

We first consider algorithm 2 for constructing the upper prediction limits for a single future measurement. In this simulation, we set the rate parameter $\theta = 1$ and let the shape parameter $k = 0.05, 0.1, 0.5, 1$. Small sample sizes, i.e., $n = 3, 5, 10$ are considered. We are interested in the 95% and 90% upper prediction limits. The coverage probabilities based on algorithm 2 and the normal-based method

TABLE 1. Coverage Probabilities of the Upper Prediction Limits Based on Algorithm 2 and the Normal-Based Method (in Parentheses); for Each Combination of k and n, the Nominal Coverage Probabilities are 95% (Left) and 90% (Right)

n	$k = 0.05$	$k = 0.1$	$k = 0.5$	$k = 1$
3	0.953 / 0.909	0.952 / 0.903	0.956 / 0.898	0.949 / 0.899
	(0.818 / 0.797)	(0.863 / 0.823)	(0.937 / 0.881)	(0.942 / 0.892)
5	0.952 / 0.901	0.950 / 0.896	0.951 / 0.896	0.949 / 0.900
	(0.859 / 0.840)	(0.887 / 0.840)	(0.938 / 0.888)	(0.941 / 0.890)
10	0.948 / 0.900	0.947 / 0.897	0.949 / 0.898	0.948 / 0.901
	(0.894 / 0.872)	(0.897 / 0.863)	(0.939 / 0.889)	(0.940 / 0.889)

TABLE 2. Coverage Probabilities of the Upper Prediction Limits Based on Algorithm 3 when $r = 5$; for Each Setting, the Nominal Coverage Probabilities are 95% (Left) and 90% (Right)

n	(m,p)	$k = 0.05$	$k = 0.1$	$k = 0.5$	$k = 1$
3	(3,1)	0.954 / 0.907	0.946 / 0.899	0.946 / 0.890	0.947 / 0.896
	(3,2)	0.954 / 0.903	0.952 / 0.908	0.949 / 0.906	0.946 / 0.897
	(4,1)	0.951 / 0.907	0.948 / 0.893	0.945 / 0.892	0.948 / 0.896
	(4,2)	0.953 / 0.902	0.952 / 0.897	0.953 / 0.895	0.950 / 0.899
5	(3,1)	0.954 / 0.895	0.955 / 0.893	0.949 / 0.890	0.951 / 0.895
	(3,2)	0.953 / 0.906	0.955 / 0.897	0.960 / 0.895	0.944 / 0.900
	(4,1)	0.955 / 0.899	0.946 / 0.897	0.909 / 0.893	0.945 / 0.897
	(4,2)	0.959 / 0.896	0.956 / 0.904	0.957 / 0.891	0.946 / 0.894
10	(3,1)	0.954 / 0.893	0.957 / 0.899	0.958 / 0.898	0.951 / 0.893
	(3,2)	0.950 / 0.903	0.946 / 0.903	0.949 / 0.900	0.951 / 0.896
	(4,1)	0.951 / 0.906	0.944 / 0.895	0.948 / 0.903	0.955 / 0.895
	(4,2)	0.948 / 0.890	0.945 / 0.897	0.945 / 0.900	0.949 / 0.899

are shown in Table 1. As can be seen, our proposed generalized fiducial method performs well in all the settings, as the coverage probabilities are very close to the nominal values. On the contrary, the coverage probabilities based on the normal-based method seem to be uniformly lower than the nominal values.

4.2. Upper Prediction Limits for at Least p of m Measurements at Each of r Locations

We then conduct a simulation to assess the performance of algorithm 3. Similar to the settings in the last subsection, we set $\theta = 1$ and $k = 0.05, 0.1, 0.5, 1$. We consider various combinations of (r, m, p). Specifically, we set $r = 1$, 5, 10 and, under each value of r, $(m,p) = (3,1), (3,2), (4,1), (4,2)$ are considered. The simulation results for $r = 5$ are shown in Table 2 and other simulation results for $r = 1, 10$ can be found in the Appendix. As can be seen, the coverage probabilities based on the generalized fiducial method are very close to the nominal values regardless of the parameter settings. It is worth mentioning that the normal-based method works poorly when the shape parameter k is small. Its performance is quite similar to that in the last subsection and hence the results are omitted here.

4.3. Tolerance Limits

We then consider tolerance limits based on algorithm 4. We set the rate parameter $\theta = 1$ and let the shape parameter $k = 0.05, 0.1, 0.5, 1$. Small sample sizes, i.e., $n = 5, 10, 15$, are considered. In addition, we consider the proportion of the population $\beta = 0.9, 0.95, 0.99$. Under each proportion, we are interested in the upper tolerance limits with confidence level $1 - \alpha = 0.99, 0.95, 0.90$. The coverage probabilities based on Algorithm 4 and the normal-based method are shown in Table 3. As we can see, our proposed methods perform very well regardless of the true values of k. On the contrary, the normal-based method performs poorly for $k < 1$ and its performance improves with the value of k. This finding is consistent with that in Krishnamoorthy et al. (2008). In conclusion, the normal-based method could be safely used for constructing the one-sided tolerance limits when $k \geq 1$ while our proposed method has a satisfactory performance for all values of k.

4.4. Stress-Strength Reliability

At last, a simulation is conducted to assess the accuracy of algorithm 5. We set $\theta_1 = \theta_2 = 1$, $k_1 =$ 0.5, 1, 3, and $k_2 = 2, 4, 6$. In addition, $(n_1, n_2) =$ (5, 10), (10, 10), (20, 10) are considered. We are interested in the 90% and 95% lower confidence limits for the stress-strength reliability R. Based on algorithm 1, $B = 10{,}000$ realizations of (k_1, θ_1) and (k_2, θ_2) are obtained. The coverage probabilities are shown in Table 4. As expected, the coverage probabilities based on algorithm 5 are quite close to the nominal values regardless of the parameter settings. We do not show the performance of the normal-based method here because it is quite similar to that in constructing-prediction limits and tolerance limits. In a nutshell, the normal-based method performs poorly when ei-

TABLE 3. Coverage Probabilities of the Upper Prediction Limits Based on Algorithm 4 and the Normal-Based Method (in Parentheses)

		$\beta = 0.9$			$\beta = 0.95$			$\beta = 0.99$		
n	k	0.99	0.95	0.90	0.99	0.95	0.90	0.99	0.95	0.90
5	0.05	0.995	0.958	0.907	0.996	0.958	0.916	0.996	0.962	0.917
		(0.705)	(0.613)	(0.556)	(0.652)	(0.542)	(0.474)	(0.604)	(0.464)	(0.400)
	0.1	0.995	0.958	0.916	0.997	0.967	0.919	0.996	0.966	0.922
		(0.838)	(0.733)	(0.665)	(0.817)	(0.691)	(0.615)	(0.808)	(0.657)	(0.558)
	0.5	0.992	0.959	0.910	0.993	0.962	0.916	0.994	0.963	0.917
		(0.983)	(0.929)	(0.868)	(0.984)	(0.931)	(0.873)	(0.987)	(0.938)	(0.871)
	1	0.993	0.953	0.904	0.991	0.955	0.910	0.991	0.954	0.905
		(0.990)	(0.945)	(0.892)	(0.988)	(0.945)	(0.897)	(0.989)	(0.946)	(0.894)
10	0.05	0.991	0.951	0.900	0.992	0.955	0.901	0.992	0.955	0.905
		(0.748)	(0.663)	(0.603)	(0.636)	(0.524)	(0.450)	(0.540)	(0.388)	(0.309)
	0.1	0.992	0.955	0.904	0.993	0.951	0.903	0.994	0.955	0.905
		(0.833)	(0.720)	(0.651)	(0.795)	(0.653)	(0.561)	(0.741)	(0.582)	(0.481)
	0.5	0.991	0.952	0.900	0.993	0.958	0.909	0.992	0.960	0.906
		(0.978)	(0.918)	(0.850)	(0.980)	(0.918)	(0.854)	(0.981)	(0.921)	(0.851)
	1	0.991	0.955	0.899	0.991	0.953	0.901	0.992	0.952	0.902
		(0.988)	(0.944)	(0.884)	(0.987)	(0.942)	(0.884)	(0.987)	(0.941)	(0.888)
15	0.05	0.990	0.953	0.902	0.991	0.950	0.903	0.990	0.953	0.904
		(0.762)	(0.684)	(0.630)	(0.625)	(0.502)	(0.423)	(0.459)	(0.301)	(0.236)
	0.1	0.991	0.948	0.900	0.991	0.952	0.900	0.992	0.953	0.907
		(0.832)	(0.707)	(0.639)	(0.761)	(0.606)	(0.521)	(0.679)	(0.496)	(0.395)
	0.5	0.993	0.955	0.896	0.990	0.954	0.903	0.993	0.956	0.906
		(0.98)	(0.912)	(0.837)	(0.975)	(0.913)	(0.837)	(0.978)	(0.905)	(0.841)
	1	0.990	0.952	0.895	0.993	0.954	0.899	0.993	0.955	0.905
		(0.987)	(0.939)	(0.877)	(0.988)	(0.940)	(0.882)	(0.989)	(0.943)	(0.886)

ther k_1 or k_2 is small and its implementation is not so straightforward.

5. Illustrative Examples

In this section, we use some real data sets to illustrate the application of our methods. In addition, we also present results based on other existing methods for the purpose of comparison.

5.1. Upper Prediction Limits for at Least p of m Measurements at Each of r Locations

We first consider upper prediction limits for at least p of m measurements at each of r locations based on algorithm 3. The vinyl chloride data from clean upgradient groundwater monitoring wells given in Table 5 are used. The data can also be found from Table 1 in Bhaumik and Gibbons (2006). As a volatile organic compound, the vinyl chloride causes great harm to human health. Therefore, it is crucially important to detect its release. Bhaumik and Gibbons (2006) showed that the gamma distribution can fit this data set very well. In order to compare our method with the methods in Bhaumik and Gibbons (2006) and Krishnamoorthy et al. (2008), we consider the same settings as those in these two papers. Specifically, $(r, m, p) = (1, 2, 1), (10, 2, 1), (10, 3, 1)$ and $(10, 3, 2)$ are considered and 95% simultaneous upper prediction limits are constructed. The results are shown in Table 6. Generally, the upper prediction limits based on algorithm 3 are larger than those based on the normal-based method, which indicates that our proposed method is more accurate. This is because the estimated gamma shape parameter for this data set is 1.063 and, when $k = 1$, the upper prediction limits based on the normal-based method

TABLE 4. Coverage Probabilities of the Lower Confidence Limits for the Stress-Strength Reliability Based on Algorithm 5

		$k_1 = 0.5$		$k_1 = 1.0$		$k_1 = 3.0$	
n_1	k_2	0.95	0.90	0.95	0.90	0.95	0.90
5	2	0.955	0.899	0.951	0.907	0.954	0.904
	4	0.949	0.896	0.946	0.898	0.951	0.903
	6	0.949	0.898	0.946	0.892	0.950	0.897
10	2	0.947	0.893	0.950	0.903	0.959	0.901
	4	0.954	0.891	0.953	0.894	0.954	0.904
	6	0.949	0.899	0.952	0.895	0.949	0.894
15	2	0.950	0.899	0.946	0.901	0.953	0.902
	4	0.951	0.891	0.949	0.898	0.953	0.899
	6	0.951	0.898	0.945	0.894	0.944	0.891
20	2	0.951	0.899	0.947	0.892	0.954	0.904
	4	0.949	0.892	0.949	0.898	0.949	0.901
	6	0.950	0.898	0.950	0.894	0.944	0.893

tend to be small as the coverage probabilities are smaller than the nominal values (see simulations in the last section).

5.2. Tolerance Limits

We then use the data set in Krishnamoorthy et al. (2008) to illustrate the application of algorithm 4. This data set in Table 7 represents the alkalinity concentrations in groundwater. The size of the data set is 27 and the gamma distribution is found to provide a good fit to this data set (Aryal et al. (2007)). The one-sided tolerance limits are of interest. In Krishnamoorthy et al. (2008), three values of (β, α) are used, i.e., $(\beta, \alpha) = (0.9, 0.05)$, $(0.95, 0.05)$, $(0.99, 0.05)$. In this section, we also use these values of (β, α) for comparison purposes. The results based on the normal-based method and our generalized fiducial method are shown in Table 8. Because the estimated shape parameter for these data is 9.37, it is not surprising that the results from these two methods are quite close.

TABLE 6. 95% Simultaneous Upper Prediction Limits for the Vinyl Chloride Data Based on the Generalized Fiducial Method (Fiducial), the Method in Bhaumik and Gibbons (2006) (Bhaumik-Gibbons), and the Method in Krishnamoorthy et al. (2008) (Normal-Based)

r	m	p	Fiducial	Bhaumik-Gibbons	Normal-based
1	2	1	2.893	2.931	2.893
10	2	1	5.442	5.224	5.203
10	3	1	3.605	3.521	3.479
10	3	2	6.569	6.330	6.369

5.3. Stress-Strength Reliability

The inferential procedure for the stress-strength reliability is demonstrated using an example from a manufacturing factory. The company uses cutting machines for production, and the drill is one of the important components in the cutting machine. Drills of different sizes are needed in the production, and we focus on drill size 1.88 mm. Currently, the factory purchases the drills from two different suppliers. After a certain period of usage, it is of interest to know which brand is more reliable so that the factory can make the subsequent purchase decision. Table 9 presents the lifetimes of the drill of size 1.88 mm from the two suppliers. We first consider fitting the drill lifetimes by three popular lifetime distributions, i.e., gamma, Weibull, and log-normal. According to Figure 1, the gamma distribution and the log-normal distribution seem to provide a better fit than the Weibull distribution. On the other hand, the Akaike information criterion (AIC) values in Table 10 indicate that the gamma distribution is the preferred model for these drill lifetimes. Clearly, $R = P(X_1 > X_2)$ is the probability that the drill lifetime by the first supplier is larger than that by the second supplier. Based on algorithm 5, the 95% lower confidence limit of R is 0.867 (0.870 by the normal-based method), indicating that drills of size 1.88 mm from the first supplier have a higher quality.

TABLE 5. Vinyl Chloride Data from Clean Upgradient Groundwater-Monitoring Wells in μ g/L

5.1	2.4	0.4	0.5	2.5	0.1	6.8	1.2	0.5	0.6	5.3	2.3	1.8	1.2	1.3	1.1	0.9
3.2	1.0	0.9	0.4	0.6	8.0	0.4	2.7	0.2	2.0	0.2	0.5	0.8	2.0	2.9	0.1	4.0

TABLE 7. Alkalinity Concentrations in Groundwater

58	82	42	28	118	96	49	54	42
51	66	89	40	51	54	55	59	42
39	40	60	63	59	70	32	52	79

TABLE 8. Tolerance Limits for the Alkalinity Concentrations Data Based on the Generalized Fiducial Method and the Normal-Based Method (in Parentheses)

$(\beta, 1-\alpha)$	Lower	Upper
(0.90,0.95)	28.278 (28.341)	97.768 (97.713)
(0.95,0.95)	23.165 (23.296)	110.90 (110.51)
(0.99,0.95)	15.493 (15.400)	137.70 (137.94)

6. Conclusion

In some environmental-monitoring problems and control problems, it is of critical importance to construct prediction limits as well as tolerance limits. The prediction and tolerance limits can help determine whether a change has occurred in the process. In these problems, a gamma distribution is often found to be more appropriate than the normal distribution in various cases. To the best of our knowledge, the most advanced work for the construction of prediction and tolerance limits for the gamma distribution is the normal-based method, where the normal approximation to a gamma random variable is used. However, this method does not work so well when the gamma shape parameter is small. In this article, we have constructed both the prediction and tolerance limits for the gamma distribution based on the generalized fiducial method. The simulation re-

TABLE 9. Lifetimes (in Minutes) of 1.88-mm Drill from Two Suppliers

X_1	135	98	114	137	138	144	99	93	115	106	132	122	94	98	127	122
	102	133	114	120	93	126	119	104	119	114	125	107	98	117	111	106
	108	127	126	135	112	94	127	99	120	120	121	122	96	109	123	105
X_2	105	105	95	87	112	80	95	97	77	103	78	87	107	96	79	91
	108	97	80	76	92	85	76	96	77	80	100	94	82	104	91	95
	93	99	99	94	84	99	91	85	86	79	89	89	100			

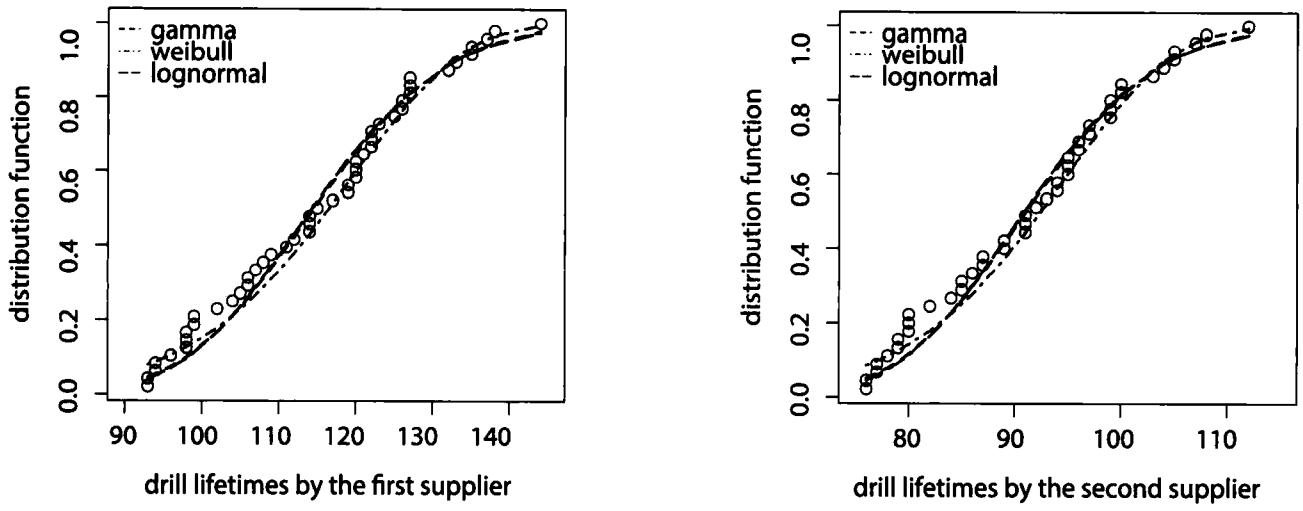

FIGURE 1. Gamma, Weibull, and Log-Normal Probability Plots for the Drill-Lifetime Data. The circle denotes the empirical distribution function and the lines denote the estimated distribution functions using the three parametric models.

TABLE 10. AIC Values for Gamma, Weibull and Log-Normal

	Gamma	Weibull	Log-normal
X_1	389.87	391.96	390.08
X_2	335.27	337.80	335.38

sults have shown the excellent performance of the proposed methods in terms of coverage probabilities. For each case, we have also illustrated how to apply the proposed methods in real problems. Another important finding is that the proposed methods can also be extended to the stress-strength reliability problem. The simulation results have again revealed the convincing performance of the proposed methods.

Acknowledgment

We are grateful to the editor and two referees for their insightful comments that have lead to a substantial improvement to an earlier version of the paper. The work is supported by Singapore AcRF Tier 1 funding #R-266-000-081-133, and also the National Research Foundation Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE).

Appendix

R Codes for Some Proposed Algorithms

In this paper, all our algorithms are coded in the R language (version 3.2.3). Here, we provide the codes for generating realizations of (k, θ) (algorithm 1), constructing upper prediction limits for a single future measurement (algorithm 2), and constructing upper tolerance limits (algorithm 4). To save space, the codes for algorithm 3 and algorithm 5 are not shown here. Nevertheless, they can be produced in a very similar vein.

Algorithm of Constructing Two-Sided Tolerance Intervals

In this section, we briefly discuss the construction of two-sided tolerance intervals by the generalized fiducial inference method. For a $(\beta, 1-\alpha)$ equal-tailed, two-sided tolerance interval (L_1, L_2), L_1 and L_2 should satisfy

$$P(L_1 \le \tau_{(1-\beta)/2} \text{ and } \tau_{(1+\beta)/2} \le L_2) = 1 - \alpha, \quad (8)$$

where τ_β is the βth quantile of a gamma distribution. Note that the solution for the above question is not unique. We may restrict ourselves with an additional condition

$$P(L_1 \le \tau_{(1-\beta)/2}) = P(\tau_{(1+\beta)/2} \le L_2) = 1 - \alpha^*, \quad (9)$$

as the one-sided tolerance limits for a given α^* can be easily obtained by algorithm 2. After generating B realizations of (k, θ), we can obtain B realizations of the quantiles $(\tau_{(1-\beta)/2}, \tau_{(1+\beta)/2})$. The value of α^* is then determined such that the corresponding one-sided tolerance limits L_1 and L_2 in Equation (9) satisfy condition (8), where the probability in condition (8) is approximated by the proportion of the pairs $(\tau_{(1-\beta)/2}, \tau_{(1+\beta)/2})$'s that satisfy $\tau_{(1-\beta)/2} \ge L_1$ and $\tau_{(1+\beta)/2} \le L_2$. The procedure is summarized below.

(1) Use algorithm 1 to obtain $k_1, \ldots, k_B$ and $\theta_1, \ldots, \theta_B$. For each (k_b, θ_b), obtain one realization of the quantiles $(\tau_{(1-\beta)/2}, \tau_{(1+\beta)/2})$, denoted as $(\tau^{(b)}_{(1-\beta)/2}, \tau^{(b)}_{(1+\beta)/2})$, $b = 1, \ldots, B$.

(2) Determine the value of α^* such that (L_1, L_2)

```
## Generating realizations of (k,theta)
## data is a sample from a gamma distribution and n is the number of data points
xtilde <- (prod(data))^(1/n); xbar <- mean(data)
s <- log(xbar)-mean(log(data))
khat <- (n-1)*sum(data)/(n*sum(data*log(data))-sum(log(data))*sum(data))
Ew <- 2*n*khat*(-log(n)+digamma(khat*n)-digamma(khat))
Vw <- 4*n^2*khat^2*(1/n*psigamma(khat,1)-psigamma(khat*n,1))
v <- 2*Ew^2/Vw; c <- Ew/v
kset <- c*rchisq(B,v)/(2*n*log(xbar/xtilde)) #B realizations of k
thetaset <- rchisq(B,2*n*kset)/2/sum(data) #B realizations of theta
## Constructing (1-alpha) upper prediction limit
PL <- quantile(rgamma(B,kset,thetaset),1-alpha,na.rm=T)
## Constructing (b,1-alpha) upper tolerance limit
L2 <- quantile(qgamma(b,kset,thetaset),1-alpha,na.rm=T)
```

TABLE 11. Coverage Probabilities of the Upper Prediction Limits Based on Algorithm 3 when $r = 1$; for Each Setting, the Nominal Coverage Probabilities are 95% (Left) and 90% (Right)

n	(m,p)	$k = 0.05$	$k = 0.1$	$k = 0.5$	$k = 1$
3	(3,1)	0.955 / 0.906	0.950 / 0.896	0.943 / 0.896	0.945 / 0.895
	(3,2)	0.951 / 0.899	0.946 / 0.892	0.949 / 0.890	0.952 / 0.894
	(4,1)	0.954 / 0.905	0.954 / 0.891	0.943 / 0.899	0.944 / 0.894
	(4,2)	0.957 / 0.899	0.949 / 0.899	0.956 / 0.898	0.947 / 0.898
5	(3,1)	0.949 / 0.908	0.946 / 0.894	0.947 / 0.904	0.944 / 0.899
	(3,2)	0.953 / 0.905	0.947 / 0.895	0.951 / 0.899	0.955 / 0.896
	(4,1)	0.957 / 0.900	0.944 / 0.891	0.947 / 0.890	0.942 / 0.891
	(4,2)	0.949 / 0.894	0.951 / 0.893	0.943 / 0.891	0.952 / 0.897
10	(3,1)	0.947 / 0.903	0.944 / 0.899	0.952 / 0.898	0.947 / 0.907
	(3,2)	0.948 / 0.894	0.948 / 0.890	0.946 / 0.901	0.958 / 0.906
	(4,1)	0.944 / 0.907	0.947 / 0.907	0.945 / 0.892	0.952 / 0.906
	(4,2)	0.950 / 0.898	0.955 / 0.893	0.941 / 0.904	0.943 / 0.897

TABLE 12. Coverage Probabilities of the Upper Prediction Limits Based on Algorithm 3 when $r = 10$; for Each Setting, the Nominal Coverage Probabilities are 95% (Left) and 90% (Right)

n	(m,p)	$k = 0.05$	$k = 0.1$	$k = 0.5$	$k = 1$
3	(3,1)	0.956 / 0.908	0.957 / 0.908	0.954 / 0.899	0.959 / 0.892
	(3,2)	0.950 / 0.945	0.951 / 0.903	0.956 / 0.904	0.957 / 0.905
	(4,1)	0.949 / 0.898	0.950 / 0.895	0.946 / 0.894	0.948 / 0.894
	(4,2)	0.955 / 0.908	0.954 / 0.908	0.942 / 0.893	0.949 / 0.891
5	(3,1)	0.946 / 0.897	0.942 / 0.903	0.945 / 0.893	0.946 / 0.893
	(3,2)	0.956 / 0.901	0.950 / 0.906	0.952 / 0.904	0.959 / 0.897
	(4,1)	0.952 / 0.896	0.942 / 0.899	0.949 / 0.894	0.947 / 0.894
	(4,2)	0.949 / 0.896	0.941 / 0.896	0.948 / 0.902	0.949 / 0.890
10	(3,1)	0.947 / 0.901	0.947 / 0.894	0.954 / 0.889	0.941 / 0.896
	(3,2)	0.944 / 0.899	0.955 / 0.896	0.957 / 0.902	0.947 / 0.892
	(4,1)	0.954 / 0.896	0.947 / 0.898	0.946 / 0.896	0.945 / 0.907
	(4,2)	0.953 / 0.901	0.945 / 0.902	0.941 / 0.899	0.947 / 0.893

satisfy condition (8), where L_1 is the $100(1-\alpha^*)\%$ lower confidence limit for $\tau_{(1-\beta)/2}$ and L_2 is the $100(1-\alpha^*)\%$ upper confidence limit for $\tau_{(1+\beta)/2}$. The confidence limits for the quantiles can be constructed by algorithm 2, and the probability in Equation (8) is approximated by the proportion of the pairs $(\tau^{(b)}_{(1-\beta)/2}, \tau^{(b)}_{(1+\beta)/2})$'s that satisfy $L_1 \le \tau^{(b)}_{(1-\beta)/2} \le \tau^{(b)}_{(1+\beta)/2} \le L_2$.

(3) (L_1, L_2) obtained in step (b) can serve as a $(\beta, 1-\alpha)$ two-sided tolerance interval.

More Simulation Results for Upper Prediction Limits for at Least p of m Measurements at Each of r Locations

In the main text, we have presented the simulation results for the simultaneous prediction limits when $r = 5$ locations are considered. Here, Table 11 and 12 further present the coverage probabilities when $r = 1$ and $r = 10$, respectively. As expected, the coverage probabilities based on our proposed method are very close to the nominal values in all the parameter settings.

References

AKSOY, H. (2000). "Use of Gamma Distribution in Hydrological Analysis". *Turkish Journal of Engineering and Environmental Sciences* 24(6), pp. 419–428.

ARYAL, S.; BHAUMIK, D.; MATHEW, T.; and GIBBONS, R. (2007). "Approximate Tolerance Limits and Prediction Limits for the Gamma Distribution". *Journal of Applied Statistical Science* 16, pp. 103–111.

ASHKAR, F. and OUARDA, T. B. M. J. (1998). "Approximate Confidence Intervals for Quantiles of Gamma and Generalized Gamma Distributions". *Journal of Hydrologic Engineering* 3(1), pp. 43–51.

ATAPATTU, S.; TELLAMBURA, C.; and JIANG, H. (2011). "A Mixture Gamma Distribution to Model the SNR of Wireless Channels." *IEEE Transactions on Wireless Communications* 10(12), pp. 4193–4203.

AVRAMIDIS, A. N.; DESLAURIERS, A.; and L'ECUYER, P. (2004). "Modeling Daily Arrivals to a Telephone Call Center." *Management Science* 50(7), pp. 896–908.

BAIN, L. and ENGELHARDT, M. (1975). "A Two-Moment Chi-Square Approximation for the Statistic $\log(\bar{x}/\tilde{x})$". *Journal of the American Statistical Association* 70(352), pp. 948–950.

BAIN, L.; ENGELHARDT, M.; and SHIUE, W.-K. (1984). "Approximate Tolerance Limits and Confidence Limits on Reliability for the Gamma Distribution". *IEEE Transactions on Reliability* 33(2), pp. 184–187.

BHAUMIK, D. K. and GIBBONS, R. D. (2006). "One-Sided Approximate Prediction Intervals for at Least p of m Observations from a Gamma Population at Each of r Locations". *Technometrics* 48 (1), pp. 112–119.

BIRNBAUM, Z. and MCCARTY, R. (1958). "A Distribution-Free Upper Confidence Bound for $\Pr(Y < X)$, Based on Independent Samples of X and Y". *The Annals of Mathematical Statistics* pp. 558–562.

CHEE, M.; YANG, R.; HUBBELL, E.; BERNO, A.; HUANG, X. C.; STERN, D., WINKLER, J.; LOCKHART, D. J.; MORRIS, M. S.; and FODOR, S. P. (1996). "Accessing Genetic Information with High-Density DNA Arrays". *Science* 274(5287), pp. 610–614.

CHEN, P. and YE, Z.-S. (2016). "Estimation of Field Reliability Based on Aggregate Lifetime Data". *Technometrics*, to appear.

COIT, D. W. and JIN, T. (2000). "Gamma Distribution Parameter Estimation for Field Reliability Data with Missing Failure Times". *IIE Transactions* 32(12), pp. 1161–1166.

DAVIS, C. B. and MCNICHOLS, R. J. (1987). "One-Sided Intervals for at Least p of m Observations from a Normal Population on Each of r Future Occasions". *Technometrics* 29(3), pp. 359–370.

FAN, T.-H. and YU, C.-H. (2013). "Statistical Inference on Constant Stress Accelerated Life Tests Under Generalized Gamma Lifetime Distributions". *Quality and Reliability Engineering International* 29(5), pp. 631–638.

FERNÁNDEZ, A. J. (2010). "Tolerance Limits for k-out-of-n Systems with Exponentially Distributed Component Lifetimes". *IEEE Transactions on Reliability* 59(2), pp. 331–337.

FISHER, R. A. (1930). "Inverse Probability". In *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 26, pp. 528–535. Cambridge, UK: Cambridge University Press.

GIBBONS, R. D.; BHAUMIK, D. K.; and ARYAL, S. (2009). *Statistical Methods for Groundwater Monitoring*, vol. 59. Hoboken, NJ: John Wiley & Sons.

GIBBONS, R. D.; BHAUMIK, D. K.; COX, D. R.; GRAYSON, D. R.; DAVIS, J. M.; and Sharma, R. P. (2005). "Sequential Prediction Bounds for Identifying Differentially Expressed Genes in Replicated Microarray Experiments". *Journal of Statistical Planning and Inference* 129(1), pp. 19–37.

GUO, H. and KRISHNAMOORTHY, K. (2004). "New Approximate Inferential Methods for the Reliability Parameter in a Stress-Strength Model: The Normal Case". *Communications in Statistics – Theory and Methods* 33(7), pp. 1715–1731.

HANNIG, J. (2009). "On Generalized Fiducial Inference". *Statistica Sinica* 19(2), pp. 491–554.

HANNIG, J.; IYER, H.; and PATTERSON, P. (2006). "Fiducial Generalized Confidence Intervals". *Journal of the American Statistical Association* 101(473), pp. 254–269.

HUSAK, G. J.; MICHAELSEN, J.; and FUNK, C. (2007). "Use of the Gamma Distribution to Represent Monthly Rainfall in Africa for Drought Monitoring Applications". *International Journal of Climatology* 27(7), pp. 935–944.

KRISHNAMOORTHY, K. and MATHEW, T. (2004). "One-Sided Tolerance Limits in Balanced and Unbalanced One-Way Random Models Based on Generalized Confidence Intervals". *Technometrics* 46(1), pp. 44–52.

KRISHNAMOORTHY, K.; MATHEW, T.; and MUKHERJEE, S. (2008). "Normal-Based Methods for a Gamma Distribution". *Technometrics* 50(1), pp. 69–78.

LAWLESS, J. F. (2002). *Statistical Models and Methods for Lifetime Data*. Hoboken, NJ: John Wiley & Sons.

LEE, H.-I. and LIAO, C.-T. (2012). "Estimation for Conformance Proportions in a Normal Variance Components Model." *Journal of Quality Technology* 44(1), pp. 63–79.

LIU, C.; MARTIN, R.; and SYRING, N. (2013). "Simulating from a Gamma Distribution with Small Shape Parameter". arXiv preprint arXiv:1302.1884.

REISER, B. (2000). "Measuring the Effectiveness of Diagnostic Markers in the Presence of Measurement Error Through the Use of ROC Curves". *Statistics in Medicine* 19(16), pp. 2115–2129.

REISER, B. and GUTTMAN, I. (1986). "Statistical Inference for $\Pr(Y < X)$: The Normal Case". *Technometrics* 28(3), pp. 253–257.

RYAN, T. P. (2011). *Statistical Methods for Quality Improvement*. Hoboken, NJ: John Wiley & Sons.

STEPHENSON, D.; KUMAR, K. R.; DOBLAS-REYES, F.; ROYER, J.; CHAUVIN, F.; and PEZZULLI, S. (1999). "Extreme Daily Rainfall Events and Their Impact on Ensemble Forecasts of the Indian Monsoon". *Monthly Weather Review* 127(9), pp. 1954–1966.

WANG, B. X. and YE, Z.-S. (2015). "Inference on the Weibull Distribution Based on Record Values". *Computational Statistics & Data Analysis* 83, pp. 26–36.

WANG, C.; HANNIG, J.; and IYER, H. K. (2012). "Fiducial Prediction Intervals". *Journal of Statistical Planning and Inference* 142(7), pp. 1980–1990.

WEERAHANDI, S. (1993). "Generalized Confidence Intervals". *Journal of the American Statistical Association* 88(423), 899.

WEERAHANDI, S. and JOHNSON, R. A. (1992). "Testing Reliability in a Stress-Strength Model when X and Y Are Normally Distributed". *Technometrics* 34(1), pp. 83–91.

YE, Z.-S. and CHEN, N. (2016). "Closed-Form Estimators for the Gamma Distribution Derived from Likelihood Equations". *The American Statistician*, to appear.

ZABELL, S. L. (1992). "RA Fisher and Fiducial Argument". *Statistical Science* 7(3), pp. 369–387.