P-ISSN 2586-9000
E-ISSN 2586-9027
Homepage : *https://tci-thaijo.org/index.php/SciTechAsia* Science & Technology Asia
Vol.28 No.3 July - September 2023 Page: [44-58]
Original research article

Methods for Testing the Rainfall Dispersion Data Fitted to a Gamma Distribution of Songkhla, Thailand

Wararit Panichkitkosolkul*

Department of Mathematics and Statistics, Faculty of Science and Technology, Thammasat University, Pathum Thani 12120, Thailand

Received 9 April 2023; Received in revised form 6 July 2023
Accepted 11 September 2023; Available online 26 September 2023

ABSTRACT

Floods are ordinary natural disasters in Songkhla, Thailand that occur almost every year due to heavy rainfall during the rainy season. The organizations related to agriculture and water resource planning can solve or reduce the flood problem in Songkhla by testing the monthly rainfall dispersion for flood monitoring. We can use the coefficient of variation (CV) to measure the dispersion of rainfall amount in different areas because the rainfall amount varies greatly depending on the region and season. The objective of this study is to propose two methods for testing the CV for a gamma distribution. The methods based on Score-type and Wald-type confidence intervals were applied for testing the monthly rainfall dispersion which these data were well-fitted with a gamma distribution. Using Monte Carlo simulations, the type I error rate and power of the test for two test statistics were estimated under several shape parameter values in a gamma distribution. The simulation results indicated that the test statistic based on Wald-type confidence interval performed better than its competitor in terms of the attained nominal significance level (0.05). The mean of the absolute differences between the empirical type I error rates and 0.05 for those based on Score-type and Wald-type confidence intervals was 0.0480 and 0.0067, respectively. Therefore, the test statistic based on Wald-type confidence interval is recommended for analysis in similar scenarios. Both test statistics were utilized to test the rainfall dispersion data from Sa Dao Meteorological Station in Songkhla. The study concluded that the population CV of the Songkhla's rainfall was not significantly different from the hypothesized setting of 0.90.

Keywords: Measure of dispersion; Power of the test; Skewed distribution; Test statistics; Type I error rate

***Corresponding author:** wararit@mathstat.sci.tu.ac.th
doi: 10.14456/scitechasia.2023.45

1. Introduction

Since damage from natural disasters has increased due to anomalous global climate changes, researchers are interested in studying their occurrences. Thailand has been divided into six geographical regions by the National Research Council: north, northeast, central, east, west, and south, many of which are inclined to the seasonal flash flood in the north, northeast, and south regions [1]. The worst flooding occurred in Thailand in 2011 and caused significant damage to agriculture sector in the 26 provinces in the central, north and northeast regions [2]. Additionally, it killed more than 800 people, destroyed 2 million dwellings, extremely affected the lives of millions, and caused economic losses to reach about US$ 45.5 billion [3].

Songkhla, one of the southern coastal provinces of Thailand, is faced with flooding every year. Fig.1 shows the map of Songkhla. Several coastal cities are increasingly at risk of flooding caused by extreme weather conditions [4]. Tropical cyclones tend to cause violent flooding in coastal cities[5, 6]. In coastal cities, flooding is main caused by two processes: surface runoff on account of inland severe rainfall and tidal flooding during extremely high tide events [7–10]. On November 24, 2020, Songkhla's four districts: Sa Dao, Hat Yai, Muang Songkhla, and Na Mom were inundated with heavy rain for several days. Although the authorities tried to drain the water, heavy flooding blocked access to some 300 households and many villages. Later, on December 1–2, 2020, the worst flooding in 12 years occurred in Songkhla. Prior to that, in 2005, there was major flooding in Songkhla [11]; water from the Songkhla Lake and Patthalung province caused a wide range of effects on all affected areas, 86 schools in the province were faced with flooding with 62 of them being temporarily closed down, Hat Yai Airport was flooded, and crops in farmers' fields were damaged. The uncontrollable natures of rainfall are changing over time affecting agricultural sustainability of Songkhla.

To show the variation of rainfall amounts within the months, the accumulated rainfall amounts over a sliding 31-day period centered around each day of the year is shown in Fig.2. Songkhla experiences extreme seasonal variation in monthly rainfall. Songkhla receives its most monthly rainfall in November with an average rainfall of approximately 12.0 inches (304.8 millimeters) while the month with the least rainfall amounts is February, with an average rainfall amount of approximately 0.9 inches (22.9 millimeters) [12]. The coefficient of variation (CV) can be used to demonstrate rainfall dispersion in different regions because the rainfall amount varies greatly depending on the season and region. The CV is a statistical measure of variability relative to the mean of the population [13], which represents the proportion of the population standard deviation σ to the population mean μ, i.e., $\theta = \sigma/\mu$, where $\mu \neq 0$. Without units, the CV is more commonly used than the standard deviation, mean deviation and the variance for the variation comparison in variables with different units.

The estimation of CV is widely used in several fields such as medicine, science, engineering, business, economics and others (see [14]). For instance, Faber and Korn [15] applied the CV for analyzing synaptic plasticity. The correlation between spatial and temporal global solar radiation data was estimated based on the CV [16]. Reed, Lynn and Meade [17] applied the CV in evaluating the variation of quantitative as-

Fig. 1. The map of Songkhla, Thailand.

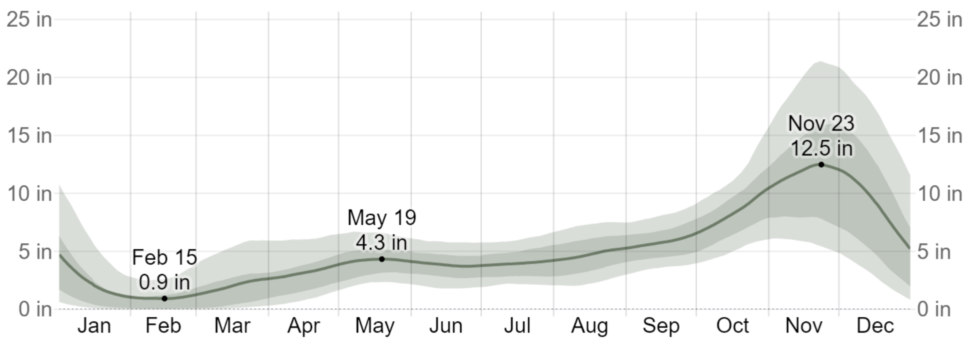

Fig. 2. The average rainfall (solid line) accumulated over the course of a sliding 31-day period centered on the day in question, with 25th to 75th and 10th to 90th percentile bands.

says. The diversity comparison in work groups using the CV was studied by Bedeian and Mossholder [18]. Kang, Lee, Seong and Hawkins [19] applied the CV for monitoring variability in statistical process control. A novel approach for monitoring the square of CV with two one-sided exponentially weighted moving average (EWMA) control charts was proposed by Castagliola, Celano and Psarakis [20]. Döring and Reckling [22] introduced an approach for the adjustment in the CV because of the methodical dependency of population variance from the population mean.

Frederick and Kut [21] studied the outcome of the CV to the operation times on the optimum allocation of storage spaces in production line procedures. Bakowskia, Radziszewskia and Žmindak [23] evaluated the injection pressure recorded throughout the time of the performance of the engine by using the CV. When the CV is known, Abu-Dayyeh, Al-Rawi and Alodat [24] discussed the interval estimation for the normal mean. The optimal sampling plan via the Fisher information numbers for normal distribution with known CV was indicated by Alodat and Omari [25]. Wang et al. [27] calculated the CV of the total amount of rain (rainfall depth) of all catchment areas in each day. Lastly, Addisu et al. [26] studied the variation of the annual total rainfall data in Ethiopia by using the CV.

From the literature review, the research on the construction of testing the CV for a gamma distribution is not much. However, there are some approaches for estimating the confidence interval for the CV. For example, McKay [28] introduced the interval estimation for the CV which depends on the chi-square distribution. The modified McKay's approximate confidence interval for the CV was proposed by Vangel [29]. For a normal distribution, the modified McKay's approximated confidence interval by the replacement of the classical sample estimator of CV with the maximum likelihood estimate was discussed by Panichkitkosolkul [30]. Albatineh et al. [31] applied a ranked set sampling method to estimate the interval estimation for the CV. For a gamma distribution, Sangnawakij and Niwitpong [32] proposed Score-type and Wald-type confidence intervals for the CV. We can apply the confidence intervals proposed Sangnawakij and Niwitpong [32] to test the hypothesis for the CV.

There are several research works on the confidence intervals for the CV, differences between the CVs, the common CV and their applications with the amount of rainfall. For example, Yosboonruang et al. [34] used the Bayesian approaches for constructing the confidence intervals for the CV of a delta-lognormal distribution. For the delta-lognormal distributions, simultaneous confidence intervals for all pairwise differences of the CVs were constucted by Yosboonruang et al. [35]. Later, Yosboonruang et al. [36] employed the fiducial generalized confidence interval, equal-tailed Bayesian credible intervals using the independent Jeffreys or uniform priors, and the method of variance estimates recovery to estimate the confidence intervals for the common CV of delta-lognormal distributions.

In some situations, we consider the testing on the monthly rainfall dispersion in terms of the CV. Recently, three statistics for testing the ratio of the CVs in the inverse gamma distributions by using the fiducial quantities (FQ) and two Bayesian approaches were proposed by Panichkitkosolkul [33]. To the best of our knowledge, there is no research conducted on testing the CV of a gamma distribution based on Score-type and Wald-type confidence intervals. Methods for testing on the CV could be advantageous for the Thailand's Ministry of Agriculture and Cooperatives and Ministry of Natural Resources and Environment, and other related organizations in order to solve the Thailand's flood problem, including the water resource planning. Based on the gamma quantile-quantile (Q-Q) plot (Fig.3(a)), a histogram (Fig.3(b)), and the Anderson-Darling goodness-of-fit test, they showed that the monthly rainfall amounts from Sa Dao Meteorological Station, located at latitude $6^\circ\ 4'\ 23.74''$ North and longitude $100^\circ\ 24'\ 35.47''$ East, from

January 2016 to December 2018 seems to fit well to a gamma distribution.

The objective of this study is to propose the methods for testing the CV in a gamma distribution and examine the efficacy of the proposed methods. For the practitioners, the larger value of CV indicates greater variation, whereas the lower value of CV indicates a lower risk [1]. This can further assist in planning for the impacts of flooding as a result of these climatic changes. This study concentrates on two confidence intervals for the CV proposed by Sangnawakij and Niwitpong [32]. Due to the unavailability of a direct theoretical comparison, we designed and conducted a Monte Carlo simulation to compare the performance of these methods.

2. Point Estimation of Parameters in a Gamma Distribution

The point estimates of shape and scale parameters for a gamma distribution are explained in this section. Let $X = (X_1, ..., X_n)$ be a random sample from the gamma distribution with the shape parameter α and scale parameter β, denoted as Gamma(α, β). Eq. (2.1) is the probability density function of X,

$$f(x; \alpha, \beta) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta}, \quad (2.1)$$

where $x > 0, \alpha > 0, \beta > 0$.

The mean and variance of X are $E(X) = \alpha\beta$ and $Var(X) = \alpha\beta^2$, respectively [37]. Therefore, the CV of X is given by $\theta = 1/\sqrt{\alpha}$.

Here, the maximum likelihood (ML) estimators for α and β were considered. From the probability density function shown in Eq. (2.1), the log-likelihood function of parameters α and β is as follows:

$$\ln L(\alpha, \beta) = -\frac{\sum_{i=1}^{n} X_i}{\beta} + (\alpha - 1) \sum_{i=1}^{n} \ln(X_i) - n \ln(\Gamma(\alpha)) - n\alpha \ln(\beta).$$

The Score-type function is obtained by partial differentiating the log-likelihood function with respect to α and β, respectively,

$$U(\alpha, \beta) = \begin{bmatrix} \sum_{i=1}^{n} \ln(X_i) - n \ln(\alpha) + \frac{n}{2\alpha} - n \ln(\beta) \\ \frac{\sum_{i=1}^{n} X_i}{\beta^2} - \frac{n\alpha}{\beta} \end{bmatrix}.$$

Then, the ML estimators can be derived for α and β, respectively,

$$\hat{\alpha} = \frac{1}{2\left(\ln(\bar{X}) - \frac{\sum_{i=1}^{n} \ln(X_i)}{n}\right)}, \quad \hat{\beta} = \frac{\bar{X}}{\hat{\alpha}},$$

where $\bar{X} = n^{-1} \sum_{i=1}^{n} X_i$ denotes the sample mean of X. Subsequently, the ML estimator of CV is given by $\hat{\theta} = 1/\sqrt{\hat{\alpha}}$.

3. Proposed Test Statistics

Let $X_1, ..., X_n$ be an independent and identically distributed (i.i.d.) random sample of size n from a population with finite mean, μ and finite variance, σ^2. In this study, the goal of testing is to determine the likelihood that a population CV of a gamma distribution is likely to be true. The null and alternative hypotheses, H_0 and H_1, for the test are written in mathematical symbols as follows:

$$H_0 : \theta = \theta_0$$

$$H_a : \theta \neq \theta_0.$$

Two methods for testing the CV based on Score-type and Wald-type confidence intervals are introduced in this section.

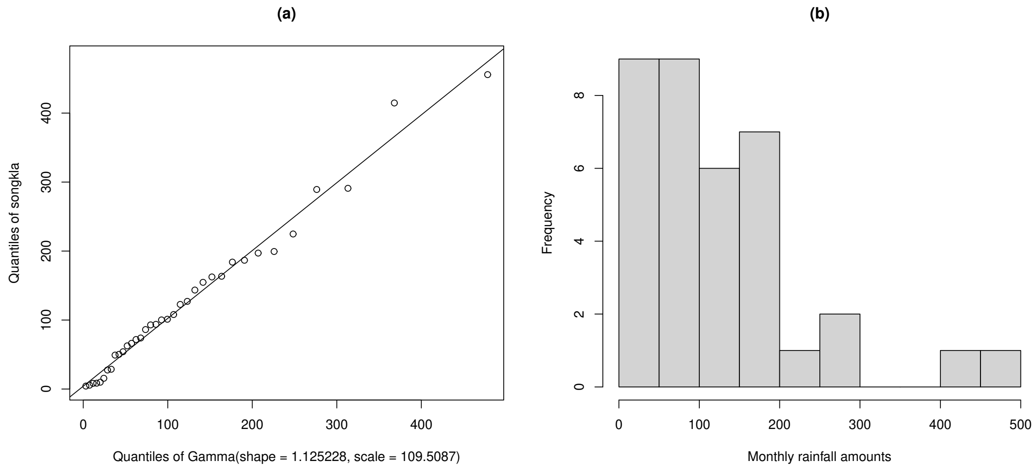

Fig. 3. (a) gamma QQ plot (b) histogram of the monthly rainfall amounts from Sa Dao Meteorological Station, Songkhla, Thailand.

3.1 Score-type method

Let α be the parameter of interest and β be the nuisance parameters. The Score-type statistic [38, 39] is defined as

$$W_1 = U^T(\alpha_0, \hat{\beta}_0) I^{-1}(\alpha_0, \hat{\beta}_0) U(\alpha_0, \hat{\beta}_0),$$

where $\hat{\beta}_0$ is the ML estimator for β, under the null hypothesis $H_0' : \alpha = \alpha_0$, $U(\alpha_0, \hat{\beta}_0)$ is the vector of the Score function and $I(\alpha_0, \hat{\beta}_0)$ is the matrix of the Fisher information. The Score function under H_0' is derived as follows:

$$U(\alpha_0, \hat{\beta}_0) = \begin{bmatrix} \sum_{i=1}^{n} \ln(X_i) + \frac{n}{2\alpha_0} - n\ln(\bar{X}) \\ 0 \end{bmatrix}.$$

The inverse of the Fisher information matrix can be found as follows:

$$I^{-1}(\alpha_0, \hat{\beta}_0) = \begin{bmatrix} \frac{2\alpha_0^2}{n} & -\frac{2\bar{X}}{n} \\ -\frac{2\bar{X}}{n} & \frac{\bar{X}^2(2\alpha_0+1)}{n\alpha_0^3} \end{bmatrix}.$$

The pivotal quantity was constructed by using the property of the Score-type function as Eq. (3.1),

$$Z_{Score} = \sqrt{\frac{2\alpha_0^2}{n}} \left(\sum_{i=1}^{n} \ln(X_i) + \frac{n}{2\alpha_0} - n\ln(\bar{X}) \right), \quad (3.1)$$

converges to the standard normal distribution as $n \to \infty$ (converges in distribution). Since the variance of $\hat{\alpha}$ is $\frac{2\alpha_0^2}{n}$, it is approximated by substituting $\hat{\alpha}$ in its variance. Under H_0', the statistic in Eq. (3.1) is given as

$$Z_{Score} \cong \sqrt{\frac{2\hat{\alpha}^2}{n}} \left(\sum_{i=1}^{n} \ln(X_i) + \frac{n}{2\alpha} - n\ln(\bar{X}) \right).$$

From the probability statement, $1 - \gamma = P(-Z_{1-\gamma/2} \le Z_{Score} \le Z_{1-\gamma/2})$, it can be written in another form as follows: $1 - \gamma = P(l_s \le \theta \le u_s)$. Therefore, the $(1-\gamma)100\%$ Score-type confidence interval for θ, CI_S, is given by

$$CI_S = [l_S, u_S] = \left[\sqrt{\frac{2}{n}\left(M - Z_{1-\gamma/2}\sqrt{\frac{n}{2\hat{\alpha}^2}} \right)}, \sqrt{\frac{2}{n}\left(M + Z_{1-\gamma/2}\sqrt{\frac{n}{2\hat{\alpha}^2}} \right)} \right],$$

where $M = n\ln(\bar{X}) - \sum_{i=1}^{n}\ln(X_i)$ and $Z_{1-\gamma/2}$ is the $(1-\gamma/2)^{\text{th}}$ quantile of the standard normal distribution. Therefore, the null hypothesis, $H_0 : \theta = \theta_0$, will be rejected if

$$\theta_0 < \sqrt{\frac{2}{n}\left(M - Z_{1-\gamma/2}\sqrt{\frac{n}{2\hat{\alpha}^2}}\right)} \quad \text{or}$$

$$\theta_0 > \sqrt{\frac{2}{n}\left(M + Z_{1-\gamma/2}\sqrt{\frac{n}{2\hat{\alpha}^2}}\right)}.$$

3.2 Wald-type method

It is well known that the Wald-type statistic is asymptotic. This statistic can be derived from the property of the ML estimator [40]. Under the null hypothesis $H_0' : \alpha = \alpha_0$, the general form of the Wald-type statistic is defined as

$$W_2 = (\hat{\alpha} - \alpha_0)^T \left[I^{\alpha\alpha}(\hat{\alpha}, \hat{\beta})\right]^{-1} (\hat{\alpha} - \alpha_0),$$

where $I^{\alpha\alpha}(\hat{\alpha}, \hat{\beta})$ is the estimated variance of $\hat{\alpha}$ obtained from the first row and the first column of $I^{-1}(\hat{\alpha}, \hat{\beta})$. Using the information of partial derivatives from the previous subsection, the inverse matrix is given by

$$I^{-1}(\hat{\alpha}, \hat{\beta}) = \begin{bmatrix} \frac{2\hat{\alpha}^2}{n} & -\frac{2\bar{X}}{n} \\ -\frac{2\bar{X}}{n} & \frac{\bar{X}^2(2\hat{\alpha}+1)}{n\hat{\alpha}^3} \end{bmatrix},$$

where $I^{\alpha\alpha}(\hat{\alpha}, \hat{\beta}) = \frac{2\hat{\alpha}^2}{n}$. Therefore, under H_0', we obtain the Wald-type statistic

$$Z_{Wald} \cong \sqrt{\frac{n}{2\hat{\alpha}^2}}(\hat{\alpha} - \alpha),$$

which has the limiting distribution of a standard normal distribution. Therefore, the $(1-\gamma)100\%$ Wald-type confidence interval for θ, CI_W, is as follows:

$$CI_W = [l_W, u_W] = \left[1/\sqrt{\hat{\alpha} + Z_{1-\gamma/2}\sqrt{\frac{2\hat{\alpha}^2}{n}}},\; 1/\sqrt{\hat{\alpha} - Z_{1-\gamma/2}\sqrt{\frac{2\hat{\alpha}^2}{n}}}\right],$$

where $Z_{1-\gamma/2}$ is the $(1-\gamma/2)^{\text{th}}$ quantile of the standard normal distribution. Therefore, the null hypothesis, $H_0 : \theta = \theta_0$, will be rejected if

$$\theta_0 < 1/\sqrt{\hat{\alpha} + Z_{1-\gamma/2}\sqrt{\frac{2\hat{\alpha}^2}{n}}} \quad \text{or}$$

$$\theta_0 > 1/\sqrt{\hat{\alpha} - Z_{1-\gamma/2}\sqrt{\frac{2\hat{\alpha}^2}{n}}}.$$

4. Simulation Study

For a gamma distribution, two methods for testing the CV based on Score-type and Wald-type confidence intervals are compared in this section. Due to the unavailability of a direct theoretical comparison, a Monte Carlo simulation study was designed and conducted using R [41] version 4.2.3 to compare the performance of two proposed test statistics. They were evaluated in terms of the accomplishment of empirical type I error rates and the powers of the test. We count the number of times for each test that the null hypothesis was rejected when H_0 was true, to obtain the empirical type I error rates. In addition, the number of times for each test, that the null hypothesis was rejected when H_0 was not true, was counted to obtain the power of the test. Since the simulation results provided the similar conclusions obtained for other values of significant level γ, the simulation results are reported only for $\gamma = 0.05$.

The sample sizes n = 10, 30, 50, 100 and 200 are employed to present small to large samples. The data were generated from a gamma distribution with $\beta = 2$ and we adjusted the value of α in order to receive the desired value of CV, θ. The values of θ are equal to 0.05, 0.10, 0.20, 0.28, 0.30 and 0.33. The number of simulations was set at 10,000.

The simulation results are summarized in Tables 1–6. For all sample sizes, the test statistic based on Wald-type confidence interval provided the empirical type I error rates close to the nominal significance level of 0.05. On the other hand, the empirical type I error rates of test statistic based on Score-type confidence interval were close to the nominal level of significance when sample sizes were large. Furthermore, we found that the statistic based on Score-type confidence interval had a high empirical type I error rate for small sample sizes. In terms of power of the test, the test statistic based on Score-type confidence interval outperformed the competition for $\theta > \theta_0$. On the other side, for $\theta < \theta_0$, the test statistic based on Wald-type confidence interval performed better. As sample size increased, the chance of empirical type I error rate of the test statistic based on Score-type confidence interval decreased and approached 0.95. However, as sample size increased, empirical type I error rate of the test statistic based on Wald-type confidence interval increased and approached 0.95. Furthermore, as the difference between the true value of CV and the hypothesized value of CV increased, the power of the test increased. For large sample sizes, it can be shown that both test statistics performed similarly well in terms of the empirical type I error rates and the powers of the test. However, both criteria of two proposed test statistics are notably different when sample sizes are small.

5. Application to Rainfall Dispersion Data

In this section, the real application of two test statistics for the CV of a gamma distribution is illustrated. This study used the monthly rainfall amounts (mm) from the Sa Dao Meteorological Station from January 2016 to December 2018 recorded by the Thai Meteorological Department. The descriptive statistics are calculated and presented as follows: sample size = 36, arithmetic mean = 123.22, standard deviation = 108.76 and CV = 0.88.

Fig.3 displays a gamma quantile-quantile (QQ) plot for monthly rainfall amounts versus fitted gamma distribution and a histogram of the monthly rainfall amounts. The density and Box-Whisker plots are shown in Fig.4. From Figs.3 and 4, we found that the monthly rainfall amounts has a right-skewed distribution. By using the Anderson-Darling (AD) test [42] via Minitab [43], the AD statistic was 0.333 and p-value > 0.25. It was found that the monthly rainfall amount fitted well to a gamma distribution with shape parameter $\hat{\alpha}$ = 1.1252 and scale parameter $\hat{\beta}$ = 109.5086, while the estimator of the CV was $\hat{\theta} = 1/\sqrt{\hat{\alpha}}$ = 0.9427. A researcher believes that the population CV of the monthly rainfall amounts is equal to 0.90. It can be written in the null and alternative hypotheses as:

$$H_0 : \theta = 0.90$$

$$H_a : \theta \neq 0.90.$$

The Score-type and Wald-type confidence intervals for the CV were calculated and are shown in Table 7. These lower and upper limits can be used as the critical values of the test statistics. There is no sufficient evidence in support of the alternative hypothesis, and therefore the null hypothesis cannot be rejected because $0.7380 \leq \theta_0 \leq 1.2166$ and $0.8321 \leq \theta_0 \leq 1.3717$ using test statistics based on Score-type confidence interval and Wald-type confidence interval, respectively. Therefore, this study concluded that the researcher's belief is correct. Hence, the population CV of the monthly rainfall

Table 1. The Empirical type I error rates (bold numeric) and powers of tests for Gamma(400,2), θ = 0.05.

n	Method	θ_0								
		0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
10	Score-type	0.9550	0.1588	0.0072	0.0548	**0.1955**	0.4325	0.6863	0.8652	0.9562
	Wald-type	0.9998	0.9640	0.6678	0.2091	**0.0286**	0.0059	0.0160	0.0432	0.0841
30	Score-type	1.0000	0.9999	0.8209	0.1068	**0.1057**	0.5182	0.8970	0.9921	0.9998
	Wald-type	1.0000	1.0000	0.9714	0.4715	**0.0417**	0.1496	0.5358	0.8782	0.9850
50	Score-type	1.0000	1.0000	0.9887	0.3471	**0.0809**	0.6283	0.9759	0.9997	1.0000
	Wald-type	1.0000	1.0000	0.9979	0.6604	**0.0446**	0.3162	0.8497	0.9946	1.0000
100	Score-type	1.0000	1.0000	1.0000	0.7646	**0.0658**	0.8432	0.9998	1.0000	1.0000
	Wald-type	1.0000	1.0000	1.0000	0.8949	**0.0483**	0.6660	0.9967	1.0000	1.0000
200	Score-type	1.0000	1.0000	1.0000	0.9811	**0.0595**	0.9786	1.0000	1.0000	1.0000
	Wald-type	1.0000	1.0000	1.0000	0.9941	**0.0493**	0.9440	1.0000	1.0000	1.0000

Table 2. The Empirical type I error rates (bold numeric) and powers of tests for Gamma(100,2), θ = 0.10.

n	Method	θ_0								
		0.06	0.07	0.08	0.09	0.10	0.11	0.12	0.13	0.14
10	Score-type	0.0093	0.0213	0.0545	0.1149	**0.1992**	0.3027	0.4245	0.5631	0.6864
	Wald-type	0.6564	0.4143	0.2137	0.0865	**0.0303**	0.0104	0.0054	0.0116	0.0159
30	Score-type	0.8271	0.4328	0.1117	0.0374	**0.1052**	0.2704	0.5157	0.7491	0.8934
	Wald-type	0.9718	0.8207	0.4807	0.1599	**0.0450**	0.0583	0.1547	0.3219	0.5362
50	Score-type	0.9885	0.8120	0.3453	0.0547	**0.0851**	0.2988	0.6249	0.8778	0.9746
	Wald-type	0.9987	0.9515	0.6577	0.2180	**0.0441**	0.0995	0.3108	0.6127	0.8588
100	Score-type	1.0000	0.9936	0.7572	0.1763	**0.0679**	0.3919	0.8291	0.9861	0.9998
	Wald-type	1.0000	0.9986	0.8955	0.3536	**0.0518**	0.2084	0.6447	0.9408	0.9974
200	Score-type	1.0000	1.0000	0.9842	0.4446	**0.0529**	0.5675	0.9761	1.0000	1.0000
	Wald-type	1.0000	1.0000	0.9939	0.5982	**0.0483**	0.4133	0.9423	0.9993	1.0000

Table 3. The Empirical type I error rates (bold numeric) and powers of tests for Gamma(25,2), θ = 0.20.

n	Method	θ_0								
		0.16	0.17	0.18	0.19	0.20	0.21	0.22	0.23	0.24
10	Score-type	0.0533	0.0784	0.1101	0.1450	**0.1982**	0.2465	0.3048	0.3635	0.4219
	Wald-type	0.2153	0.1390	0.0887	0.0499	**0.0319**	0.0176	0.0114	0.0059	0.0065
30	Score-type	0.1152	0.0461	0.0389	0.0567	**0.1045**	0.1730	0.2690	0.3839	0.5055
	Wald-type	0.4805	0.2978	0.1704	0.0841	**0.0434**	0.0377	0.0563	0.0913	0.1460
50	Score-type	0.3471	0.1577	0.0616	0.0451	**0.0812**	0.1554	0.2983	0.4545	0.6232
	Wald-type	0.6686	0.4374	0.2276	0.0990	**0.0474**	0.0464	0.0981	0.1817	0.3025
100	Score-type	0.7837	0.4739	0.1862	0.0558	**0.0656**	0.1743	0.3814	0.6244	0.8233
	Wald-type	0.9056	0.6897	0.3755	0.1379	**0.0487**	0.0743	0.1966	0.4082	0.6453
200	Score-type	0.9849	0.8553	0.4755	0.1176	**0.0535**	0.2109	0.5550	0.8507	0.9716
	Wald-type	0.9934	0.9239	0.6267	0.2136	**0.0495**	0.1222	0.4049	0.7412	0.9357

amounts are not different from 0.90 at the significance level of 0.05. It is interpreted that the value of CV is 0.90 times of the population mean.

6. Discussion

The objective of this study is to propose two test statistics based on Score-type and Wald-type confidence intervals for the

Table 4. The Empirical type I error rates (bold numeric) and powers of tests for Gamma(12.76,2), θ = 0.28.

n	Method	θ_0								
		0.24	0.25	0.26	0.27	0.28	0.29	0.30	0.31	0.32
10	Score-type	0.0801	0.0999	0.1269	0.1577	**0.1914**	0.2167	0.2623	0.2968	0.3510
	Wald-type	0.1349	0.1009	0.0684	0.0414	**0.0294**	0.0209	0.0141	0.0099	0.0087
30	Score-type	0.0512	0.0397	0.0438	0.0640	**0.0944**	0.1463	0.2005	0.2711	0.3590
	Wald-type	0.2969	0.1943	0.1175	0.0732	**0.0458**	0.0390	0.0399	0.0600	0.0816
50	Score-type	0.1457	0.0705	0.0436	0.0531	**0.0740**	0.1220	0.2074	0.3011	0.4110
	Wald-type	0.4135	0.2702	0.1496	0.0812	**0.0518**	0.0408	0.0599	0.0984	0.1566
100	Score-type	0.4381	0.2438	0.0988	0.0500	**0.0598**	0.1276	0.2418	0.3922	0.5768
	Wald-type	0.6606	0.4451	0.2350	0.1113	**0.0491**	0.0547	0.1117	0.2095	0.3514
200	Score-type	0.8251	0.5524	0.2528	0.0924	**0.0529**	0.1313	0.3288	0.5765	0.8001
	Wald-type	0.9066	0.6987	0.3903	0.1601	**0.0530**	0.0739	0.2000	0.4225	0.6728

Table 5. The Empirical type I error rates (bold numeric) and powers of tests for Gamma(11.11,2), θ = 0.30.

n	Method	θ_0								
		0.26	0.27	0.28	0.29	0.30	0.31	0.32	0.33	0.34
10	Score-type	0.0863	0.1031	0.1303	0.1609	**0.1904**	0.2163	0.2585	0.2962	0.3361
	Wald-type	0.1240	0.0904	0.0637	0.0446	**0.0311**	0.0242	0.0144	0.0103	0.0078
30	Score-type	0.0406	0.0359	0.0463	0.0592	**0.0946**	0.1403	0.1962	0.2614	0.3254
	Wald-type	0.2638	0.1765	0.1128	0.0684	**0.0462**	0.0386	0.0390	0.0557	0.0739
50	Score-type	0.1307	0.0612	0.0468	0.0478	**0.0765**	0.1280	0.1914	0.2824	0.3815
	Wald-type	0.3791	0.2384	0.1444	0.0829	**0.0495**	0.0417	0.0552	0.0916	0.1453
100	Score-type	0.3787	0.2074	0.0884	0.0471	**0.0605**	0.1167	0.2259	0.3631	0.5273
	Wald-type	0.6043	0.4007	0.2230	0.1065	**0.0503**	0.0527	0.1040	0.1872	0.3102
200	Score-type	0.7659	0.5054	0.2300	0.0768	**0.0516**	0.1275	0.2898	0.5231	0.7423
	Wald-type	0.8635	0.6479	0.3643	0.1488	**0.0541**	0.0685	0.1778	0.3734	0.5973

Table 6. The Empirical type I error rates (bold numeric) and powers of tests for Gamma(9.18,2), θ = 0.33.

n	Method	θ_0								
		0.29	0.30	0.31	0.32	0.33	0.34	0.35	0.36	0.37
10	Score-type	0.0931	0.1119	0.1358	0.1571	**0.1862**	0.2050	0.2487	0.2781	0.3200
	Wald-type	0.1190	0.0866	0.0618	0.0443	**0.0341**	0.0242	0.0159	0.0091	0.0089
30	Score-type	0.0421	0.0370	0.0492	0.0641	**0.0967**	0.1416	0.1764	0.2365	0.2990
	Wald-type	0.2311	0.1599	0.0985	0.0663	**0.0497**	0.0395	0.0378	0.0497	0.0600
50	Score-type	0.1038	0.0564	0.0452	0.0477	**0.0714**	0.1126	0.1727	0.2449	0.3336
	Wald-type	0.3269	0.2116	0.1377	0.0738	**0.0468**	0.0412	0.0525	0.0709	0.1146
100	Score-type	0.3277	0.1734	0.0848	0.0448	**0.0586**	0.1077	0.1946	0.3030	0.4489
	Wald-type	0.5465	0.3513	0.2104	0.0915	**0.0550**	0.0497	0.0866	0.1456	0.2488
200	Score-type	0.6908	0.4291	0.2004	0.0711	**0.0485**	0.1067	0.2423	0.4376	0.6612
	Wald-type	0.8121	0.5799	0.3273	0.1334	**0.0556**	0.0543	0.1426	0.2898	0.5053

CV in a gamma distribution. From the Monte Carlo simulation results, the test statistic based on Wald-type confidence interval outperformed the competitor in terms of the empirical type I error rate for all situations. As expected, as the sample size increased, the power of the test also increased and the empirical type I error rates

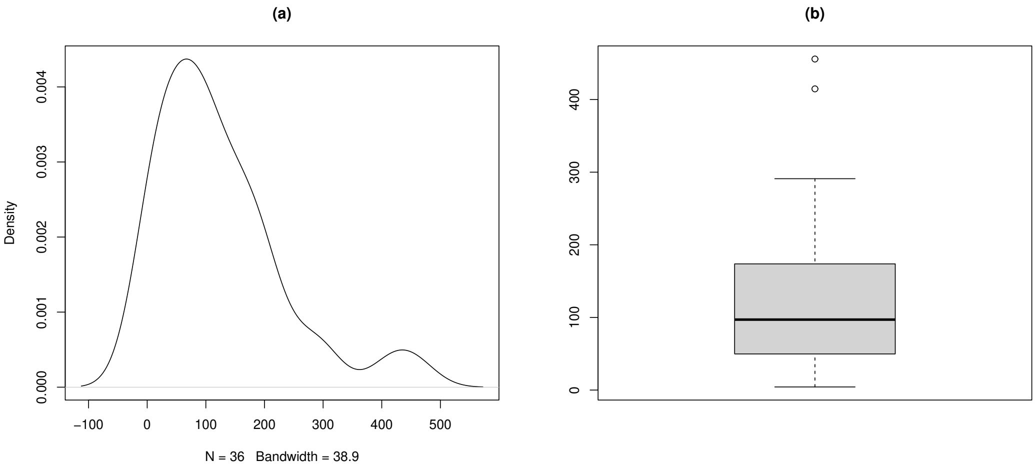

Fig. 4. (a) density plot (b) Box-Whisker plot of the monthly rainfall amounts from Sa Dao Meteorological Station.

Table 7. Critical values of test statistics based on Score-type and Wald-type confidence intervals for the significance level of 0.05

Method	Critical values	
	Lower	Upper
Score-type	0.7380	1.2166
Wald-type	0.8321	1.3717

were close to the nominal significance level of 0.05. The performance of both statistics are not considerably different in terms of the power of the test and the attainment of the nominal significance level when sample sizes are large. However, the empirical type I error rate and the power of the test were not significantly different for small sample sizes. For the inverse gamma distribution, Kaewprasert et al. [48] concluded that the interval estimator for the CV based on Wald-type method performed better challenger. The conclusions of the current study were followed with the study of Kaewprasert et al. [48].

When the statistical methods for testing the population CV was conducted, the environmental and agricultural organizations can know the rainfall dispersion and plan to prevent the heavy rainfall problem, including water resources planning. Other data sets identified potential distribution as a gamma distribution can be applied with the proposed test statistics for the CV. The gamma distribution is identified as an effective probability density function for meteorology data analysis [44–47]. In addition, the gamma distribution is often applied to model precipitation and streamflow data [49].

7. Conclusions

Due to the unavailability of a direct theoretical comparison, a simulation study was conducted via Monte Carlo simulation to evaluate the performance of test statistics. Based on the results, test statistic based on Wald-type confidence interval achieved a good result in terms of the empirical type I error rate. The test statistic based on Score-type confidence interval performed well in the sense of the power of the test when $\theta > \theta_0$. On the other hand, the test statistic based on Wald-type confidence interval performed better when $\theta < \theta_0$. In conclusion, this study would suggest the test statistic based on Wald-type confidence interval because its empirical type I error rates are close to the nominal significance level of 0.05 for all circumstances.

Songkhla relies primarily on resources that are sensitive to climate variability, including rain-fed agriculture. There may be repercussions of increased or decreased rainfall variability on current social and economic interactions throughout Songkhla. The severity of future economic, societal, and human damage can be reduced by studying rainfall dispersion. For the planning and design of crop scheduling and the design of water management in Songkhla, the rainfall dispersion was investigated from data on the monthly rainfall of Sa Dao Meteorological Station. In the current study, the test statistics for testing the CV of a gamma distribution were utilized to test the rainfall dispersion data from Sa Dao Meteorological Station in Songkhla, Thailand. The current study can make a summary that the CV of the Songkhla's rainfall was insignificantly different from the hypothesized value of 0.90. Consequently, the concerned stakeholders ought to think about specifically rainfall dispersion for developing and implementing sustainable development strategies.

The limitation of this study is the fact that the proposed methods for testing the CV can only be used for the gamma distribution. They cannot be applied to other distributions. In future work, the construction of likelihood-based, Score-type, Wald-type, and Bayesian confidence intervals for the difference between two CVs of gamma distributions will be investigated.

8. Acknowledgements

The author is grateful for the financial support provided by the Faculty of Science and Technology, Thammasat University, Thailand under Grant no. SciGR21/2563.

References

[1] Thangjai W, Niwitpong SA, Niwitpong S. Confidence intervals for the common coefficient of variation of rainfall in Thailand. Peer J 2020; 8:e10004.

[2] Rapid Assessment for Resilient Recovery and Reconstruction Planning. Thai flood 2011. Bangkok: The World Bank; 2012.

[3] Loc HH, Emadzadeh A, Park E, Nontikansak P, Deo RC. The Great 2011 Thailand flood disaster revisited: Could it have been mitigated by different dam operations based on better weather forecasts?. Environ Res. 2023; 216: 114493.

[4] Hanson S, Nicholls R, Ranger N, Hallegatte S, Corfee-Morlot J, Herweijer C, Chateau J. A global ranking of port cities with high exposure to climate extremes. Clim Change 2011; 104: 89–111.

[5] Tanim AH, Goharian E. Developing a hybrid modeling and multivariate analysis framework for storm surge and runoff interactions in urban coastal flooding. J Hydrol 2021; 595: 125670.

[6] Shen Y, Morsy MM, Huxley C, Tahvildari N, Goodall JL. Flood risk assessment and increased resilience for coastal urban watersheds under the combined impact of storm tide and heavy rainfall. J Hydrol 2019; 579: 124159.

[7] Dawson RJ, Speight L, Hall JW, Djordjevic S, Savic D, Leandro J. Attribution of flood risk in urban areas. J Hydroinformatics 2008; 10(4): 275-88.

[8] Archetti R, Bolognesi A, Casadio A, Maglionico M. Development of flood probability charts for urban drainage network in coastal areas through a simplified joint assessment approach. Hydrol Earth Syst Sci 2011; 15(10): 3115-22.

[9] Xu K, Ma C, Lian J, Bin L. Joint probability analysis of extreme precipitation and storm tide in a coastal city under changing environment. PLoS ONE 2014; 9(10): 0109341.

[10] Wahl T, Jain S, Bender J, Meyers SD, Luther ME. Increasing risk of compound flooding from storm surge and rainfall for major US cities. Nat Clim Change 2015; 5(12): 1093-7.

[11] Thairesidents. Songkhla flood, 62 schools closed [Internet]. [cited 2023 Mar 1]. Available from: https://thairesidents.com/local/songkhla-flood-62-schools-closed.

[12] Weather Spark. Climate and average weather year round in Songkhla Thailand [Internet]. [cited 2023 Mar 1]. Available from: https://weatherspark.com/y/113364/Average-Weather-in-Songkhla-Thailand-Year-Round.

[13] Albatineh AN, Boubakari I, Kibra BMG. New confidence interval estimator of the signal-to-noise ratio based on asymptotic sampling distribution. Commun Stat-Theory Methods 2017; 46(2): 574-90.

[14] Nairy KS, Rao KA. Tests of coefficients of variation of normal population. Commun Stat-Simul Comput 2003; 32(3): 641-646.

[15] Faber DS, Korn H. Applicability of the coefficient of variation method for analyzing synaptic plasticity. Biophys J 1991; 60(5): 1288-94.

[16] Calif R, Soubdhan T. On the use of the coefficient of variation to measure spatial and temporal correlation of global solar radiation. Renew Energ 2016; 88: 192-9.

[17] Reed GF, Lynn F, Meade BD. Use of coefficient of variation in assessing variability of quantitative assays. Clin Diagn Lab Immunol 2002; 9(6): 1235-1239.

[18] Bedeian AG, Mossholder KW. On the use of the coefficient of variation as a measure of diversity. Organ Res Methods 2000; 3(3): 285-97.

[19] Kang CW, Lee MS, Seong YJ, Hawkins DM. A control chart for the coefficient of variation. J Qual Technol 2007; 39(2): 151-8.

[20] Castagliola P, Celano G, Psarakis S. Monitoring the coefficient of variation using EWMA charts. J Qual Technol 2011; 43(3): 249-65.

[21] Frederick SH, Kut CS. The effect of machine breakdowns and interstage storage on the performance of production line systems. Int J Prod Res 1991; 29: 2043-55.

[22] Döring TF, Reckling M. Detecting global trends of cereal yield stability by adjusting the coefficient of variation. Eur J Agron 2018; 99: 30-6.

[23] Bakowskia A, Radziszewskia L, Žmindak M. Analysis of the coefficient of variation for injection pressure in a compression ignition engine. Procedia Eng 2017; 177: 297-302.

[24] Abu-Dayyeh WA, Al-Rawi ZR, Alodat MT, Comparison of several confidence intervals for normal distribution with known coefficient of variation. Stat e Appl 2009; 12(2): 75-85.

[25] Alodat MT, Omari DA. Estimating the normal population with known coefficient of variation using the optimal ranked set sampling scheme. Stat e Appl 2012; 10(1): 33-42.

[26] Addisu S, Selassie YG, Fissha G, Gedif B. Time series trend analysis of temperature and rainfall in lake Tana Sub-basin, Ethiopia. Environ Syst Res 2015; 4(1): 1-12.

[27] Wang H-J, Merz R, Yang S, Tarasova L, Basso S. Emergence of heavy tails in streamflow distributions: The role of spatial rainfall variability. Adv Water Resour 2023; 171: 104359,

[28] McKay AT. Distribution of the coefficient of variation and the extended t distribution. J R Stat Soc 1932; 95(4): 695-98.

[29] Vangel MG. Confidence intervals for a normal coefficient of variation. Am Stat 1996; 50(1): 21-6.

[30] Panichkitkosolkul W., Improved confidence intervals for a coefficient of variation of a normal distribution. Thail Stat 2009; 7(2): 193-99.

[31] Albatineh AN, Kibria BMG, Wilcox ML, Zogheib B. Confidence interval estimation for the population coefficient of variation using ranked set sampling. J Appl Stat 2014; 41(4): 733-51.

[32] Sangnawakij P, Niwitpong SA, Niwitpong S. Confidence intervals for the ratio of coefficients of variation of the gamma distributions. In: Huynh VN, Inuiguchi M, Demoeux T, editors. Integrated Uncertainty in Knowledge Modelling and Decision Making. IUKM 2015. Lecture Notes in Computer Science. Cham: Springer International Publishing; 2015. p. 193-203.

[33] W. Panichkitkosolkul, Testing the ratio of the coefficients of variation for the inverse gamma distributions with an application to rainfall dispersion in Thailand. NIDA Development Journal 2022; 62(2): 80-103.

[34] Yosboonruang N, Niwitpong SA, Niwitpong S. Measuring the dispersion of rainfall using Bayesian confidence intervals for coefficient of variation of delta-lognormal distribution: a study from Thailand. PeerJ 2019; 7: e7344.

[35] Yosboonruang N, Niwitpong SA, Niwitpong S. Simultaneous confidence intervals for all pairwise differences between the coefficients of variation of rainfall series in Thailand. PeerJ 2021; 9: e11651,

[36] Yosboonruang N, Niwitpong SA, Niwitpong S. Bayesian computation for the common coefficient of variation of delta-lognormal distributions with application to common rainfall dispersion in Thailand. PeerJ 2022; 10: e12858.

[37] Casella G, Berger RL. Statistical Inference. Pacific Grove: Thomson Learning; 2002.

[38] Rao CR. Large sample tests of statistical hypotheses concerning several parameters with applications to problems of estimation. Math Proc Camb 1948; 44(1): 50-7.

[39] Rao CR. Score test: Historical review and recent developments. In: Advances in ranking and selection, Multiple Comparisons, and Reliability, Birkhäuser, Boston, 2005.

[40] Gaffke N, Steyer R, von Davier AA. On the asymptotic null-distribution of the Wald statistic at singular parameter points. Stat Decis 1999; 17(4): 339-58.

[41] Ihaka R, Gentleman R. R: A language for data analysis and graphics. J Comput Graph Stat 1996; 5(3): 299-314.

[42] Banik S, Kibria BMG. Estimating the population coefficient of variation by confidence intervals. Commun Stat Simul 2011; 40(8): 1236-61.

[43] McKenzie JD. Minitab student release 14: Statistical software for education. Boston: Pearson Addison-Wesley; 2004.

[44] Ines AVM, Hansen JW. Bias correction of daily GCM rainfall for crop simulation studies. Agric For Meteorol 2006; 138(1-4): 44-53.

[45] Block PJ, Souza Filho FA, Sun L, Kwon HH. A streamflow forecasting framework using multiple climate and hydrological models. J Am Water Resour Assoc 2009; 45(4): 828-43.

[46] Piani C, Haerter JO, Coppola E. Statistical bias correction for daily precipitation in regional climate models over Europe. Theoretical and Applied Climatology 2010; 99(1): 187-92.

[47] Sangnawakij P, Niwitpong S. Interval estimation for the common coefficient of variation of Gamma distributions. Thail Stat 2020; 18(3): 340-53.

[48] Kaewprasert T, Niwitpong SA, Niwitpong S. Confidence interval for coefficient of variation of inverse gamma distributions. In: Huynh VN, Entani T, Jeenanunta C, Inuiguchi M, Yenradee P. Integrated Uncertainty in Knowledge Modelling and Decision Making. IUKM 2020. Lecture Notes in Artificial Intelligence: Integrated Uncertainty in Knowledge Modelling and Decision Making. Cham: Springer International Publishing; 2020. p. 407-18.

[49] Coe R, Stern RD. Fitting models to daily rainfall data. J Appl Meteorol 1982; 21(7), 1024–1031.