Interevent Time Distribution of Renewal Point Process, Case Study: Extreme Rainfall in South Sulawesi

Nurtiti Sunusi
Mathematics Department, Faculty of Mathematics and Natural Sciences, Hasanuddin University
Jl. Perintis Kemerdekaan Km 10 Tamalanrea, Makassar, South Sulawesi, Indonesia 90245
E-mail: nurtitisunusi@unhas.ac.id

Abstract.
The study of time distribution of occurrences of extreme rain phenomena plays a very important role in the analysis and weather forecast in an area. The timing of extreme rainfall is difficult to predict because its occurrence is random. This paper aims to determine the inter event time distribution of extreme rain events and minimum waiting time until the occurrence of next extreme event through a point process approach. The phenomenon of extreme rain events over a given period of time is following a renewal process in which the time for events is a random variable \( \tau \). The distribution of random variable \( \tau \) is assumed to be a Pareto, Log Normal, and Gamma. To estimate model parameters, a moment method is used. Consider \( R_t \) as the time of the last extreme rain event at one location is the time difference since the last extreme rainfall event. if there are no extreme rain events up to \( t_0 \), there will be an opportunity for extreme rainfall events at \( (t_0, t_0 + \delta t_0) \). Furthermore from the three models reviewed, the minimum waiting time until the next extreme rainfall will be determined. The result shows that Log Normal model is better than Pareto and Gamma model for predicting the next extreme rainfall in South Sulawesi while the Pareto model can not be used.

1. Introduction
Many phenomena in nature are random and complex, such as extreme rainfall events which can affect the occurrence of extreme weather. This is very influential on the early determination of the growing season of various agricultural products. Knowledge of the accuracy of the time distribution of rain events is a fundamental issue in meteorological hydrological studies such as the use of the surface runoff in the design of the hydraulic structures. This requires the right approximation to be able to explain the phenomenon. One stochastic model that can explain the random nature phenomena both in space and time is known by the name of the process. In the point process, each event is considered a point in a particular location [1]. Many ways of assessing point process of modeling random phenomena, for example: point process as homogeneous Poisson process, non homogeneous Poisson process, renewal process, and others.

Some definitions of the occurrence of extreme rainfall are described as events characterized by deviations from great intensity from the average condition [2]. The Indian Meteorological Department says that rain with intensity over 64.5 mm per day is referred to as heavy rain
Three different methods have commonly been used to identify extreme rainfall events [4]. The first method is based on the actual rainfall amounts. A second way to define extreme precipitation events is to use specific thresholds such as the 90th and 99th percentiles of precipitation days to define heavy and very heavy events, respectively. A third way of defining extreme precipitation events is to calculate return periods of the event based on the annual maximum 24-h precipitation series [5]. This research used the first method to identify extreme rainfall events.

Several approaches were proposed to determine the temporal distribution of rainfall. In 1975, Pilgrim and Cordery proposed temporal rainfall distribution by selection of intensity rainfalls with different time bases in Sidney of Australia [6]. Some researchers have proposed theoretical methods using statistical approach, that are [6], [4], and [7]. Several previous studies on the prediction of the timing of recurrence on some events have been done by [8], [9], [10], and [11]. Although various distribution models have been used in modeling the timing of recurrence of an event, it has not found the most appropriate distribution to state the time of occurrence of events at a site. This paper aims to determine the inter event time distribution of extreme rain events and minimum waiting time until the occurrence of next extreme event through a renewal point process approach. In this study, the process point is approximated by a renewal process with inter events time are independent. For that matter, this paper will examine the forecast of the time of occurrence of extreme rainfall for two extreme rain events follows Gamma, Log Normal, and Pareto distributions.

2. Point Process

One dimensional point process can be expressed as a model that useful for random timing sequences when an event occurs. A process of spatial point is a useful model for a random pattern of dots in the d dimensional space, where d ≥ 2. As an example, if a map of the location of extreme rain events at any given time, then this map is a random pattern of dots in two dimensions.

There are several differences between the one-dimensional point process and the dimension point process higher. In point process one-dimensional, it has found natural properties which is not found on the point process of higher dimension. One-dimensional point process can expressed mathematically in a variety of different ways, including: arrival times \( t_i < t_{i+1} < \cdots \) with \( t_i \) is the occurrence time of event \( i \). Another way that can be used to define a point process is through time between events \( \tau_i = t_{i+1} - t_i \). The length of time interval between subsequent events are known as inter event times.

One-dimensional point process can expressed mathematically in a variety of different ways, including: arrival times \( t_i < t_{i+1} < \cdots \) with \( t_i \) is the occurrence time of event \( i \). Another way that can be used to define a point process is through time between two events \( \tau_i = t_{i+1} - t_i \). The lengths of the time intervals between subsequent events are known as inter event times. For some special models (Poisson and renewal process), \( \tau_1, \tau_2, \cdots \) are mutually exclusive.

2.1. Renewal Process

An important type of temporal point process is renewal process. Let \( t_1, t_2, \cdots, t_n \) is the series of occurrences events. Consider \( \{\tau_i\} \) in interval \( (0, \infty) \) where \( \tau_i = t_{i+1} - t_i \) is time difference between two consecutive events and \( \tau_i \) as a stochastic process that has a positive value. Let \( \{\tau_i\} \) is Independent and identically distributed random variables, then \( \tau_i \) is a renewal process. If its marginal is exponentially distributed, then \( \{\tau_i\} \) is Poisson stationer. A renewal process in line is simple point process so that inter event time \( \{\tau_1, \tau_2, \cdots, \tau_n\} \) are independent random variables. Temporal renewal process with probability density function \( f \), conditional intensity is defined by \( \lambda(t) = v(t - \bar{t}) \) where \( \bar{t} \) is the last time before time \( t \), \( S(t) \) is survival function that associated
Table 1. Renewal Point Process Model.

| No | Model       | pdf                                  | cdf                                  |
|----|-------------|--------------------------------------|--------------------------------------|
| 1  | Eksponential| $f(\tau) = \theta e^{-\theta \tau}$ | $F(\tau) = 1 - e^{-\theta \tau}$   |
| 2  | Weibull     | $f(\tau) = (\alpha \beta)^{\beta-1} e^{-\alpha \tau \beta}$, $\alpha > 0$, $\beta > 0$ | $F(\tau) = 1 - e^{-\alpha \tau \beta}$ |
| 3  | Gamma       | $f(\tau) = \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} e^{-\beta \tau}$, $\tau > 0$ | $F(\tau) = \Gamma(\tau, c\tau) / \Gamma(r)$, $\Gamma(\alpha, x) = \int_0^x e^{-u,u} \cdot x^{k-1} du$ |
| 4  | Log normal  | $f(\tau) = \frac{1}{\sigma \tau \sqrt{2\pi}} \exp\left\{ \frac{-(\ln \tau - Y)^2}{2\sigma^2} \right\}$, $m > 0$ | $F(\tau) = \Phi\left\{ \frac{\ln \tau - m}{\sigma} \right\}$; $\Phi = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\tau} e^{-u^2} du$ |
| 5  | Pareto      | $f(\tau) = \frac{\alpha x_m^\alpha}{\tau^{m+1}}$, $\alpha > 0$, $x_m > 0$ | $F(\tau) = 1 - \left( \frac{x_m}{\tau} \right)^\alpha$ |
| 6  | Rayleigh    | $f(\tau) = \frac{\sigma}{\alpha} e^{-\tau^2 / 2\sigma^2}$, $\tau > 0$, $\alpha > 0$ | $F(\tau) = 1 - e^{-\tau^2 / 2\sigma^2}$ |

by survival function $f$. Survival function is defined by $S(t) = \frac{f(t)}{1 - F(t)}$, where $F(t) = \int_0^t f(u) du$ is a cumulative distribution function of $f$.

The renewal model is often used for estimate the long-term time dependent probability. In the renewal model, it is assumed that the time between successive extreme rain events follows a certain probability distribution. Some distribution opportunities have used as an renewal model in reliability issues and failure time are Pareto, Weibull, exponential, Gamma, log normal, and Rayleigh [8].

If $f(\tau)$ denotes the probability density function (pdf), $F(\tau)$ as the cumulative distribution function (cdf), and $\tau$ as the inter event time, then renewal models for the extreme rainfall occurrence are described in Table 1.

Notes:
- $\tau$: inter event time extreme rainfall.
- pdf: probability density function of renewal models.
- cdf: cumulative distribution function.

3. Results and Discussion
3.1. Time Recurrence of Extreme Rainfall
Time recurrence of extreme rainfall defined as the time interval until the next rain extreme event [7]. In this section we will model the timing of the occurrence of extreme rainfall with the inter event time follows Log Normal, Pareto, and Gamma distribution.

3.1.1. Log Normal Model  Let $T$ as random variable of the inter event time of extreme rainfall is log normal distributed with parameters $\sigma^2$ and $Y$. The probability density function (pdf) is defined by [12]:

$$ f(\tau) = \frac{1}{\tau \sigma \sqrt{2\pi}} e^{-\frac{(\ln \tau - Y)^2}{2\sigma^2}}. \tag{1} $$
Cumulative distribution of log normal random variable is:

\[ F(\tau) = \int_0^\tau \frac{1}{u\sqrt{2\pi}} e^{-\left(\ln u - \bar{Y}\right)^2/2\sigma^2} du. \quad (2) \]

Cumulative distribution of equation (2) have to be evaluated by numerical method. Conditional probability density function of extreme rainfall as follows:

\[ f(\tau|T \geq t_0) = \frac{f(\tau)}{1 - F(t_0)} \quad (3) \]

so that

\[ f(\tau|T \geq t_0) = \frac{\frac{1}{\tau\sigma\sqrt{2\pi}} e^{-\left(\ln \tau - \bar{Y}\right)^2/2\sigma^2}}{1 - \int_0^{\tau} \frac{1}{u\sqrt{2\pi}} e^{-\left(\ln u - \bar{Y}\right)^2/2\sigma^2} du}. \quad (4) \]

Cumulative distribution of \( F(\tau) \) for log normal distribution is constant \( D \). Thus, conditional probability of log normal distribution for (4) is defined as:

\[ f(\tau|T \geq t_0) = Q \frac{1}{\tau\sigma\sqrt{2\pi}} e^{-\left(\ln \tau - \bar{Y}\right)^2/2\sigma^2} \quad (5) \]

where \( Q = \frac{1}{\frac{\sigma}{\tau\sqrt{2\pi}}}. \)

To estimate recurrence time for \( \hat{\tau} \) until the occurrence of the next extreme rainfall, then the maximum of conditional probability density function of log normal \( f(\tau|T \geq t_0) \) will be determined. Next, we have solution as follows:

\[ \frac{d}{d\tau} f(\tau|T \geq t_0) = Q \frac{d}{d\tau} \left\{ \frac{1}{\tau\sigma\sqrt{2\pi}} e^{-\left(\ln \tau - \bar{Y}\right)^2/2\sigma^2} \right\} = 0. \quad (6) \]

Based on equation (6) we have

\[ [(\ln \tau - \bar{Y}) + \sigma^2] e^{-\left(\ln \tau - \bar{Y}\right)^2/2\sigma^2} = 0. \quad (7) \]

and

\[ (\ln \tau - \bar{Y}) + \sigma^2 = 0. \quad (8) \]

Thus, the solution is:

\[ \ln \tau = \bar{Y} - \sigma^2. \quad (9) \]

or,

\[ \hat{\tau} = e^{\bar{Y} - \sigma^2}. \quad (10) \]

Equation (10) can be used to estimate waiting time \( \hat{\tau} \) until the next extreme rainfall. By equation (10), it can be knowing that the forecasting waiting time \( \hat{\tau} \) until the next extreme rainfall does not depend on elapsed time since the last extreme rainfall. To estimate waiting time \( \hat{\tau} \) until the next extreme rainfall using log normal model, then model parameters should be estimated. Furthermore, let mean and variance of the log normal model is given by

\[ \mu = e^{\bar{Y} + \frac{1}{2}\sigma^2} \quad \text{dan} \quad \nu = e^{2\bar{Y} + 3\sigma^2} \left( e^{\sigma^2} - 1 \right). \quad (11) \]

Based on equation (11) we have

\[ \mu^2 = \left\{ e^{\bar{Y} + \frac{1}{2}\sigma^2} \right\}^2 = e^{2\bar{Y} + \sigma^2} \quad (12) \]

and

\[ \sigma^2 = \ln \left\{ \frac{\nu}{\mu^2} + 1 \right\}. \quad (13) \]

so

\[ \bar{Y} = \ln \mu - \frac{1}{2} \ln \left\{ \frac{\nu}{\mu^2} + 1 \right\}. \quad (14) \]

Thus, we have:

\[ \bar{Y} = \ln \mu - \frac{1}{2} \ln \left\{ \frac{\nu}{\mu^2} + 1 \right\}. \quad (15) \]
3.1.2. Pareto Model  Let $T$ is random variable for inter event time of extreme rainfall that modeled by Pareto distribution with pdf as\cite{12}:

$$f(\tau) = \frac{\alpha x_m^\alpha}{\tau^{\alpha+1}}, \quad \alpha > 0, \quad x_m > 0 \quad (16)$$

Mean and variance sample are $E[T] = \mu_T = \alpha x_m$, for $\alpha < 1$ and variance $Var[T] = \frac{\alpha x_m^2}{(\alpha-1)(\alpha-2)}$ respectively. Cumulative distribution for Pareto random variable is given by:

$$F(\tau) = 1 - \left(\frac{x_m}{x}\right)^\alpha \quad (17)$$

Based on equation (4), we have conditional pdf as:

$$f(\tau|T \geq t_0) = \frac{\alpha x_m^\alpha \tau^{-(\alpha+1)}}{1 - \int_0^\tau \alpha x_m^\alpha \tau^{-(\alpha+1)} d\tau}. \quad (18)$$

Since cumulative distribution function for random variable of Pareto distribution in equation (17) is constant $C$. Thus, pdf of extreme rainfall occurrence can be written as

$$f(\tau|T \geq t_0) = W\alpha x_m^\alpha \tau^{-(\alpha+1)}, \quad (19)$$

with $W = \frac{1}{1-C}$. Furthermore, we determined waiting time for $\tau$ until the occurrence of the next extreme rainfall by maximizing the conditional pdf for the extreme rainfall occurrence as follows

$$\frac{\partial f(\tau|T \geq t_0)}{\partial \tau} = -\alpha x_m^\alpha (\alpha + 1) \tau^{-\alpha-2}. \quad (20)$$

so,

$$-\alpha x_m^\alpha (\alpha + 1) \tau^{-\alpha-2} = 0. \quad (21)$$

Since $\alpha > 0$ and $x_m > 0$, then $\tau^{-\alpha-2} = 0$, so we have $\tau = 0$. Thus it can be said that the Pareto model can not be used to forecast time recurrence of extreme rainfall.

3.1.3. Gamma Model  Gamma is a distribution that plays an important role in statistics. Let random variable $T$ is the inter event time of extreme rainfall, the pdf of Gamma distribution is \cite{12}:

$$f(\tau) = \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} e^{-\beta \tau}, \quad \tau > 0. \quad (22)$$

This distribution depends on shape parameter $\alpha$ and scale parameter $\beta$, where $\alpha > 0$ and $\beta > 0$. The mean and variance for this distribution as follows

$$E[T] = \frac{\alpha}{\beta} \quad \text{dan} \quad Var[T] = \frac{\alpha}{\beta^2}. \quad (23)$$

Cumulative distribution for random variable of Gamma is

$$F(\tau) = \int_0^\tau \frac{\beta^\alpha}{\Gamma(\alpha)} u^{\alpha-1} e^{-\beta u} du. \quad (24)$$

Substitute equation (22) and equation (24) to equation (3), we found conditional probability density function of Gamma for extreme rainfall occurrence is

$$f(\tau|T \geq t_0) = \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} e^{-\beta \tau} \frac{1}{1 - \int_0^\tau \frac{\beta^\alpha}{\Gamma(\alpha)} u^{\alpha-1} e^{-\beta u} du}. \quad (25)$$

Cumulative distribution function for random variable which Gamma distributed in equation (25) is constant $C$, then pdf of extreme rainfall occurrence can be written by:
Table 2. Renewal Model for Extreme Rainfall Occurrence.

| No | Model    | Mean and Variance | Parameters       |
|----|----------|-------------------|------------------|
| 1  | Log Normal | 62.85 and 8557.608 | $\sigma^2 = 1.154$ and $\bar{Y} = 3.564$ |
| 2  | Gamma    | 62.85 and 8557.608 | $\alpha = 0.459$ and $\beta = 0.007$ |
| 5  | Pareto   | 62.85 and 8557.608 | $\alpha > 0$ and $x_m > 0$ |

$$f(\tau|T \geq t_0) = W \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} e^{-\beta \tau},$$  \hspace{1cm} (26)

with $W = \frac{1}{1-C}$. Furthermore, we determined recurrence time for $\hat{\tau}$ until the occurrence of the next extreme rainfall by maximizing the conditional pdf for the extreme rainfall occurrence as follows

$$\frac{\partial f(\tau)}{\partial \tau} = \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} \beta^\alpha \left\{ \left( \frac{\alpha-1}{\tau} - \beta \right) \right\} = 0.$$  \hspace{1cm} (27)

Using equation (27) we found that $f(\tau|T \geq t_0)$ value reaches maximum at

$$\hat{\tau} = \frac{\alpha - 1}{\beta}.$$  \hspace{1cm} (28)

If $\alpha < 1$, then maximum value occurs outside the definition range and for this case the conditional probability density for occurrence of extreme rainfall is a decreasing monotonous function to zero for $t_0$.

3.2. Case Study

The data used in this case study is the historical extreme daily rainfall data onto the South Sulawesi region period 1 January 2000 to 17 February 2017. We have 124 extreme rainfall data. There are 70% for training data and 30% for validation data. From the sample used it is obtained that the mean sample are 62.85 and the sample variance is 8557.608. Summary of parameter estimation of the three models is described in the following Table 2:

Based on the parameter values of each model given in the Table 2, we know that using log Normal model and equation $\hat{\tau} = e^{\bar{Y} - \sigma^2}$, we have the approximate of waiting time $\hat{\tau}$ until the next extreme rainfall is $\hat{\tau} = 11.14$ days. It is therefore possible to estimate the time of occurrence of the next extreme rainfall as follows: It is known that the time of the last extreme rainfall occurrence is February 17, 2017. Since $\hat{\tau} = 11.14$, then the next extreme rainfall prediction occurs before March 1, 2017.

For Gamma model and equation $\hat{\tau} = \frac{\alpha - 1}{\beta}$, it is found that $\alpha = 0.459$. Because of $\alpha < 1$ then the maximum occurs outside the definition range. Especially for this case, conditional probability density of extreme rainfall is a monotone decreasing function to zero for $t_0 \rightarrow \infty$.

In Pareto model, $-\alpha x^2_m (\alpha + 1) \tau^{-\alpha - 2} = 0$. Since $\alpha > 0$ and $x_m > 0$, then the equation is true if $\tau^{-\alpha - 2} = 0$. So that $\frac{x_m}{\tau^2} = 0$, or $\tau^{-\alpha} = 0$. Thus, $\hat{\tau} \rightarrow \infty$. In other words, the inter events time that Pareto distributed can not be used to forecast the recurrence time of extreme rain.

3.3. Acknowledgments

The authors are grateful to all those who had provided suggestions for the completion of this paper. The authors would like to thank the ministry of technology research and higher education Indonesia for financial funding PUPT research grant 005/SP2H/LT/DRPM/IV/2017 through decentralization research.

4. Conclusion

1. The minimum waiting time until the next extreme rainfall is not depend on elapsed time since the last extreme rainfall.
2. The Log Normal model is better than Gamma model for predicting the next extreme rainfall in South Sulawesi. The minimum waiting time until next extreme rainfall is approximately 11 days.

3. Gamma model can not be used for predicting the next extreme rainfall in South Sulawesi.

References
[1] Daley D J and Vere-Jones D 2003 An Introduction to the Theory of Point Processes Springer Vol(I) ed 2
[2] Meehl G A, Karl T, Easterling D R, Changnon, et.al 2000 An Introduction to Trends in Extreme Weather and Climate Events: Observations, Socioeconomic Impacts, Terrestrial Ecological Impacts, and Model Projections Bulletin of AMS
[3] Massachusetts Institute of Technology 2017 Climate Change: Extreme Rainfall will Vary Between Regions: Intensification of Extreme Rainfall Varies from Region, Study Shows. ScienceDaily
[4] Chu P S, Zhao X, Ruan Y and Grubbs M 2009 Extreme Rainfall Events in the Hawaiian Islands J. App. Meteorology and Climatology 48 3
[5] 9. Kunkel K E, Andsager K and Easterling D R 1999 Long-term trends in extreme precipitation events over the conterminous United States and Canada J. Climate 12 2515-2527.
[6] Pilgrim D H and Cordery I 1975 Rainfall Temporal Patterns for Design Floods Journal of Hydraulics, 101 1 81-95
[7] Sunusi N, Herdiani E T and Nirwan 2017 Modeling of Extreme Rainfall Recurrence Time by Using Point Process Models J. Environ. Sci. Technol., 10 6 320-324
[8] Ellsworth W L 1955 Characteristic Earthquake Forecasts: Implications of Central California Seismicity: in Urban Disaster Mitigation: the Role of Science and Technology Elsevier Science Ltd
[9] Ogata Y 1998 Estimating the Hazard of Rupture Using Uncertain Occurrence Times of Paleoeartquakes Journal of Geophysical Research, 104 17 995
[10] Sunusi N, Darwis S, Triyoso W and Mangku I W 2010 Study of Earthquake Forecast through Hazard Rate Analysis International Journal of Applied Mathematics and Statistics 17 J10 96-103
[11] Sunusi N, Herdiani E T, Giarno and Nawawi F 2015 Exploratory Analysis of Rainfall Occurrence in South Sulawesi Region Using Spatial Point Process Journal of Mathematics and Statistics, 11 4 113-118
[12] Jugman S A, Panjer H H and Wilmot G E 2002 Loss Models: From Data to Decisions Wiley