



Research article

The role of parameters in Bayesian Online Changepoint Detection: detecting early warning of mount Merapi eruptions

Seli Siti Sholihat^a, Sapto Wahyu Indratno^{a,b,*}, Utriweni Mukhaiyar^a^a Statistics Research Group, Faculty of Mathematics and Natural Sciences, Institut Teknologi Bandung, Bandung, 40132, Indonesia^b U-CoE AI-VLB (Artificial Intelligence for Vision, Natural Language Processing and Big Data Analytics), Indonesia

ARTICLE INFO

Keywords:

Online early warning system
 Volcanic eruption
 Changepoint detection
 Mount Merapi eruption

ABSTRACT

Indonesia is a country that is surrounded by active volcanoes, which may erupt at any time; therefore, an online early warning system of volcanic eruption is crucial. In this paper, an online early warning system is constructed based on the changepoints detection on earthquake magnitude time series. This online early warning system is built using a Bayesian Online Changepoint Detection (BOCPD) method. One of the method's advantages is that one can customize the parameters (initial hyper-parameters and hazard-rate parameter) of BOCPD to follow a chosen constraint. These parameters determine the time and number of changepoints. An algorithm, called Appropriate Parameters of Bayesian Online Changepoint Detection for Early Warning (APBOCPD-EW), is proposed to get the parameters that lead the detection to the early warning points before eruption. We apply the algorithm for online early warning of mount Merapi eruptions. The results show that the proposed method produces parameters that give good estimation time for early warnings of mount Merapi's eruptions.

1. Introduction

Early warning detection of volcanic eruptions is an important study in volcanology to avoid the risks and damage due to explosions. Monitoring and studying volcanoes can develop information, actions, and evacuation plans for people living around when an upcoming eruption occurs. Volcano-monitoring studies are at the core of an effective early warning system [1]. If the early warning system fails to detect an abrupt event, it could result in severe damage [2]. In statistical analysis, the abrupt event is known as a changepoint.

Some early warnings of volcanic eruptions have been studied. Bertagnini et al. (2006) published an early warning of Vesuvius eruption by learning its precursors (gas emission, ground deformation, and seismic activity) period: the first period, lasting from months before the eruption; a second period, represented by the 15–20 days before the eruption [3]. The precursory geophysical patterns are also studied by Mothes et al. (2017) for early warning at Cotopaxi volcano [4]. Gambino et al. (2014) reviewed that the time–amplitude tilt ranges before the eruptions are categorized in large variations (up to over 100 micro radians) and minor variations (not exceeding 2.5 micro radians) [5]. The studies of mount Etna's infrasound for detecting imminent eruption are discussed by Hall (2018) and Ripepe et al. (2018) that the volcano

often produces infrasound waves before it erupts [6] where the infrasound parameter (IP) thresholds are $IP > 0.2$ for 20 minutes (a warning stage), and $IP < 0.02$ for 24 hours (a normal stage) [7]. Matoza et al. (2019) and Marcheti et al. (2019) analyzed the infrasound of volcano at regional and global scales [8] [9]. Some studies proposed the early warning using the data from satellite: total ozone mapping [10], the airborne volcanic ash cloud [11], volcano thermal [12], and the volume of volcanic sulfur dioxide emissions [13].

Scientists also study that increased seismic activity may predict an eruption [14]. According to the United States geological survey, magma and volcanic gas movement caused an earthquake. The study in estimating intruded magma volume from the cumulative seismic moment by tectonic earthquakes is another tool for predicting eruption [15]. Although earthquakes in volcano areas can occur at different places (longitude and latitude) and depth, these studies conclude that magma activity causes the earthquakes before an eruption/explosion [16]. Further, the earthquakes at the volcano become the signal of the eruption [17] and a sequence of earthquakes may indicate an imminent eruption [18], [19]. Therefore, a changepoints analysis in the earthquake magnitude data for eruption's early warning needs further research.

In this paper, early warnings are constructed based on changepoints in the time series of earthquake magnitude. The changepoint

* Corresponding author at: Statistics Research Group, Faculty of Mathematics and Natural Sciences, Institut Teknologi Bandung, Bandung, 40132, Indonesia.
 E-mail address: sapto@math.itb.ac.id (S.W. Indratno).

<https://doi.org/10.1016/j.heliyon.2021.e07482>

Received 6 January 2021; Received in revised form 27 February 2021; Accepted 1 July 2021

defines an eruption signal, where the time of changepoint represents the early-warning time. Changepoint detection methods have been successfully applied in some cases, such as the satellite fault prediction using Bayesian Changepoint Detection [20] and change detection of the UAV (Unmanned Aerial Vehicle) fuel system [21]. Changepoint detection using BOCPD successfully detects a change in the stratification of the earth's crust, the Dow Jones return rate, and coal mine disaster data [22].

The changepoint for early warning in our research uses Bayesian Online Changepoint Detection (BOCPD) method. BOCPD can detect the changes incrementally as data arrives. This model's advantage is an efficient computational cost by an online scheme. Another benefit of BOCPD is that we can set its parameters (initial hyperparameters and hazard-rate parameter in BOCPD) to produce the expected time and number of the changepoints. We utilize this to adjust the time of changepoint before the time of the eruption.

The BOCPD parameters are initial hyper-parameters and hazard-rate parameter. The hazard-rate parameter is a parameter to determine a conditional prior on changepoint. Here, we use a constant parameter derived from a Geometric distribution. The Geometric distribution is a common model for calculating the probability of step number failures before the first success. It represents the probability of the run-length until the success to detect changepoint. Meanwhile, initial hyper-parameters are used to define the initial prediction-probability density function that initiates the conjugate-exponential models in BOCPD. These initial hyper-parameters are updated incrementally as new data arrives. For detecting changepoints before the eruption time, we should adjust these parameters appropriately.

The role of BOCPD parameters has been studied. We design an algorithm for selecting the parameters for the early warning system of a volcano eruption. The algorithm is called Appropriate Parameter Bayesian Online Changepoint Detection for Early Warning (APBOCPD-EW). The APBOCPD-EW algorithm outputs are appropriate initial hyper-parameters and a hazard rate parameter for accurately detecting early warning points. We define a minimum number of changepoints (early warnings) before an eruption and a minimum total lag day from explosion time to get the proper detection of early warning time.

According to the earthquake's magnitude data, we should use its distribution that qualifies a Gutenberg-Richter relation. A truncated generalized exponential distribution is one of the magnitude distributions that follow a Gutenberg-Richter relation [23]. BOCPD with conjugate exponential models supports the earthquake's magnitude data type (continuous and bounded). The conjugate exponential models that use bounded continuous data type is a truncated generalized Student-t distribution. Thus, the initial hyper-parameters refer to parameters in a probability density function of truncated generalized Student-t.

The algorithm is designed for general data types with various eruption numbers. So, this study can also use other volcanic data types or even apply the method to other volcanoes. For using another data type, we need to confirm that BOCPD supports its distribution. The data type for early warnings of an eruption is selected based on the signs that usually initiate an eruption on that volcano.

We choose mount Merapi in research for good reasons. The volcano is located in a tectonically active region of south-central Java, and mount Merapi is the most active volcano in Indonesia that frequently erupts [24]. During 2012-2018, mount Merapi has erupted four times within earthquake sequences. Therefore, we use earthquake magnitude data in the mount Merapi area to obtain information for its early-warning system. Some researches of mount Merapi eruption have been done in last five decades: the eruption and earthquake events at Merapi volcano during 2006 [25], the 2010 explosive eruption of Java's Merapi volcano in a 100-year event [26], correlation of seismic activity and fumarole temperature at mount Merapi in 2000 [27], classification of seismic signals of volcanic origin at mount Merapi [28] and monitoring a temporal change of seismic velocity before the eruption of mount Merapi in 1992 [29]. The seismic activity changes at mount Merapi can

be treated as probabilistic changes, which the changepoint detection methods can solve.

The paper consists of five sections. Section 2 discusses the theory of BOCPD mathematically and the analysis of parameters of BOCPD. Section 3 describes the methodology of APBOCPD-EW as the proposed algorithm and shows the BOCPD algorithm. Section 4 discusses the data, analysis sensitivity of hyper-parameter on BOCPD as part of our research, and analysis of APBOCPD-EW results of the APBOCPD-EW in detecting early warning. We use training data to get early warnings of three eruptions, and then we use testing/validation data to detect early warnings of the fourth eruption. It confirms that using appropriate parameters gives the proper estimation of the early warning. The conclusions of the paper are provided in Section 5.

2. Theory

This section discusses the BOCPD method mathematically and the role of updating parameters in truncated generalized Student-t.

2.1. Bayesian Online Changepoint Detection Method

Bayesian Online Changepoint Detection (BOCPD) is an online method to detect the changepoints on a time-discrete data sequence. BOCPD uses Bayesian statistical analysis of run-length posterior distribution [22].

Assume we have data of earthquake magnitude y_t . Generally, run-length at time t (r_t) is a non-negative discrete variable denoting the number of time steps elapsed at time $t = 1, 2, \dots, N$ from the last changepoint. A changepoint is the data at early warning time. Thus, the run-length at time t defines the number of time steps elapsed at time t since the last early warning time.

Initially at time $t = 1$, y_1 is a changepoint then the step number to the changepoint at time $t = 1$ is zero ($r_1 = 0$). Prior probability of a changepoint given previous run-length is written as $P(r_t = 0 | r_{t-1})$. If the next observation is not a changepoint, the run-length is increased by one. Thus, $P(r_t = r_{t-1} + 1 | r_{t-1})$ is the probability of the growth run-length given the previous run-length. Posterior of run-length given data sequence $y_{1:t} = y_1, y_2, \dots, y_t$ is written as follows

$$P(r_t | y_{1:t}) = \frac{P(y_{1:t}, r_t)}{P(y_{1:t})}, \quad (1)$$

where $P(y_{1:t}, r_t)$ is a joint probability function of the observation $y_{1:t}$ and the run-length (r_t) that is formulated by marginalizing the run-length r_{t-1} as follows

$$P(y_{1:t}, r_t) = \sum_{r_{t-1}} P(r_t | r_{t-1}) P(y_t | r_t, y_{t-1}^{(r)}) P(y_{1:(t-1)}, r_{t-1}), \quad (2)$$

where $y_{t-1}^{(r)} = y_{(t-1-r):t-1}$ are the data sequence from the last changepoint. The probability $P(r_t | r_{t-1})$ is formulated in the following way

$$P(r_t | r_{t-1}) = \begin{cases} H(r_t), & r_t = 0 \text{ (changepoint)} \\ 1 - H(r_t), & r_t = r_{t-1} + 1 \text{ (not changepoint),} \\ 0, & \text{otherwise,} \end{cases} \quad (3)$$

where $H(r_t)$ is a hazard function. We use a geometric distribution with time scale λ for the inter-arrival time of changepoint which gives $H(r_t) = \frac{1}{\lambda}$. The joint probability function of data observation $P(y_{1:t})$ is derived by marginalizing run-length r_t , as follows

$$P(y_{1:t}) = \sum_{(r_t)} P(y_{1:t}, r_t) \quad (4)$$

Substituting Equation (3) into Equation (2); and then substituting Equation (2) and Equation (4) into Equation (1), one obtains

$$P(r_t = 0 | y_{1:t}) = \frac{\sum_{r_{t-1}} H(r_t) P(y_t | r_{t-1}, y_{t-1}^{(r)}) P(y_{1:(t-1)}, r_{t-1})}{\sum_{(r_t)} P(y_{1:t}, r_t)}, \quad (5)$$

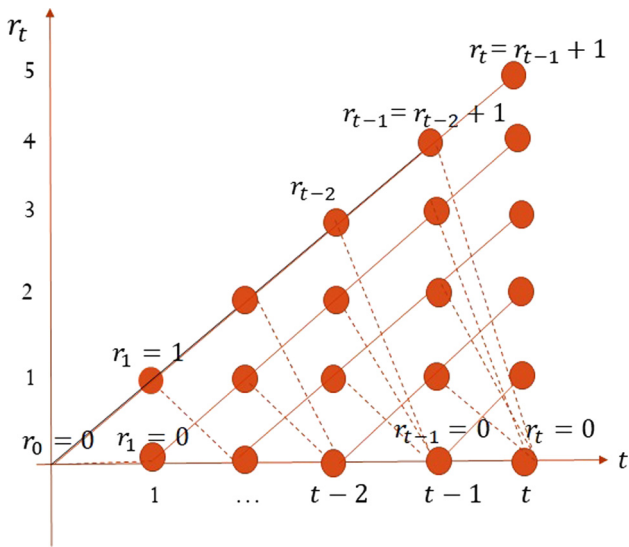


Fig. 1. The figures illustrate how the run-length at time t has two possibility of run-length; the solid lines show how the run-length growth $r_t = r_{t-1} + 1$ and the dotted lines show the run-length time t that drop to zero. Zero run-length $r_t = 0$ is defined when changepoint occurs.

where $P(y_t | r_{t-1}, y_{t-1}^{(r)})$ is an Underlying Predictive Model (UPM) function of y_t given $y_{t-1}^{(r)}$ and r_{t-1} ; $t = 1, 2, \dots, N$; $r = 0, 1, 2, \dots, r_{t-1}$. The changepoint is detected when posterior of zero run-length $P(r_t = 0 | y_{1:t})$ is higher than the posterior of growth run-length $P(r_t = r_{t-1} + 1 | y_{1:t})$. In this case, y_t is a changepoint and time t is occurrence time of the changepoint.

An illustration of the run-length at time t in BOCPD, r_t , is depicted in Fig. 1. The run-length becomes zero when the changepoint is detected. The run-length is increased by one when the changepoint is not detected.

The online scheme is built by updating the hyper-parameter of the predictive distribution. We choose a truncated generalized Student-t distribution for UPM.

2.2. Role of BOCPD hyper-parameters

The probability density function of a generalized Student- t distribution is defined in Equation (6)

$$P(y_t | \mu_t, \sigma_t, \nu_t) = \frac{\Gamma\left(\frac{\nu_t+1}{2}\right)}{\sqrt{\pi}\sigma_t\Gamma\left(\frac{\nu_t}{2}\right)} \left(1 + \left(\frac{y_t - \mu_t}{\sigma_t}\right)^2\right)^{-\frac{\nu_t+1}{2}}, \tag{6}$$

where mean (μ_t), scale (σ_t) and degrees of freedom (ν_t) are its parameters. Every time a new data point arrives, the generalized Student- t distribution is updated by updating μ_t , σ_t , and ν_t . Those parameters depend on parameters α_t and β_t , called hyper-parameters.

Meanwhile, a truncated generalized Student- t with lower bound a and upper bound b , is defined in Equation (7); we choose fixed $F(a)$ and $F(b)$.

$$P_T(y_t | \mu_t, \sigma_t, \nu_t) = \frac{P(y_t | \mu_t, \sigma_t, \nu_t)}{F(b) - F(a)}; F(x) = \int_{-\infty}^x P(y_t | \mu_*, \sigma_*, \nu_*) \tag{7}$$

The formulas for updating the parameters and hyper-parameters are given in Equation (8), (9), and (10):

$$\mu_t^{r+1} = \frac{\mu_{t-1}^r r + y_t}{r + 1}, \tag{8}$$

$$\sigma_t^{(r+1)} = \sqrt{\frac{2\beta_t^{r+1}(r+1)}{r}}; \beta_t^{r+1} = \beta_{t-1}^r + \frac{r(y_t - \mu_{t-1})^2}{2(r+1)}, \tag{9}$$

$$\nu_t^{r+1} = 2\alpha_t^{r+1}; \alpha_t^{r+1} = \alpha_{t-1}^r + 0.5 \tag{10}$$

Those updating parameters are based on the conjugate Bayesian analysis of Gaussian distribution [30]. The hyper-parameter α_t^{r+1} is used to define the new degrees of freedom ν_t^{r+1} in Equation (10). Meanwhile, the hyper-parameters β_t^{r+1} define the scale $\sigma_t^{(r+1)}$ in Equation (9). By the increasing iteration, the degrees of freedom and variance of generalized Student- t distribution increase.

The iterative method by updating predictive distribution is sensitive to the initial hyperparameters. Let the initial hyper-parameters of predictive distribution be α_0 and β_0 , these parameters are updated until the changepoint is detected. The left Curves in Fig. 2 show the updating of generalized Student- t distribution from $t = 1$ until $t = 30$. The figure shows two different updates by two different initial hyper-parameters; $\beta_0 = 1$ (red curve) and $\beta_0 = 7$ (blue curve).

The higher iteration produces a curve with a higher peak and a thinner tail. The possibility of the changepoints increases because the data with a small probability density value increases. Therefore, the red curve will be faster than the blue curve in detecting changepoints. The initial hyper-parameter α_0 and β_0 determine the changepoints detection.

The other important parameter in BOCPD is λ , which formulates the hazard rate, $\frac{1}{\lambda}$. If λ increases, then the hazard rate and the probability of changepoint decrease. Further, decreasing the change-point probability increases the inter-arrival time of the changepoint. As the inter-arrival time of the changepoint increases, the number of detected changepoints decreases. On the contrary, if λ decreases, then the changepoints inter-arrival time decreases, and the number of changepoints increases. Clearly, that initial hyper-parameter α_0 , β_0 and parameter λ have the role on BOCPD.

In the next section, the proposed algorithm to find the appropriate parameters ($\alpha_0, \beta_0, \lambda$) in BOCPD is discussed for detecting the early warning points.

3. Methodology

In this section, two main algorithms for the early warning volcano eruption are presented. The proposed algorithm, Appropriate Parameters Bayesian Online Changepoint Detection for Early Warning algorithm (APBOCPDEW), is presented in Algorithm 1 for finding appropriate initial hyper-parameters (α_0 and β_0) and parameter hazard rate (λ). Meanwhile, the BOCPD algorithm with the updating of truncated generalized Student- t parameters is presented in Algorithm 2. The changepoints results provide the time and number of changepoints. Times of the changepoints refer to the times of early warnings.

The APBOCPD-EW algorithm runs BOCPD with input earthquake magnitude data, the number of eruptions, the date of eruptions, hazard rates (λ), and initial hyper-parameters (α_0 and β_0). The results are the appropriate hazard rate and initial hyper-parameters by screening all changepoints detection results from some observation parameters. The screening rule is to have at least one early warning point before eruption with a minimum number of changepoints and total lag-day to the eruption-time.

Algorithm 1 Appropriate Parameters Bayesian Changepoint Detection for Eruption Early Warning (APBOCPD-EW).

```

The earthquake magnitude data =  $y_{1:N}$ 
Input: The earthquake magnitude data =  $y_{1:N} = y_1, y_2, \dots, y_N$ ;  $n_e$  = eruptions number,  $d_e$  = eruption date
Observe for some  $\lambda$ , and initial  $\alpha_0$  and  $\beta_0$  in domain
changepoint( $Cp_{\lambda, \alpha_0, \beta_0}$ )  $\leftarrow$  BOCPD algorithm
if  $\exists Cp_{\lambda, \alpha_0, \beta_0}$  before eruption date ( $d_e$ )  $\forall d_e$  then
     $G_{Cp_{\lambda, \alpha_0, \beta_0}}$   $\leftarrow$   $Cp_{\lambda, \alpha_0, \beta_0} G_{Cp}$  : good change point
     $N_{G_{Cp_{\lambda, \alpha_0, \beta_0}}}$  = change points number of  $G_{Cp_{\lambda, \alpha_0, \beta_0}}$ 
     $totallag_{Cp_{\lambda, \alpha_0, \beta_0}} = \sum_{i=1}^{n_e} lag_{Cp_{\lambda, \alpha_0, \beta_0}}(i)$ ; total lag-day to eruption i-th
end if
 $S \leftarrow$   $argmin_{\lambda, \alpha_0, \beta_0} (N_{G_{Cp_{\lambda, \alpha_0, \beta_0}}})$ ;  $S = \{\lambda, \alpha_0, \beta_0 | \min(N_{G_{Cp_{\lambda, \alpha_0, \beta_0}}})\}$ 
Appropriate  $\{\lambda, \alpha_0, \beta_0\} \leftarrow$   $argmin_{\lambda, \alpha_0, \beta_0} \{totallag_{Cp_{\lambda, \alpha_0, \beta_0}}\} (\{\lambda, \alpha_0, \beta_0\} \in S)$ 

Result: Appropriate initial hyper-parameter  $\alpha_0, \beta_0$  and  $\lambda$  for early warning
    
```

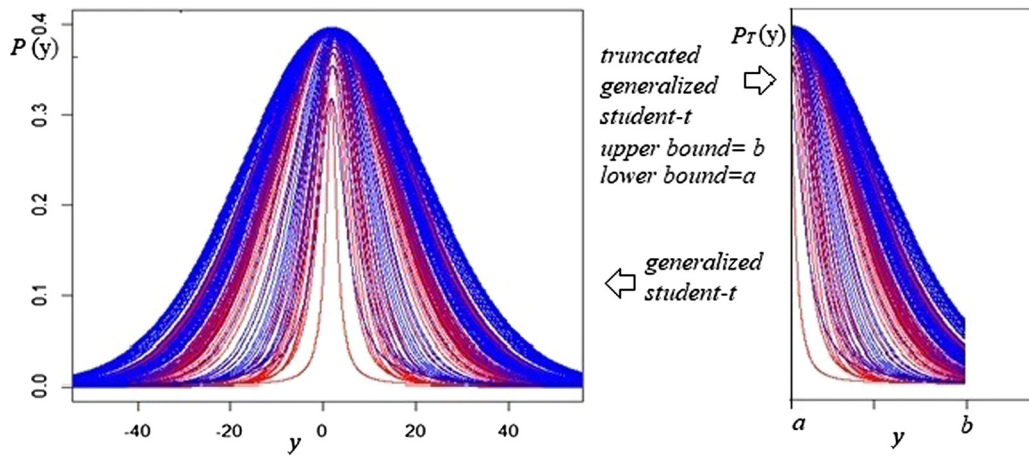


Fig. 2. The updating curve of generalized Student-t distribution and truncated generalized Student-t by 30 iterations. The left curves are probability density function ($P(y) = P(y_i|r_{t-1}, y'_{t-1})$) of generalized Student-t that are started by two condition initial value: $\beta_0 = 1; \alpha_0 = 1$ (red curves) and $\beta_0 = 7; \alpha_0 = 1$ (blue curves). The right curves are illustrations of truncated generalized Student-t with lower bound a and upper bound b . Since first iteration, the peak and variance of red curve is below the blue curve. Decreasing $P(y_i|r_{t-1}, y'_{t-1})$ will decrease growth probabilities which increase the possibility of changepoint. Thus, changepoint detection on the red curve is faster than the blue curve.

Algorithm 2 Bayesian Online Changepoint Detection.

```

Require:  $y_{1:N} = y_1, y_1, \dots, y_N; \lambda$ 
Initialization :  $\mu_0 = y_1; \sigma_0^2 = 1; r_0 = 0, H(r_0) = 1;$ 
 $\alpha_0, \beta_0; P(y_1|y^{(0)}, r_0 = 0) = P(y_1|\mu_0, \sigma_0^2);$ 
while changepoint is not detected do
  Observe new datum:  $(y_t)$ 
  Evaluate predictive probability;  $P(y_t|r_t, y^{(t)}) = P(y_t|\mu_{t-1}, \sigma_{t-1}, v_{t-1})$ 
  Calculate joint probability of growth run-length:
   $P(y_{1:t}, r_t = r_{t-1} + 1) = (1 - H(r_t))P(y_t|r_t, y^{(t)})P(y_{1:(t-1)}, r_{t-1})$ 
  Calculate joint probability for changepoint:
   $P(y_{1:t}, r_t = 0) = \sum_{r_{t-1}} (H(r_t))P(y_t|r_t, y^{(t)})P(y_{1:(t-1)}, r_{t-1})$ 
  Marginal distribution:  $P(y_{1:t})$ 
  Determine run-length distribution:
   $P(r_t = 0|y_{1:t}) = \frac{\sum_{r_{t-1}} (H(r_t))P(y_t|r_t, y^{(t)})P(y_{1:(t-1)}, r_{t-1})}{P(y_{1:t})}$ 
   $P(r_t = r_{t-1} + 1|y_{1:t}) = \frac{(1-H(r_t))P(y_t|r_t, y^{(t)})P(y_{1:(t-1)}, r_{t-1})}{P(y_{1:t})}$ 
  if  $P(r_t = 0|y_{1:t}) \geq P(r_t = r_{t-1} + 1|y_{1:t}); y_t$  is a changepoint then
     $r_t = 0$ 
  else  $\{P(r_t = 0|y_{1:t}) < P(r_t = r_{t-1} + 1|y_{1:t}); y_t\}$ 
     $r_t = r_{t-1} + 1$ 
  update hyper-parameter and parameters:  $\alpha_t, \beta_t, \mu_t, \sigma_t, v_t$ 
end if
end while
Result: Changepoint (CP)
Repeat algorithm for new data coming
    
```

We observe some $\lambda, \alpha_0, \beta_0$ in BOCPD using Algorithm 1, then collect all changepoints results. Good changepoints are chosen from the results that have at least one detection before the eruption. For every good changepoint, a number of changepoints $N_{G_{cp}(\lambda, \alpha_0, \beta_0)}$ and total lag-day ($total_{lag_{cp}(\lambda, \alpha_0, \beta_0)}$) are calculated; then choose a minimum $N_{G_{cp}(\lambda, \alpha_0, \beta_0)}$. We choose the minimum number of changepoints for minimum early warnings, which can avoid panic to public. In algorithm, a set S is a collection set of $(\lambda, \alpha_0, \beta_0)$ that have the minimum number of changepoint. A total of lag day is the total day starting from changepoints (as early warnings points) to the eruptions; summing the lag-day of each early warning points to eruptions ($lag_{cp}(\lambda, \alpha_0, \beta_0)(i)$ for $i = 1, 2, \dots, n_e$). The early warnings time should be close to the eruptions, then we choose the collection of $(\lambda, \alpha_0, \beta_0)$ from S that have the minimum total lag-day. By this technique, BOCPD with the chosen $\lambda, \alpha_0, \beta_0$ detects early warning points properly. It detects the early warning points before eruptions with the closest eruptions time. The chosen $\lambda, \alpha_0, \beta_0$ are called the appropriate BOCPD parameters for early warning.

4. Result and discussion

4.1. Data

Mount Merapi is located in the south of Central Java, Indonesia. Four eruptions have taken place at mount Merapi, from July 8, 2012, until June 23, 2018. The first eruption was on November 18, 2013; the second eruption was on March 10, 2014; the third on November 4, 2016, and the fourth eruption on May 11, 2018.

Earthquakes magnitude data of mount Merapi were collected from July 8, 2012, until June 23, 2018, which is available at the link <https://doi.org/10.6084/-/m9.figshare.11559279.v1>. The earthquake magnitude data sequences are collected inside the rectangular geographic area, with a longitude interval of 106.27 – 109.91 and latitude interval $(-7.72566) - (-6.14046)$. The magnitude range is $Mlv1.8 - Mlv5.2$ that depicted in Fig. 3 for the variant depth from 10 to 303 km.

Data is divided into training data and testing data. The training data is data from July 8, 2012, until one day before the third eruption of mount Merapi, on November 3, 2016. The testing data is data from the third eruption in November 4, 2016 until June 23, 2018. The testing data is used for detecting an early warning for the fourth eruption. It is to confirm that parameters from APBOCPD-EW work properly in detecting early warning before the eruption.

Some time-discrete data cases are managed from real-time earthquake magnitude data for comparison in results. The first case is the daily data of earthquake magnitude inside the rectangular area, 1-day data. The second case is maximum earthquake magnitude in two-day intervals, 2-day data. The data cases are continued until the last case (10-day data, the maximum earthquake magnitude in ten-day intervals).

Fig. 4 shows the magnitude data in two cases data of mount Merapi from July 8, 2012, until June 23, 2018. Data are shown for 1-day data and 10-day data. The 10-day case has fewer data points to analyze than the 1-day case.

4.2. Analysis sensitivity of BOCPD parameters

We analyze the impact of initial hyper-parameters α_0, β_0 , and hazard rate parameter λ on change-point detection. The purpose is to support our idea about choosing the appropriate hyper-parameter through the minimum number of the changepoints and minimum lag-day of the changepoints to eruptions.

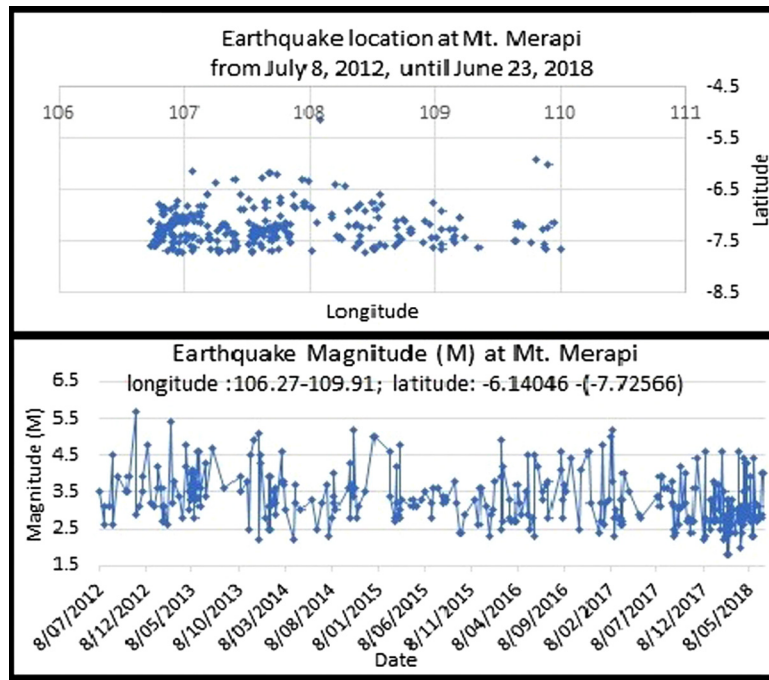


Fig. 3. Scatter plot of earthquake locations (longitude and latitude) and the scatter plot of earthquake magnitude (Mlv).

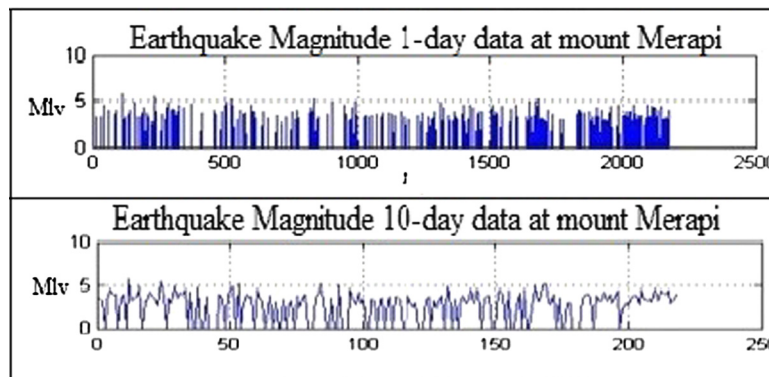


Fig. 4. 1-day data and 10-day data of earthquakes magnitude at mount Merapi from July 8, 2012, until June 23, 2018. 10-day case has less data points to analyze than 1-day case.

The observing results are summarized in Tables 1, 2, and 3. These tables list the numbers of detected changepoints with some observing parameter λ and initial hyper-parameters α_0, β_0 . The tables show that the number of changepoints decreases significantly with increasing $\lambda = 3, \dots, 10, 20$. Meanwhile, the number of change points decrease slowly with increasing initial hyper-parameter $\alpha_0 = 1, \dots, 10$ or $\beta_0 = 1, \dots, 10$. It concludes that the parameter λ affects more significantly than α_0 and β_0 . Discretized data into some data cases also affect the detection. Number of detected changepoints for different data cases with $\lambda = 3, \dots, 10, 20$ are listed in Table 3. The case 10-day data has a minimum number of changepoints than other data cases because of the fewer data number that caused changepoints. The observing results perform that different parameters and data cases detect changepoints at different numbers and times.

BOCPD with random parameters (without APBOCPD-EW) could not detect changepoints properly for early-warning points. Using the random values for initial hyper-parameters and hazard rate, BOCPD may give too many detections, e.g., the results in Fig. 5. The figure shows BOCPD results detection with parameters $\lambda = 20, \alpha_0 = 1$, and $\beta_0 = 1$. The drop of run-length to zero ($r = 0$) corresponds to the changepoint detection.

Table 1. Changepoints numbers (n_{cp}) of earthquake magnitude sequence at mount Merapi by given initial $\alpha_0 = 1, \beta_0 = 1, 2, 3, 4, 5, 6, 7, 9, 10$ and constant $\lambda = 3, 4, 5, 6, 7, 8, 9, 10$.

	β_0									
	1	2	3	4	5	6	7	8	9	10
$\lambda = 3$	68	60	59	59	59	59	58	58	57	56
$\lambda = 4$	37	35	31	29	29	27	25	25	24	22
$\lambda = 5$	25	22	19	18	17	16	14	14	13	11
$\lambda = 6$	16	15	14	13	13	10	9	9	7	6
$\lambda = 7$	14	11	11	10	8	8	6	5	4	4
$\lambda = 8$	11	10	8	7	6	5	4	4	3	3
$\lambda = 9$	9	7	6	6	4	4	4	3	2	2
$\lambda = 10$	7	6	5	5	4	3	3	2	2	2

The analysis of these simulations gives the idea of APBOCPD-EW algorithm that is described in the methodology section.

4.3. Analysis of APBOCPD-EW results

The APBOCPD-EW algorithm works finding appropriate parameters with domain setting in Table 4 and the number of the eruption $n_e = 3$

Table 2. Changepoints number (n_{cp}) of earthquake magnitude sequence at mount Merapi by 10-day data given initial $\beta_0 = 1$, $\alpha_0 = 1, 2, \dots, 10$ and constant $\lambda = 3, 4, \dots, 10$.

	α_0									
	1	2	3	4	5	6	7	8	9	10
$\lambda = 3$	68	73	73	73	73	72	72	72	72	72
$\lambda = 4$	37	38	39	42	47	51	53	56	60	61
$\lambda = 5$	25	28	26	28	28	30	31	33	38	39
$\lambda = 6$	16	17	19	20	20	23	23	28	29	31
$\lambda = 7$	14	14	14	14	16	16	17	17	18	20
$\lambda = 8$	11	11	10	11	13	14	15	15	16	16
$\lambda = 9$	9	9	9	8	10	12	11	13	14	16
$\lambda = 10$	7	7	7	8	8	8	11	11	12	13

Table 3. Number of changepoints using mount Merapi earthquake magnitude sequence; given $\alpha_0 = 1$, $\beta_0 = 5$ for some constant $\lambda = 3, 4, 5, 6, 7, 8, 9, 10, 20$ and data cases 1-day until 10-day.

	λ									
	3	4	5	6	7	8	9	10	20	
1-day data	514	224	158	130	109	91	80	76	44	
2-day data	272	141	90	67	47	39	30	28	17	
3-day data	185	102	58	43	27	23	20	17	9	
4-day data	145	77	46	28	20	17	14	13	5	
5-day data	115	63	35	23	16	13	13	11	3	
6-day data	99	51	28	17	14	12	10	8	3	
7-day data	83	44	24	15	14	11	8	9	3	
8-day data	74	39	22	15	12	7	9	5	2	
9-day data	66	30	18	13	8	6	5	4	2	
10-day data	59	29	17	13	8	6	4	4	2	

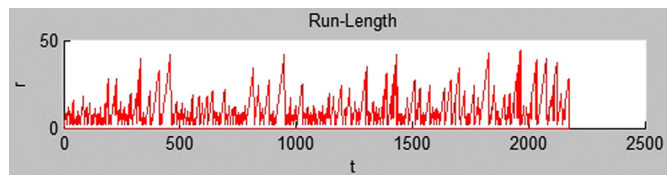


Fig. 5. The run-length results by applying BOCPD using 1-day data with $\lambda = 20$, $\alpha_0 = 1$, and $\beta_0 = 1$. These parameters could not be used to detect early warning of mount Merapi eruption because the data consists of four eruptions. Meanwhile, these parameters are resulting in 125 changepoints detection. The number of eruptions is too far from the number of changepoints. We need appropriate parameters of BOCPD to produce minimum changepoints (at least one changepoint before an eruption).

(data training includes three eruptions data). The results are summarized in Table 5.

Table 5 lists the parameters results of APBOCPD-EW for each data case using training data. Different data cases produce different detection results. Using the parameters results of APBOCPD-EW, BOCPD can detect all early warnings point for volcanic eruptions of mount Merapi. The appropriate hyperparameter λ , α_0 , and β_0 were chosen based on the minimum number of changepoints and total lag day of changepoints to eruptions. Even though the number of changepoints is minimal, they should guarantee that all outbreaks have early warning points. An ideal condition is that the number of early warnings is equal to the number of eruptions (if possible).

By APBOCPD-EW algorithm, it finds that appropriate parameters with 1-day data are $\lambda = 96$, $\alpha = 1$, and $\beta = 6$ (Table 5). This provided 15 changepoints that detected three good early warnings for three eruptions: warning on September 20, 2013, December 20, 2013, and October 26, 2016 (Table 6).

Using appropriate parameters in Table 5, good early warnings time (date) for each case are shown in Table 6. The table shows that each appropriate hyper-parameters produce early warnings before mount Merapi's eruptions that at least a changepoint is detected before the eruptions.

Table 4. Domain of hazard rate ($\frac{1}{\lambda}$) and initial hyper-parameter (α_0 and β_0) from daily until 10-day data.

	Data cases(-day)									
	1	2	3	4	5	6	7	8	9	10
λ	3-200	3-142	3-31	3-12	3-7	3-6	3-6	3-6	3-6	3-5
α_0	1	1	1	1	1	1	1	1	1	1
β_0	1-30	1-30	1-30	1-30	1-30	1-30	1-30	1-30	1-30	1-30

Table 5. The results of APBOCPD-EW: appropriate hazard rate parameter and initial hyper-parameter for detecting early warnings in training data (with three eruptions).

Data Cases	λ	α_0	β_0	N_{cp}
1-day data	96	1	6	15
2-day data	18	1	6	14
3-day data	7	1	10	13
4-day data	5	1	14	14
5-day data	5	1	10	17
6-day data	5	1	9	15
7-day data	4	1	22	12
8-day data	6	1	8	18
9-day data	6	1	1	13
10-day data	4	1	18	11

The best early warning with the minimum number of changepoints is given by a data case 10-day using $\lambda = 4$, $\alpha_0 = 1$, $\beta_0 = 18$, in Table 5. It provided 11 early warnings on three eruptions. Meanwhile, considering the minimum total lag-day to the explosion, the best early warning is given by the 5-day data using $\lambda = 5$, $\alpha_0 = 1$, $\beta_0 = 10$. It is shown in Table 6. It provided 17 early warnings on three eruptions. The number of changepoints in the 5-day case is six more than the 10-day case.

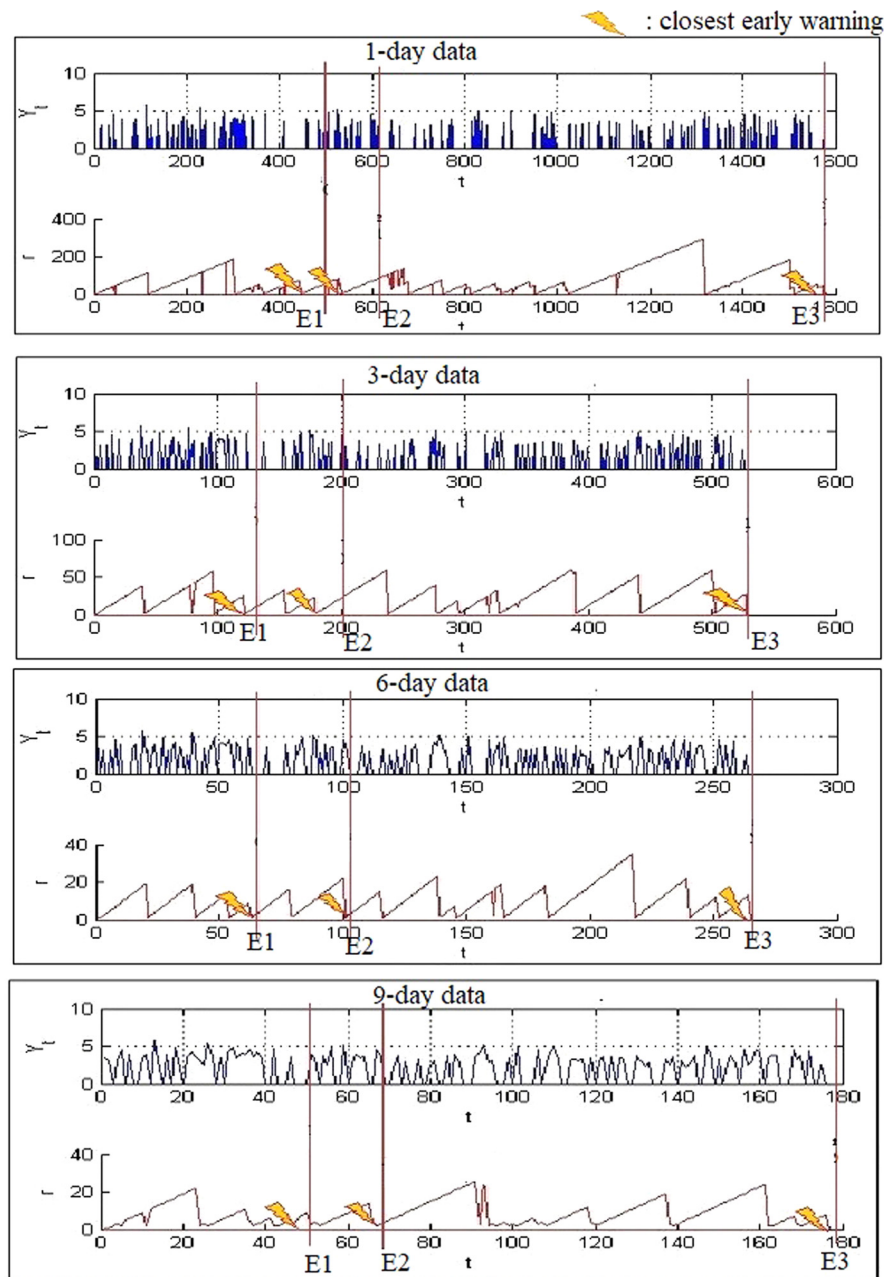
Fig. 6 shows the sequence of changepoints as the early-warning points using data cases 1-day, 3-day, 6-day, and 9-day. The detection results show that the early warnings are detected more than once, where the minimum number of early warning points is for detecting a second eruption.

4.4. Validation

This subsection discusses the validation that BOCPD with the appropriate parameters λ , α_0 , β_0 from Table 5 can detect early warnings point of mount Merapi eruption. We applied it to the testing data, earthquake magnitude data from the third eruption until June 23, 2018. The result, early warnings time are listed in Table 7. The table lists early warnings time before the fourth eruption with a number of the changepoints (N_{cp}) and the lag day $total_{lag_{cp}}$. The BOCPD with the appropriate hyperparameters detects the early warnings on all ten data cases.

Three data cases give the closest early warning predictions: Data 1-day, Data 3-day, and Data 9-day, with only have 6-12 lag days before the eruption. Detection using data cases 1-day, 3-day, 5-day, 6-day, 8-day, and 9-day give proper early warnings for the fourth eruption. Their lag-day is not too high. Meanwhile, other data cases detect the fourth eruption not well enough because the detections' date was too far from the explosions time, having 275-297 lag days.

Using the testing data, the detection is successful in detecting at least one changepoint before the outbreak because the number of changepoints for all data cases is more than one detection of early warnings ($N - cp > 1$). Some other detections with a large lag-day from the eruption are false early warnings (no outbreak occurs close to the early warning points). The detections on data cases 2-day, 4-day, 7-day, and 10-day are false early warnings (no eruption). These could be a result of failed eruptions. A failed explosion is defined as an instance in which magma reaches shallow depths but does not reach the surface [31].



E1: Eruption in 18 Nov' 2013 ; E2: Eruption in 10 March 2014; E3: Eruption in 4 Nov' 2016

Fig. 6. Closest changepoints for early warning detection based on a minimum number of changepoints and total lag-day. The drop run-length ($r_t = 0$) is the early warning points. Each eruption has more than one early warnings in each data cases 1-day data, 3-day data, 6-day data, and 9-day data. The maximum number of early warnings is for third eruption (E3). Meanwhile the second eruption (E2) has minimum number of early warning points.

5. Conclusion

BOCPD with appropriate parameters can be an alternative statistical method for early warnings of eruptions. The appropriate parameters are provided using the APBOCPD-EW algorithm. The algorithm succeeds in getting the parameters (hazard rate (λ) and initial hyper-parameter (α_0) and β_0) for BOCPD, then it detects the early warnings time properly. BOCPD with appropriate parameters detects early warning points before the explosions for all data cases. The results also show that different data cases give different results in early warnings time.

The magma at mount Merapi eruptions runs 17-40 meters per day to reach the surface on each explosion, with the magma depth is 3-4 Kilometers (3000-4000 meters). It concludes that it is reasonable that the maximum lag-day from the changepoint to the eruption is 4000 meters

divided by 17 meters, around 235 days. In all cases 1-day until 10-day data, early warning points on training data are well detected three eruptions because the total lag day is less than three times 235 days. The lag-day for each eruption is listed in Table 6. Some early warnings are very close to the eruptions time, about 10-12 days before the eruption. The early warnings of the fourth eruption using testing data are listed in Table 7. The number of lag-day before the fourth eruption that higher than 235 days are reached by 2-day, 4 day, 7 day, 8-day, and 10-day data cases. But, the 8-day case has 251 lag-day that close to 235. We conclude that there are at least six out of ten cases detecting early warning well. By testing data results, we have a valuable conclusion that the opportunity of the detection using appropriate parameters on BOCPD is 60% over ten data cases.

Table 6. Closest early warnings time of mount Merapi eruptions compare to the eruptions time using the appropriate parameters and training data.

Actual eruption	(1-day)	(2-day)	(3-day)	(4-day)	(5-day)	(6-day)	(7-day)	(8-day)	(9-day)	(10-day)
	$\lambda = 96$	$\lambda = 18$	$\lambda = 7$	$\lambda = 5$	$\lambda = 5$	$\lambda = 5$	$\lambda = 4$	$\lambda = 6$	$\lambda = 6$	$\lambda = 4$
	nCP = 15	nCP = 14	nCP = 13	nCP = 14	nCP = 17	nCP = 15	nCP = 12	nCP = 8	nCP = 13	nCP = 11
18 Nov'13	26 Sep'13	20 Sep'13	13 Oct'13	15 Oct'13	6 Oct'13	12 Oct'13	17 Nov'13	23 Sep'13	30 Sep'13	21 Sep'13
10 Mar'14	19 Dec'13	20 Dec'13	21 Dec'13	22 Dec'13	24 Feb'14	21 Feb'14	24 Feb'14	28 Jan'14	3 Feb'14	20 Dec'13
4 Nov'16	26 Oct'16	27 Oct'16	12 Aug'16	7 Aug'16	20 Oct'16	16 Aug'16	14 Aug'16	23 Oct'16	19 Aug'16	16 Aug'16
N_{cp}	15	14	13	14	17	15	12	18	13	11
total lag day	140 days	146 days	201 days	200 days	73 days	132 days	98 days	120 days	159 days	218 days

Table 7. Closest early warnings time of mount Merapi fourth eruption using the appropriate parameters and testing data (data magnitude earthquake from third eruption 11 May 2018 until 23 June 2018).

Actual eruption	(1-day)	(2-day)	(3-day)	(4-day)	(5-day)	(6-day)	(7-day)	(8-day)	(9-day)	(10-day)
	$\lambda = 96$	$\lambda = 18$	$\lambda = 7$	$\lambda = 5$	$\lambda = 5$	$\lambda = 5$	$\lambda = 4$	$\lambda = 6$	$\lambda = 6$	$\lambda = 4$
11 May'18	1 May'18	17 Jul'17	29 Apr'18	23 Jul'17	2 Feb'18	2 Mar'18	26 Jul'17	2 Sep'17	5 May'18	9 Aug'17
N_{cp}	4	2	4	4	5	5	3	2	6	3
before the eruption										
$total_{lag}_{cp}$	10 days	297 days	12 days	291 days	95 days	69 days	289 days	251 days	6 days	275 days

Some early warning methods mostly studied thresholds and period of the patterns before an eruption. Compared to other early warning methods, our approach is an online statistical changepoint method that uses the eruption probability increased over time. If the early warning is not detected, then the early warning probability is increased over time. It represents the volcano's time needed to through several stages before the eruption. Times of the changepoints are the early warnings signal, which detects the earthquake pattern's change since the last eruption. It provides the study that mount Merapi has a temporal change of seismic velocity, 4 months before eruption [29].

We use earthquake magnitude data to detect mount Merapi's eruption because mount Merapi often has earthquakes before the explosions. Further works, this study can be adapted to other volcano eruptions using its earthquake magnitude data or different data types. The data type is chosen based on the volcano characteristics (e.g., volume data of gases emitted). Using volume data of volcano gases emitted should consider how often these gases sign the outbreak. Further, the study about detected changepoints as true or false early warnings is needed. The eruption follows the true early warning, but no eruption follows the false early warning for a long time. We need the scheme to decide that the detection is a true early-warning or false early warning by analyzing the sequence of changepoints before the eruption.

Declarations

Author contribution statement

Seli Siti Sholihat: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper. Sapto Wahyu Indratno: Conceived and designed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data. Utriweni Mukhaiyar: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data.

Funding statement

This research was supported by Artificial Intelligence for Vision, Natural Language Processing and Big Data Analytics (U-CoE AI-VLB), Institut Teknologi Bandung, Indonesia.

Data availability statement

Data included in article/supplementary material/referenced in article.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

Acknowledgements

The authors also would like to express their sincerest thanks and appreciation to Dr. Irwan Meilano as an expert for discussing about the earthquakes data and volcano eruption. The first author would like to thank LPDP, through the scheme of BUDI-DN, for providing the scholarship.

References

- [1] R.I. Tilling, *Volcano monitoring and eruption warnings*, in: *Early Warning Systems for Natural Disaster Reduction*, Springer, 2003, pp. 505–510.
- [2] P.A. Mothes, H.A. Yepes, M.L. Hall, P.A. Ramón, A.L. Steele, M.C. Ruiz, *The scientific–community interface over the fifteen-year eruptive episode of Tungurahua Volcano, Ecuador*, *J. Appl. Volcanol.* 4 (1) (2015) 9.
- [3] A. Bertagnini, R. Cioni, E. Guidoboni, M. Rosi, A. Neri, E. Boschi, *Eruption early warning at Vesuvius: the ad 1631 lesson*, *Geophys. Res. Lett.* 33 (18) (2006).
- [4] P.A. Mothes, M.C. Ruiz, E.G. Viracucha, P.A. Ramón, S. Hernández, S. Hidalgo, B. Bernard, E.H. Gaunt, P. Jarrín, M.A. Yépez, et al., *Geophysical footprints of Cotopaxi's unrest and minor eruptions in 2015: an opportunity to test scientific and community preparedness*, in: *Volcanic Unrest, 2017*, p. 241.
- [5] S. Gambino, G. Falzone, A. Ferro, G. Laudani, *Volcanic processes detected by tiltmeters: a review of experience on Sicilian volcanoes*, *J. Volcanol. Geotherm. Res.* 271 (2014) 43–54.
- [6] S. Hall, *World's first automated volcano forecast predicts mount Etna's eruptions: system tracks infrasound waves to determine when an eruption is imminent—and alerts the Italian government*, *Nature* 563 (7732) (2018) 456–457.
- [7] M. Ripepe, E. Marchetti, D. Delle Donne, R. Genco, L. Innocenti, G. Lacanna, S. Valade, *Infrasound early warning system for explosive eruptions*, *J. Geophys. Res., Solid Earth* 123 (11) (2018) 9570–9585.
- [8] R. Matoza, D. Fee, D. Green, P. Mialle, *Volcano infrasound and the international monitoring system*, in: *Infrasound Monitoring for Atmospheric Studies*, Springer, 2019, pp. 1023–1077.
- [9] E. Marchetti, M. Ripepe, P. Campus, A. Le Pichon, N. Brachet, E. Blanc, P. Gaillard, P. Mialle, P. Husson, T. Arnal, *Infrasound monitoring of volcanic eruptions and contribution of arise to the volcanic ash advisory centers*, in: *Infrasound Monitoring for Atmospheric Studies*, Springer, 2019, pp. 1141–1162.
- [10] S. Carn, A. Krueger, G. Bluth, S. Schaefer, N. Krotkov, I. Watson, S. Datta, *Volcanic eruption detection by the total ozone mapping spectrometer (toms) instruments: a 22-year record of sulphur dioxide and ash emissions*, *Geol. Soc. (Lond.) Spec. Publ.* 213 (1) (2003) 177–202.
- [11] T.P. Wey, D.L. Woodell, *Airborne volcanic ash cloud and eruption detection system and method*, US Patent 7,383,131 (Jun. 3 2008).
- [12] F. Marchese, C. Filizzola, G. Mazzeo, R. Paciello, N. Pergola, V. Tramutoli, *Robust satellite techniques for thermal volcanic activity monitoring, early warning and possible prediction of new eruptive events*, in: *2009 IEEE International Geoscience and Remote Sensing Symposium*, vol. 2, IEEE, 2009, pp. II–953.

- [13] M. Rix, P. Valks, N. Hao, J. Van Geffen, C. Clerbaux, L. Clarisse, P.-F. Coheur, T. Erbertseder, W. Zimmer, S. Emmadi, et al., Satellite monitoring of volcanic sulfur dioxide emissions for early warning of volcanic hazards, *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 2 (3) (2009) 196–206.
- [14] B.A. Chouet, Long-period volcano seismicity: its source and use in eruption forecasting, *Nature* 380 (6572) (1996) 309–316.
- [15] R. White, W. McCausland, Volcano-tectonic earthquakes: a new tool for estimating intrusive volumes and forecasting eruptions, *J. Volcanol. Geotherm. Res.* 309 (2016) 139–155.
- [16] S. Gresta, F. Ghisetti, E. Privitera, A. Bonanno, Coupling of eruptions and earthquakes at Mt. Etna (Sicily, Italy): a case study from the 1981 and 2001 events, *Geophys. Res. Lett.* 32 (5) (2005).
- [17] D.P. Hill, F. Pollitz, C. Newhall, Earthquake-volcano interactions, *Phys. Today* 55 (11) (2002) 41–47.
- [18] A.J. Harris, M. Ripepe, Regional earthquake as a trigger for enhanced volcanic activity: evidence from modis thermal data, *Geophys. Res. Lett.* 34 (2) (2007).
- [19] T.R. Walter, R. Wang, M. Zimmer, H. Grosse, B. Lühr, A. Ratdomopurbo, Volcanic activity influenced by tectonic earthquakes: static and dynamic stress triggering at Mt. Merapi, *Geophys. Res. Lett.* 34 (5) (2007).
- [20] R. Turner, Bayesian change point detection for satellite fault prediction, in: *Diverse Engagement: Drawing in the Margins*, 2010, p. 213.
- [21] O. Niculita, Z. Skaf, I.K. Jennions, The application of Bayesian change point detection in UAV fuel systems, *Proc. CIRP* 22 (2014) 115–121.
- [22] R. Adams, D. Mackay, Bayesian online change point detection, University of Cambridge, Cambridge, 2007, Preprint at <http://arxiv.org/abs/0710.3742v1> [Google Scholar].
- [23] M. Raschke, Modeling of magnitude distributions by the generalized truncated exponential distribution, *J. Seismol.* 19 (1) (2015) 265–271.
- [24] L.J. Cobar, D. Legono, K. Miyamoto, Modeling of information flow for early warning in mount Merapi area, Indonesia, *J. Disaster Res.* 11 (1) (2016) 60–71.
- [25] V.R. Troll, D.R. Hilton, E.M. Jolis, J.P. Chadwick, L.S. Blythe, F.M. Deegan, L.M. Schwarzkopf, M. Zimmer, Crustal CO₂ liberation during the 2006 eruption and earthquake events at Merapi volcano, Indonesia, *Geophys. Res. Lett.* 39 (11) (2012).
- [26] P. Jousset, J. Pallister, M. Boichu, M.F. Buongiorno, A. Budisantoso, F. Costa, S. Andreastuti, F. Prata, D. Schneider, L. Clarisse, et al., The 2010 explosive eruption of Java's Merapi volcano—a '100-year' event, *J. Volcanol. Geotherm. Res.* 241 (2012) 121–135.
- [27] G. Richter, J. Wassermann, M. Zimmer, M. Ohrnberger, Correlation of seismic activity and fumarole temperature at the Mt. Merapi volcano (Indonesia) in 2000, *J. Volcanol. Geotherm. Res.* 135 (4) (2004) 331–342.
- [28] M. Ohrnberger, Continuous automatic classification of seismic signals of volcanic origin at Mt. Merapi, Java, Indonesia, Dissertation, 2001, p. 168.
- [29] A. Ratdomopurbo, G. Poupinet, Monitoring a temporal change of seismic velocity in a volcano: application to the 1992 eruption of Mt. Merapi (Indonesia), *Geophys. Res. Lett.* 22 (7) (1995) 775–778.
- [30] K.P. Murphy, Conjugate Bayesian analysis of the Gaussian distribution, Technical report, UBC, 2007.
- [31] S.C. Moran, C. Newhall, D.C. Roman, Failed magmatic eruptions: late-stage cessation of magma ascent, *Bull. Volcanol.* 73 (2) (2011) 115–122.