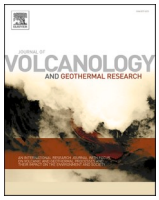
Journal of Volcanology and Geothermal Research 455 (2024) 108185

Contents lists available at ScienceDirect

Journal of Volcanology and Geothermal Research

journal homepage: www.journals.elsevier.com/journal-of-volcanology-and-geothermal-research



Statistical analysis of the ground deformation of Vulcanian explosions at Sakurajima volcano, Japan

Kyoka Ishii [a,b,*], Masato Iguchi [a,c]

[a] Sakurajima Volcano Research Center, Disaster Prevention Research Institute, Kyoto University, 1722-19 Sakurajima Yokoyama-cho, Kagoshima, Japan
[b] Aso Volcanological Laboratory, Institute for Geothermal Sciences, Graduate School of Science, Kyoto University, 5280 Kawayo, Minami-Aso, Kumamoto, Japan
[c] Crisis Management Bureau, Kagoshima City, 11-1, Yamashita-cho, Kagoshima, Japan

ARTICLE INFO

Keywords:
Vulcanian explosions
Ground deformation
Stochastic modeling
Probabilistic forecast

ABSTRACT

The forecast of pulsatory explosions during volcanic unrest periods is an essential issue for the assessment and mitigation of volcanic hazards. Although various precursors are detectable through geophysical and geochemical monitoring, difficulties remain in precisely constraining possible scenarios. A probabilistic approach is effective in assessing risk while considering various uncertainties. Sakurajima volcano characterized by frequent Vulcanian activity is one of the suitable fields for the probabilistic forecast of pulsatory explosions. Their inflation-deflation patterns of ground deformation related to Vulcanian explosions are useful for evaluating the imminence and size of the next event. The large database obtained from its vigorous activity can contribute to statistical analysis. In this study, aiming the probabilistic forecast of the timing and size of explosions, we investigated the duration of inflation and volume changes at the pressure source using strain records of over 5000 events of Sakurajima volcano. Then, a stochastic model was estimated to explain the distribution of these events. The log-logistic distribution was found to be an appropriate model for data distribution, indicating the presence of competing processes, such as pressurization and depressurization, in the conduit. The model parameters of the log-logistic distribution temporally fluctuated reflecting the volcanic activity, especially increasing the magma supply from a deep region. We also suggested a methodology to constrain the probabilities of the likely timing and size of an imminent explosion using real-time strain monitoring and an estimated model distribution. Although some improvements would be needed for practical forecasting, our approach could be useful in predicting possible ash hazards.

1. Introduction

During an eruptive period, pulsatory eruptive events are observed with short time intervals from minutes to days in many volcanoes. Strombolian explosions and Vulcanian explosions are representatives of pulsatory eruptive styles (Taddeucci et al., 2015; Voight et al., 1998). Pyroclastic flows due to lava dome collapse (Nakada et al., 1999) and the lava overflows (Bonaccorso et al., 2020) could also occur intermittently with short-term intervals. Those pulsatory events tend to occur in the same area keeping the similar eruptive style. Therefore, information on the timing and size of the next single event is important for the assessment and mitigation of volcanic hazard. This study focuses on the forecast of such recurring individual events.

Geophysical and geochemical monitoring of volcanic activity is an effective way to forecast pulsatory events by detecting precursors, such as seismicity, ground deformation, and gas emissions (Swanson et al., 1983; Sparks, 2003). Especially continuous monitoring of ground deformation can directly capture pressure build-up due to magma movement to shallow regions before explosions (Voight et al., 1998; Nishimura et al., 2012; Lyons et al., 2012). At Stromboli volcano, tilt change is observed prior to each Strombolian explosion (Genco and Ripepe, 2010), and the duration and amplitude of ground inflation scales with the volume of erupted materials (Ripepe et al., 2021). The borehole dilatometer network on Etna volcano also shows significant advantages in detecting precursory ground inflation and deflation related to repeated lava fountains (Currenti and Bonaccorso, 2019; Bonaccorso et al., 2020). The amount of strain change correlates with the erupted volume calculated from satellite images (Bonaccorso et al.,

* Corresponding author at: Aso Volcanological Laboratory, Institute for Geothermal Sciences, Graduate School of Science, Kyoto University, 5280 Kawayo, Minami-Aso, Kumamoto, Japan.
E-mail address: ishii.kyoka.7v@kyoto-u.ac.jp (K. Ishii).

https://doi.org/10.1016/j.jvolgeores.2024.108185
Received 5 September 2023; Received in revised form 30 August 2024; Accepted 1 September 2024
Available online 3 September 2024

2023). As these previous studies suggest, monitoring of durations and/or amounts of ground deformation leads to predicting the timing and size of the next explosions.

Sakurajima volcano in Japan is also a suitable field for the forecast of pulsatory explosions using ground deformation observation. Over the last 60 years, its volcanic activity has been characterized by frequent Vulcanian explosions (Iguchi et al., 2013). Ground deformation is typically observed before and after an explosion (Ishihara, 1990; Iguchi et al., 2008, 2013). Prior to the explosions, the inflation of the pressure source (<4 km depth) can usually be observed using tilt meters and extensometers installed in the underground observation tunnels. Immediately after the explosion, the pressure source is deflated owing to depressurization in the conduit. Deformation data, particularly pre-eruptive inflation data, have contributed to forecasting the occurrence of explosions since the 1980s (Kamo and Ishihara, 1989; Le et al., 2020). In addition, Iguchi (2016) revealed a linear relationship between the deflation volume of the pressure source estimated from strain data and the weight of erupted volcanic ash. Iguchi (2016) proposed an empirical formula to estimate the weight of erupted ash using ground deformation data. It is also possible to perform ash fall numerical simulations using the estimated ash weight as an input parameter (Iguchi et al., 2022a).

However, it is still challenging to deterministically forecast the timing and size of the next pulsatory event when ground inflation is currently being detected. All we know is that the next explosion is "soon," but in many cases, we cannot definitively constrain how long this inflation will last and how much the amount of inflation will reach. This is due to the lack of a well-developed physical model of pulsatory explosions and the inherent uncertainty of the eruptive phenomenon itself. Therefore, a probabilistic approach is needed to forecast the explosions.

The methodology of the probabilistic forecast has been particularly advanced in the explosion timing prediction using repose interval databases (Marzocchi and Bebbington, 2012). Repose interval is usually defined as a time interval between the onsets of consecutive explosions (Dominguez et al., 2016). The repose intervals of Vulcanian explosions were investigated at Soufrière Hills (Connor et al., 2003), Tungurahua (Varley et al., 2006), Santiaguito (Dominguez et al., 2016), Sabancaya (Fries et al., 2023), and Sakurajima (Udagawa et al., 1999; Dominguez et al., 2016; Jenkins et al., 2019). The distributions of repose intervals at those volcanoes were fitted by several stochastic models, such as Weibull, log-logistic, log-normal, gamma, or exponential distributions, to calculate the probability of the next event. The most important point of this approach is that it allows us to obtain a probability density function, which makes it possible to calculate the probability of an explosion based on past activity trends and quantify the risk of explosions. This type of time-series data analysis is called survival analysis, which is a statistical method that deals with the length of time it takes for a particular event to occur. It is applied in a wide range of fields, including the financial and engineering fields, regardless of volcanology.

We aim to apply a probabilistic approach to the forecast of the timing and size of Vulcanian explosions using ground deformation data of Sakurajima volcano. The duration of the ground deformation preceding the explosion can be viewed as the time to event occurrence, which can be treated in the survival analysis. Furthermore, as previous studies have pointed out, the amount of ground deformation is correlated with the amount of erupted material. Therefore, the statistical analysis of the ground deformation data of past explosions should lead to a probabilistic forecast of the timing and size of explosions. This paper focused on clarifying the statistical properties of the time scale and amount of ground deformation, particularly estimating their probability density functions. We investigated four variables estimated from strain data of Vulcanian explosions: (i) duration of inflation, (ii) inflation volume at the pressure source, (iii) deflation volume at the pressure source, and (iv) ratio of deflation volume to inflation volume. Then, we estimated appropriate stochastic models to explain the distributions of these variables. We finally discussed some insights from the estimated model and a method to evaluate the probabilities of explosions.

2. Data and methods

2.1. Data acquisition and overview

We created a database for statistic analysis that included (i) duration of inflation, (ii) inflation volume of the pressure source, (iii) deflation volume of the pressure source, and (iv) deflation ratio (= deflation volume/inflation volume). These four variables were calculated from strain data of Vulcanian explosions observed at the underground observation tunnel. Sakurajima volcano has three underground observation tunnels with 28-m long extensometers (Fig. 1a). The strain data of AVOT station is used for this study since the strain data of this station were more sensitive to pressure changes in the shallow region than data of other tunnels (Iguchi et al., 2013). The data were automatically processed for the removal of ocean tide effects using sea-level observations and a Bayesian method (BAYTAP-G; Tamura et al., 1991) and can be viewed in real-time online with a 1-min resolution.

Fig. 1b shows two components of strain data associated with typical Vulcanian explosions. The radial strain is a strain of the direction toward craters, and the tangential strain is that of the perpendicular direction to the radial strain. Before the explosion, a contraction in the radial strain and an extension in the tangential strain were typically observed (Fig. 1b). These features are interpreted as inflation of a shallow spherical source under the crater (Iguchi et al., 2008, 2013). From the time the explosion occurred, the radial strain rapidly increased. Simultaneously, the tangential strain contracted. These trends conversely correspond to the deflation of the shallow spherical pressure source due to the emission of ash and gas into the atmosphere.

There are two active craters at Sakurajima volcano, Minami-dake and Showa craters (Fig. 1a). Minami-dake crater began its Vulcanian activity since 1955 until now (2024). The Vulcanian explosions at Showa crater occurred from 2006 to 2018 with a brief increase in explosions in 2023. This study targeted the explosions at Showa crater from February 2009 to September 2015 and at Minami-dake crater from November 2017 to December 2020 (Fig. 1c). Because the period from October 2015 to October 2017 corresponds to the period when eruptive activity shifted from Showa crater to Minami-dake crater, it was excluded from the analysis.

We listed explosion events with large infrasound signals (>10 Pa) recorded at the HAR or AVOT stations (Fig. 1a; Iguchi et al., 2022b). Then, a database was created from the strain data of the listed events as written in the next section. Some exceptional cases were not used for the following analysis, such as missing strain data and explosions with unclear deformation. We identified 4970 events at Showa crater and 710 events at Minami-dake crater as the analysis events, which accounted for 86 % of all listed explosions (Fig. 1c).

2.2. Database creation

The four variables included in the database (duration of inflation, inflation volume, deflation volume, and deflation ratio) were defined as follows. The duration of inflation is the time since the strain change due to ground inflation is detected until the explosion starts (Fig. 1b). The inflationary phase is defined by a contraction of radial strain and an extension of tangential strain. Note that the duration of inflation is not the same as the repose interval which corresponds to the time interval between the onsets of consecutive explosions. The duration of inflation does not include the period when the explosion is occurring, in which deflation is detected. In addition, the inflation for the next explosion does not always begin immediately after the end of the previous explosion. In such cases, the strain data remain flat for a while after a previous eruption or small-scale fluctuations that are not counted as explosions occur (Fig. 1b). Thus, the duration of inflation tends to be shorter than the repose interval. Later discussion will deal with the

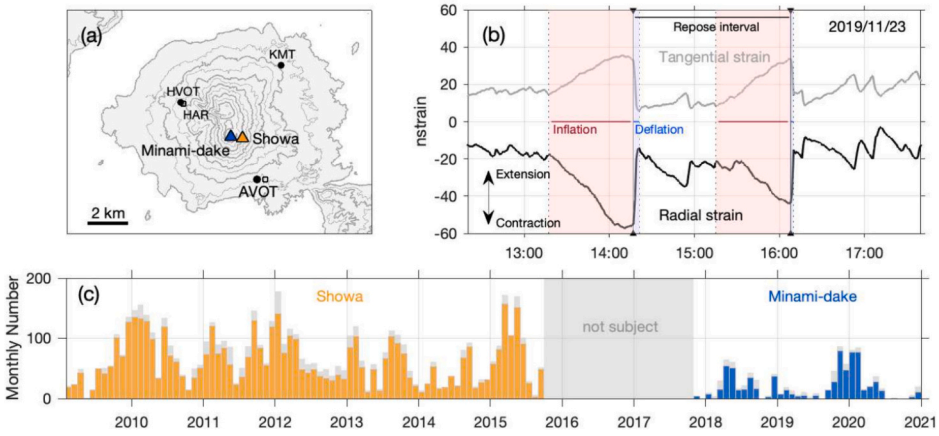


Fig. 1. Vulcanian explosions at Sakurajima volcano.
(a) Map around Sakurajima volcano. Two triangles indicate the active craters, Minami-dake and Showa craters. Solid circles represent underground tunnels (HVOT, AVOT, and KMT). In this study, we used the data recorded at AVOT. Open squares indicate low-frequency microphones.
(b) Snapshot of strain data at AVOT on November 23, 2019 (black: radial strain, gray: tangential strain). Black triangles indicate the time of the explosion. In the inflation phase of the pressure source before the explosion, the radial strain contracts and the tangential strain extends (hatched by pink). On the contrary, the radial strain extends and the tangential one contracts in the deflation phase after the explosion.
(c) Monthly number of explosions. Orange bars are events that occurred at Showa crater and blue bars are those that occurred at Minami-dake crater. Gray bars indicate events that are not used for analysis owing to unclear deformation data. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

repose intervals investigated in several previous studies, which will be used only for a comparison with the duration of inflation.

The second and third variables, the inflation and deflation volumes of the pressure source, were calculated using both the radial and tangential strains (Iguchi et al., 2013, 2022b) assuming a spherical source located beneath the crater (Mogi model; Mogi, 1958). Radial strain (ϵ_r) and tangential strain (ϵ_t) derived from a spherical pressure source are written below (Ishihara, 1990):

$$\epsilon_r = K\frac{D^2 - 2r^2}{\left(D^2 + r^2\right)^{5/2}}, \tag{1}$$

$$\epsilon_t = K\frac{1}{\left(D^2 + r^2\right)^{3/2}}, \tag{2}$$

where D is the depth of the source and r is the horizontal distance from the source to the station. K is a function of Lame constants (λ, G), source radius (a), and pressure change (ΔP):

$$K = \frac{\lambda + 2G}{2G(\lambda + G)}a^3\Delta P. \tag{3}$$

Although shear modulus G is often written as μ, G is used here since μ is used as a parameter of the statistical model in the next section. By dividing Eq.(1) by Eq.(2), the source depth is obtained (Iguchi et al., 2013):

$$D = \sqrt{\frac{2+\beta}{1-\beta}}r, \tag{4}$$

where $\beta \equiv \epsilon_r/\epsilon_t$. The volume change on the surface (ΔV_s) is represented by integrating the vertical displacement as $2\pi K$ (Ishihara, 1990). The volume change of the pressure source (ΔV_p) is given the relationship between volume change on the surface and that at the pressure source (Delaney and McTigue, 1994):

$$\Delta V_p = \frac{\Delta V_s}{2(1-\nu)} = \frac{2\pi K}{2(1-\nu)}, \tag{5}$$

where ν is the Poisson rate (0.25 in this study). Therefore, ΔV_p can be described as a function of the strain changes using Eqs.(1) and (4),

$$\Delta V_p = \frac{\pi\epsilon_r\left(D^2 + r^2\right)^{5/2}}{(1-\nu)\left(D^2 - 2r^2\right)}. \tag{6}$$

The inflation volume and deflation volume included in our database were defined as ΔV_p in the inflation and deflation phases, respectively (Fig. 1b). Iguchi et al. (2019) reported that the inflation volumes of the pressure source for explosions at Minami-dake crater are 4×10^3–10^5 m^3. Inflation volumes for explosions at Showa crater are typically less than 2×10^4 m^3, so they are generally smaller than the volume changes for Minami-dake craters' explosions.

In the case of Sakurajima volcano, deflation volume is strongly related to the size of the explosion. Iguchi (2016) revealed that the deflation volume of Vulcanian explosions and the seismic amplitude of volcanic tremors correlate positively with the monthly summed weight of ejected volcanic ash. An empirical formula was proposed to estimate the monthly weight of volcanic ash, which is composed of a linear combination of terms of the deflation volume and seismic amplitude. Concerning the weight of volcanic ash derived from an individual Vulcanian explosion, it strongly depends on the deflation volume rather than on the seismic amplitude because the seismic amplitude of volcanic tremors indicates the contribution of non-explosive continuous ash emission lasting for a long time.

Therefore, the fourth variable is the deflation ratio which is defined as the deflation volume relative to the inflation volume. All we can know before an explosion is not the deflation volume but the inflation volume so that the ratio between inflation and deflation volumes is needed to predict the amount of erupted materials. In Sakurajima volcano, the explosion does not always cancel all amounts of the precursory inflation, and a part of the inflation may remain.

2.3. Statistical analysis

After creating the database, we estimated stochastic models for the distributions of the four listed variables. Assuming several candidate stochastic models, the best model was determined. This procedure follows the conventional approach for the survival analysis of repose intervals (Marzocchi and Bebbington, 2012). Survival analysis is a field of statistics dealing with the time until an event occurs, such as the time to recover from disease and the time to failure of mechanics. In

volcanology, survival analysis is often applied to repose intervals since intermittent eruptive activity is assumed as a renewal process where the elapsed time since the last event is the only control factor for the time of the next event. The duration of inflation dealt with in this study is related to the time until the events, so survival analysis can be applied. The other three variables of our study are not the variables of time. However, we estimated the stochastic models for these three variables as well. This is because our purpose is to statistically describe the characteristics of the distribution of ground deformation data, in particular, to estimate the probability density function.

Model estimation was performed on three different-sized data sets: 1) all events of each crater, 2) sub-datasets extracted from eruptive episodes, and 3) every 100 events. At first, we estimated stochastic models for the four variables using the data of 1) all events for Minami-dake and Showa craters in order to get an overall picture of the data. However, it generally needs to confirm the stationarity and independence of the data when conducting survival analysis (Cox and Oakes, 1984). Therefore, as a second step, we divided the analysis period into several eruptive episodes. Then, 2) sub-datasets were extracted from each episode, and the best stochastic model for sub-datasets was estimated.

The eruptive episodes were defined considering the temporal change in deflation rate which corresponds to the slope of the cumulative deflation volume due to individual Vulcanian explosions (Fig. 2, Supplementary material 1). The analysis period was divided into 14 episodes, of which nine were related to the activity at Showa Crater and five to the activity at Minami-dake. The deflation rate was high in even-numbered episodes, and that was low in odd-numbered episodes.

Sub-dataset is defined as a shorter period included in every eruptive episode indicated in the top panels of Fig. 2, where independence and stationarity of the data are guaranteed. The stationary of the sub-datasets were evaluated by using moving-average (Dominguez et al., 2016). Moving averages for the sub-datasets were confirmed to be within the mean $\pm\sigma$ for each episode. As a simple test of data independence, autocorrelations were calculated using each dataset (Supplementary material 2).

After the model estimation for the sub-datasets, we also estimated the best model for 3) every 100 events by sliding the analysis window by one event. This is intended to eliminate arbitrariness in extracting the sub-datasets and to investigate more detailed temporal changes in the best model.

Three candidate stochastic models were assumed: log-logistic, Weibull, and exponential distributions. These three models were selected because they are often used as survival analysis for repose intervals (Marzocchi and Bebbington, 2012). The log-logistic distribution is a probability distribution of a variable whose logarithm has a logistic distribution. It can describe a system where two competing processes work: some parameters work to increase the probability of explosions, and others work to decrease the probability (Connor et al., 2003). The Weibull distribution (Weibull, 1951) is a model for the time to failure used to forecast brittle fractures and machine failure. The exponential distribution is related to the Poisson process.

The probability density functions (PDFs) of the three candidate models are listed in Table 1. The PDF of the exponential distribution is controlled by one parameter μ, whereas those of log-logistic and Weibull distributions are controlled by two parameters, μ and k. μ is the scale parameter, which corresponds to the mean (Weibull and exponential) or median (log-logistic). The shape of the distribution depends on the shape parameter k. If $k \leq 1$, then the distribution monotonically decreases. If $k > 1$, the distribution exhibits a single peak. A higher k value indicates a narrow distribution. The shape parameter is defined as an indicator of the variability of a distribution like standard deviation, and Dominguez et al. (2016) argued that the shape parameter for the distribution of the repose intervals could be interpreted as an indicator of the regularity of

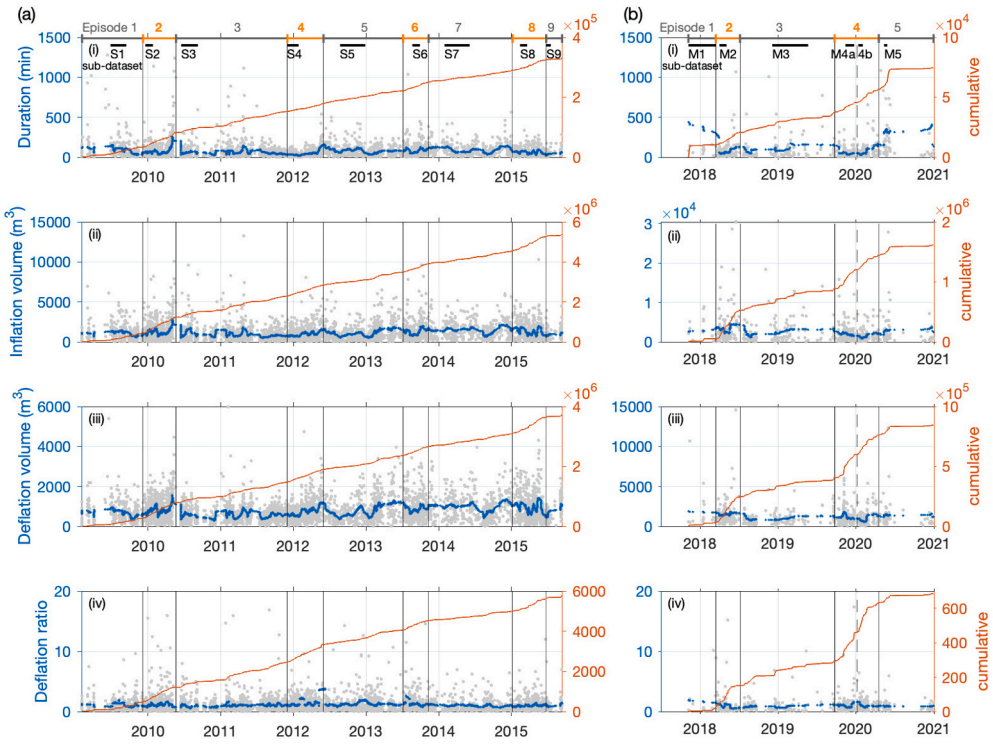


Fig. 2. Temporal evolution of the four variables of our dataset: (i) duration of inflation, (ii) inflation volume, (iii) deflation volume, and (iv) deflation ratio (=iii/ii). Gray markers are the values of each event and blue plots are moving average every 50 events. The orange lines show their cumulative value. (a) The active period of Showa crater and (b) that of Minami-dake crater. These analysis periods were divided into eruptive episodes. Even-numbered episodes correspond to periods of high deflation rates. Sub-datasets (S1–S9 and M1–M5) were extracted from every episode, which is indicated by black bars in the top panels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
Stochastic models and their model parameters.

		Log-logistic		Weibull		Exponential	
PDF		$f(t) = \frac{kt^{k-1}\left(\frac{1}{\mu}\right)^k}{\left[1+\left(\frac{t}{\mu}\right)^k\right]^2}$		$f(t) = \left(\frac{k}{\mu}\right)\left(\frac{t}{\mu}\right)^{k-1} e^{-\left(\frac{t}{\mu}\right)^k}$		$f(t) = \frac{1}{\mu} e^{-\frac{t}{\mu}}$	
		parameter	AIC	parameter	AIC	parameter	AIC
Showa	(i) Inflation duration	$\mu = 44.45$ $k = 1.97$	50,713	$\mu = 68.90$ $k = 1.09$	51,569	$\mu = 66.34$	51,638
	(ii) Inflation volume	$\mu = 774.7$ $k = 2.04$	78,643	$\mu = 1161$ $k = 1.22$	79,051	$\mu = 1081$	79,377
	(iii) Deflation volume	$\mu = 601.6$ $k = 2.38$	74,508	$\mu = 834.2$ $k = 1.48$	74,619	$\mu = 750.0$	75,746
	(iv) Deflation ratio	$\mu = 0.78$ $k = 2.10$	9854.1	$\mu = 1.17$ $k = 1.02$	11,392	$\mu = 1.16$	11,397
Minami-dake	(i) Inflation duration	$\mu = 43.92$ $k = 1.83$	7741.0	$\mu = 79.29$ $k = 0.76$	7839.9	$\mu = 105.0$	8031.0
	(ii) Inflation volume	$\mu = 1355$ $k = 1.76$	12,278	$\mu = 2253$ $k = 0.98$	12,398	$\mu = 2272$	12,396
	(iii) Deflation volume	$\mu = 883.4$ $k = 2.30$	11,279	$\mu = 1289$ $k = 1.24$	11,421	$\mu = 1191$	11,479
	(iv) Deflation ratio	$\mu = 0.65$ $k = 2.08$	1170.4	$\mu = 0.99$ $k = 1.06$	1372.4	$\mu = 0.97$	1375.2

the pulsatory explosions: a higher value of k means less variability and more regularity of the explosions. The model parameters for each model are estimated by using maximum likelihood estimation.

After estimating these model parameters, we determined the best model among the three using the Akaike information criterion (AIC; Akaike, 1974). AIC is defined as AIC $= 2n - 2lnL$, where n is the number of model parameters, and L is the maximum likelihood. A smaller AIC value indicates a better model. In addition to determining the best model, Kolmogorov-Smirnov test (KS test) was also applied to confirm the assumption that the data follow the model distribution. If the p-value obtained in the KS test is less than the significance level (here $p = 0.05$), it can be concluded that the distribution of the data is different from the assumed model.

3. Results

The four variables of our database, (i) duration of inflation, (ii) inflation volume, (iii) deflation volume, and (iv) deflation ratio are shown as a time series in Fig.2. For most of the analysis period, the moving averages of the durations of inflation are less than 200 min, while some events have long durations of over a half day. The inflation volumes of the explosions at the Showa crater tend to be smaller than those at Minami-dake crater. The moving averages of inflation volumes of Showa crater are less than 2000 m^3 for most periods, and those of Minami-dake are more than 2000 m^3. The trend of fluctuations in the moving averages of the deflation volume was similar to that of the inflation volume. The absolute values of the moving averages of the deflation volume tended to be smaller than those of the inflation volume.

Model estimation was performed using three-sized datasets for four variables. These results are presented in order below. Supplementary material 3 also summarizes the datasets used for figures and tables. Fig. 3 shows the probability distributions of the four variables for all events of each crater. We found the log-logistic distribution to be the best among the three candidate models for all data from both craters. Because all data distributions had a single peak, the exponential distribution did not match the data. The log-logistic distribution has a heavier tail than that of the Weibull distribution, making it suitable for

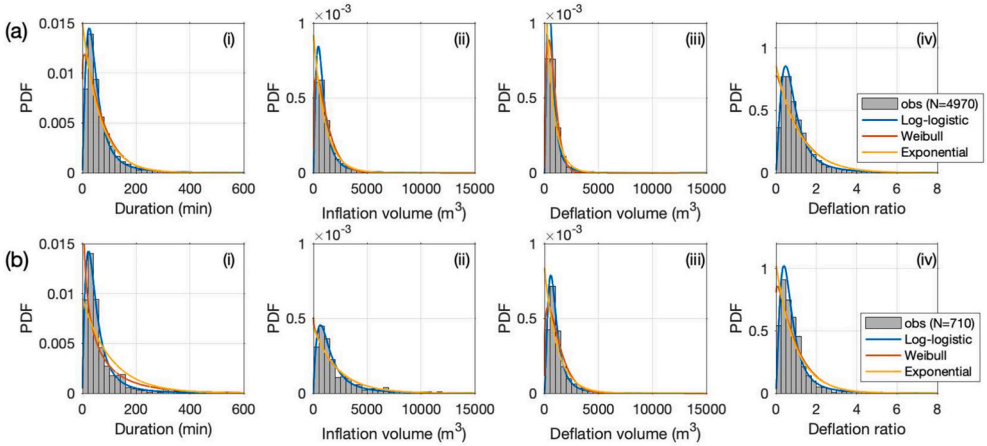


Fig. 3. Distributions of observed data and stochastic models for all explosions at each crater.
(i) Duration of inflation, (ii) inflation volume, (iii) deflation volume, and (iv) deflation ratio. The gray histograms are the distributions of the observed data of events that occurred at (a) Showa and (b) Minami-dake craters. The vertical axis represents the probability density function (PDF). The blue, red, and yellow lines represent the estimated stochastic models of the log-logistic, Weibull, and exponential distributions, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

describing our dataset. The parameters and AICs of the three models are presented in Table 1. In terms of the log-logistic distribution, the values of the shape parameter k were estimated to be approximately 2 for all variables. The scale parameter μ of the duration of inflation did not significantly differ between the Showa and Minami-dake craters; μ was approximately 44 min, and over 80 % of all events had durations of less than 100 min. The scale parameter of the deflation ratio also showed no significant differences between the two craters, with 60 % of the events being less than one. However, the scale parameters of inflation and deflation volume of Minami-dake were larger than those of Showa crater.

Next, we set sub-datasets for the statistical analysis, where the moving average did not fluctuate significantly and the autocorrelation function fluctuated around 0. The best models of the four variables for each sub-dataset are shown in Table 2 and Supplementary material 4. Most datasets follow a log-logistic distribution in terms of the duration of inflation, inflation volume, and deflation ratio. Especially concerning the deflation ratio, the best models are the log-logistic distribution for all sub-datasets. On the other hand, the best models for the deflation volume are often estimated to be Weibull distribution (Table 2). Since the p-value of KS test for all subset are over 0.05, all the best models were not rejected. In addition, KS test shows that log-logistic distribution are not rejected for all the subset (Supplementary material 5). The model parameters of log-logistic distribution, μ and k, for sub-datasets are shown in Fig. 4. The shape parameter k for all datasets is within 1–3. It described unique temporal change: k increased in even-numbered sub-datasets (S2, S4, S6, S8, M2, and M4), especially for the duration of inflation. In addition, an inverse correlation between μ and k is observed for the duration of inflation: as k increases, μ becomes smaller.

More detailed temporal changes in the best model every 100 events are shown in Fig. 5. The analysis windows in which the log-logistic distribution was the best model stood out. The Weibull distribution was better than the log-logistic distribution in some periods, particularly with respect to the deflation volume. Very few analysis windows were present in which the exponential distribution was the best model. Focusing on the active period of the Minami-dake crater (since 2017), the log-logistic distribution was appropriate for almost all periods for all four variables. According to the Kolmogorov-Smirnov test, the exponential distribution is rejected in almost all the analysis windows unlike the log-logistic distribution (Supplementary material 5).

The model parameters of the log-logistic distribution for every 100 events temporally fluctuated as shown in Fig. 6. The shape parameter k fluctuated within 1–4 over several months and reached large values in Episode 4 of Showa crater and Episode 4 of Minami-dake crater. During these episodes, k increased and then decreased. At the same time, the scale parameter μ decreased and then increased. Particularly in terms of duration of inflation and inflation volume, when k becomes larger, μ takes a smaller value (Fig. 6c, d). This suggests that the high regularity of explosions reduces their duration of inflation and inflation volume. We also recognize a similar feature during the other even-numbered episodes.

4. Discussion

Based on the results, three topics are discussed. First, we discuss the insights for the eruptive activities derived from the estimated stochastic model (Section 4.1). We then suggest a methodology to constrain the probabilities of the timing and size of the explosions using the estimated model distribution (Section 4.2). Finally, in Section 4.3, we highlight certain considerations for the practical application of eruption forecasts.

4.1. Best stochastic model and eruptive activities

Our results demonstrated that a log-logistic distribution is often a suitable model for various-sized datasets of the time scale and deformation volumes derived from the ground deformation of Vulcanian

Table 2
Model distributions fitted to the data of subsets.

Sub-datasets	Number of events	Model*	μ	k	p value
(i) Inflation duration					
S1 (2009/6/26–2009/9/13)	141	W	104.3	1.21	0.618
S2 (2009/12/18–2010/1/25)	215	L	45.6	2.19	0.664
S3 (2010/6/16–2010/9/8)	266	L	42.4	2.07	0.866
S4 (2011/12/1–2012/1/27)	256	L	25.8	2.83	0.616
S5 (2012/8/23–2012/12/27)	150	L	70.8	2.17	0.300
S6 (2013/8/21–2013/9/28)	134	L	54.4	2.62	0.976
S7 (2014/1/29–2014/6/4)	120	L	61.7	2.06	0.275
S8 (2015/2/13–2015/3/21)	87	L	64.6	2.75	0.856
S9 (2015/6/25–2015/7/18)	54	L	23.6	2.05	0.989
M1 (2017/11/7–2018/3/15)	16	L	78.4	1.48	0.491
M2 (2018/4/1–2018/5/4)	62	W	59.3	1.68	0.765
M3 (2018/12/4–2019/5/22)	75	L	50.7	1.79	0.798
M4a (2019/11/13–2019/12/24)	94	L	33.7	2.58	0.952
M4b (2020/1/12–2020/2/2)	82	L	28.9	2.55	0.892
M5 (2020/5/11–2020/5/26)	26	E	161.8		0.799
(ii) Inflation volume					
S1		W	1374	1.27	0.726
S2		L	883.0	2.25	0.767
S3		L	674.3	1.98	0.937
S4		L	588.7	2.52	0.769
S5		W	1029	1.33	0.783
S6		L	775.9	2.56	0.818
S7		W	1754	1.59	0.940
S8		W	1799	1.52	0.670
S9		L	607.7	2.54	0.991
M1		E	3394		0.806
M2		L	2189	2.08	0.789
M3		E	2558		0.375
M4a		L	1400	2.20	0.697
M4b		L	762.2	2.31	0.917
M5		W	1897	1.30	0.619
(iii) Deflation volume					
S1		W	865.0	1.84	0.676
S2		L	633.2	2.86	0.845
S3		W	623.3	1.71	0.525
S4		W	629.9	2.11	0.924
S5		W	787.1	1.43	0.520
S6		L	700.2	2.79	0.593
S7		W	1159	1.93	0.706
S8		W	1238	1.54	0.796
S9		W	532.2	2.00	0.982
M1		L	1074	2.17	0.396
M2		W	1706	1.99	0.918
M3		L	754.9	2.63	0.969
M4a		W	1132	1.58	0.712
M4b		W	813.2	1.98	0.556
M5		W	1165	1.95	0.998
(iv) Deflation ratio					
S1		L	0.711	2.08	0.870
S2		L	0.712	2.23	1.00
S3		L	0.685	1.93	0.972
S4		L	0.849	2.26	0.911
S5		L	0.781	1.72	0.923
S6		L	0.884	2.08	0.939
S7		L	0.697	2.10	0.305
S8		L	0.714	2.15	0.985
S9		L	0.647	2.41	0.996

(continued on next page)

Table 2 (*continued*)

Sub-datasets	Number of events	Model*	μ	k	p value
M1		L	0.574	1.36	0.811
M2		L	0.619	2.12	0.942
M3		L	0.484	1.87	1.00
M4a		L	0.596	2.12	0.976
M4b		L	0.818	2.48	0.502
M5		L	0.724	2.28	0.994

* L: Log-logistic, W: Weibull, E: Exponential.

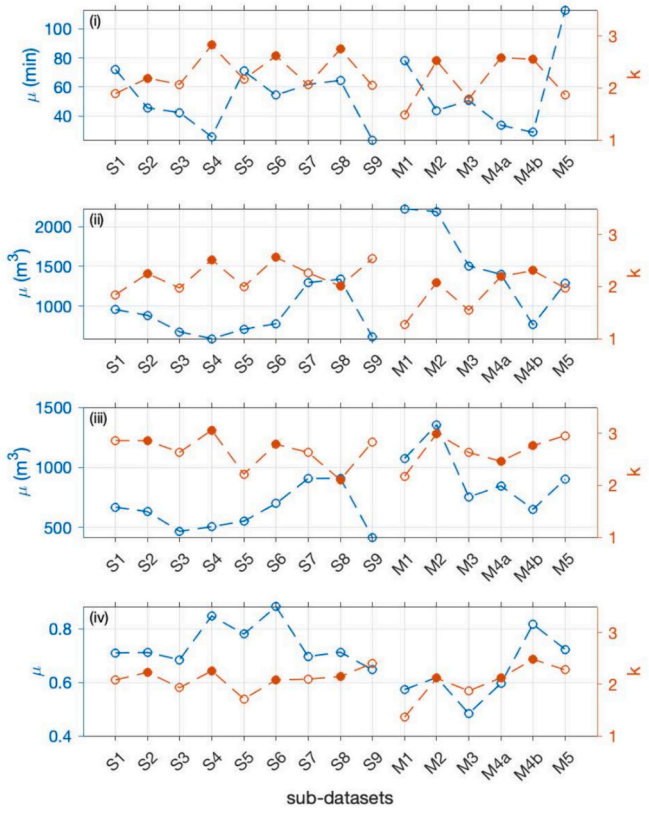


Fig. 4. Model parameters of the log-logistic distribution for sub-datasets. The blue markers indicate the scale parameter μ, and the red markers indicate the shape parameter k. Filled markers of k correspond to the sub-datasets in even-numbered episodes with high deflation rates. (i) Duration of inflation, (ii) inflation volume, (iii) deflation volume, and (iv) deflation ratio. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

explosion at Sakurajima volcano. The log-logistic distribution is a model in which logarithmic variables follow a logistic distribution. The logistic distribution is originally derived from a model of population growth which models a system in which some factors increase the population and others decrease the population (Verhulst, 1838). Thus, the log-logistic distribution is inherently suitable to model systems where competing processes coexist. The fact that the log-logistic distribution is the best model for the ground deformation database of Sakurajima volcano suggests that competing processes are the main factor controlling the time scale and amount of ground deformation. Connor et al. (2003) showed that the repose intervals of Vulcanian explosions at Soufrière Hills volcano are explained by the log-logistic distribution. It was interpreted that the success of modeling by the log-logistic distribution is due to competing processes with different time scales operating in the upper volcano conduit, such as increasing and decreasing pressure. Prior to Vulcanian explosions, magma ascends and stiffens, thus increasing pressure in the shallow conduit. On the other hand, upward and lateral degassing through permeable magma and fractures of plumbing systems decreases the pressure (Diller et al., 2006). In the case of Sakurajima volcano, it has been proposed that magma intrusion to the shallow conduit and gas accumulation provide pressurization, and gas leakage provides depressurization in the conduit (Iguchi et al., 2008; Yokoo et al., 2013). Therefore, our findings support the idea that the balance of such competing processes in the conduit can control the time scale and magnitude of ground deformation.

The Weibull distribution was the best model in some periods. In particular, the distributions of deflation volume follow the Weibull distribution distinctly more often than the other variables. It may suggest that the controlling factors of the deflation process are different from those of the inflation process. Once the explosion occurs, the destruction of the lava plug disrupts the previous balance of pressurization and depressurization in the conduit, releasing the system. Such changes in the system might be reflected in the estimated models.

Although it is difficult to judge which is more appropriate between log-logistic or Weibull distributions considering only AIC (Supplementary material 5), in terms of applicability for the evaluation and forecast of explosions at Sakurajima volcano, the log-logistic distribution is likely to be better. One reason is that many previous studies dealing with repose intervals of Vulcanian explosions use the log-logistic distribution (e.g., Dominguez et al., 2016). In addition, the log-logistic distribution was not rejected throughout the analysis period by the KS test, so it is highly versatile for the ground deformation data of Sakurajima volcano. Moreover, the cost to calculate the probability of the explosion is lower when the log-logistic distribution is used. This is because the cumulative distribution function and survival function (explained in the next section) of the log-logistic distribution can be described analytically so that no numerical integration is required. Therefore, we propose the use of log-logistic distribution as a probability distribution for the time scale and volume change of Sakurajima's ground deformation data.

A log-logistic distribution is often applied to the repose interval data of other volcanoes (e.g., Connor et al., 2003; Varley et al., 2006; Watt et al., 2007; Dominguez et al., 2016; Fries et al., 2023). Although the repose interval between explosions has a different definition from the duration of inflation, both were used to evaluate the over-pressure build-up process, so a comparison of them would provide some insight. Fig. 7 shows the model parameters of the log-logistic distribution for the repose intervals of various volcanoes reported in the previous studies. Note that Supplementary material 6 provides correspondence of the notations of the model parameters between this study and the previous studies. We recognize that μ and k are negatively correlated. This trend progressively represents the more frequent and regular nature of Strombolian and violent Strombolian eruptions and the less frequent and irregular nature of Vulcanian eruptions (Dominguez et al., 2016). The parameters estimated from the inflation duration in this study were within the range of Vulcanian eruptions. Compared with the parameters estimated from the repose interval of Sakurajima volcano in previous studies, k became larger (higher regularity), and μ became smaller (shorter duration) in this study. The inflation duration more directly reflects the over-pressure build-up than the repose interval, which may have resulted in a shorter timescale. The repose phase included periods without clear ground deformation and those with small-scale ground deformation that were not counted as explosions (Fig. 1b). This implies that the repose intervals obtained from the event catalogs may overestimate the time scale of the over-pressure build-up process. Using the duration of inflation rather than the repose interval reduces the effect of such an overestimation, which may have resulted in a clearer regularity of the phenomenon and a larger k. Nevertheless, ground deformation data is not necessarily available for all volcanoes. A more accurate understanding of the over-pressure build-up process may become possible even for volcanoes without ground deformation data by closely examining the correspondence between the time scale of ground deformation and repose intervals.

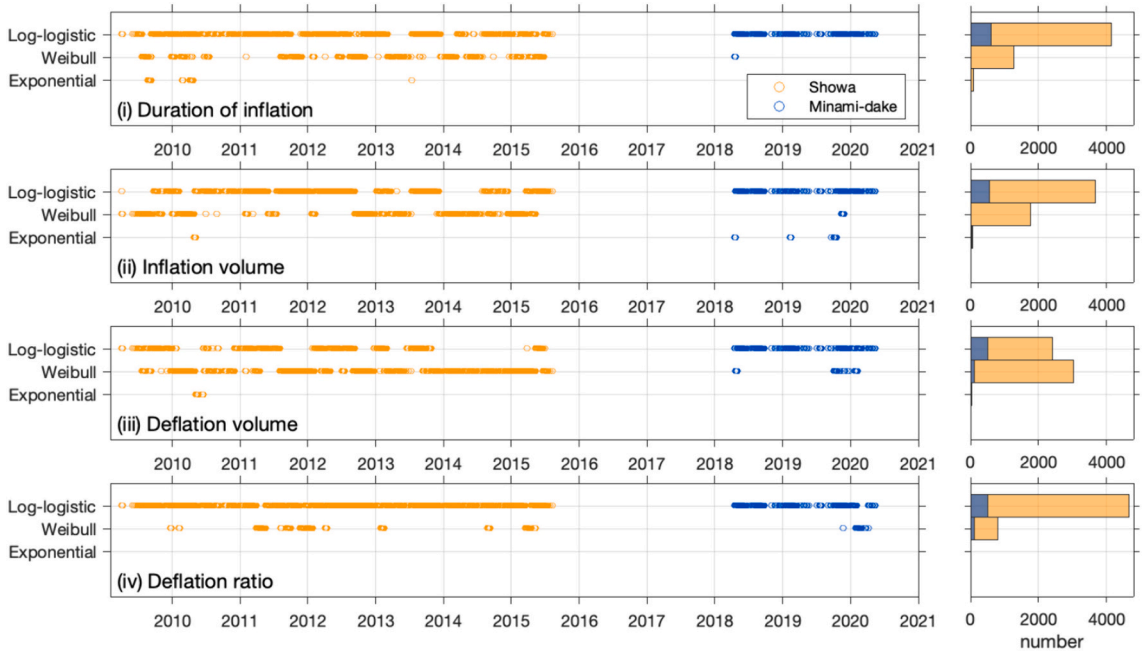


Fig. 5. Temporal evolution in the best model every 100 events.
Each marker represents the best model (i.e., the smallest AIC model) for every 100 events, sliding the analysis window by a single event. The horizontal location of the marker was the time of the 50th event of 100 events. The right panels show the number of analysis windows in which each model was estimated to be the best model. (i) Duration of inflation, (ii) inflation volume, (iii) deflation volume, and (iv) deflation ratio.

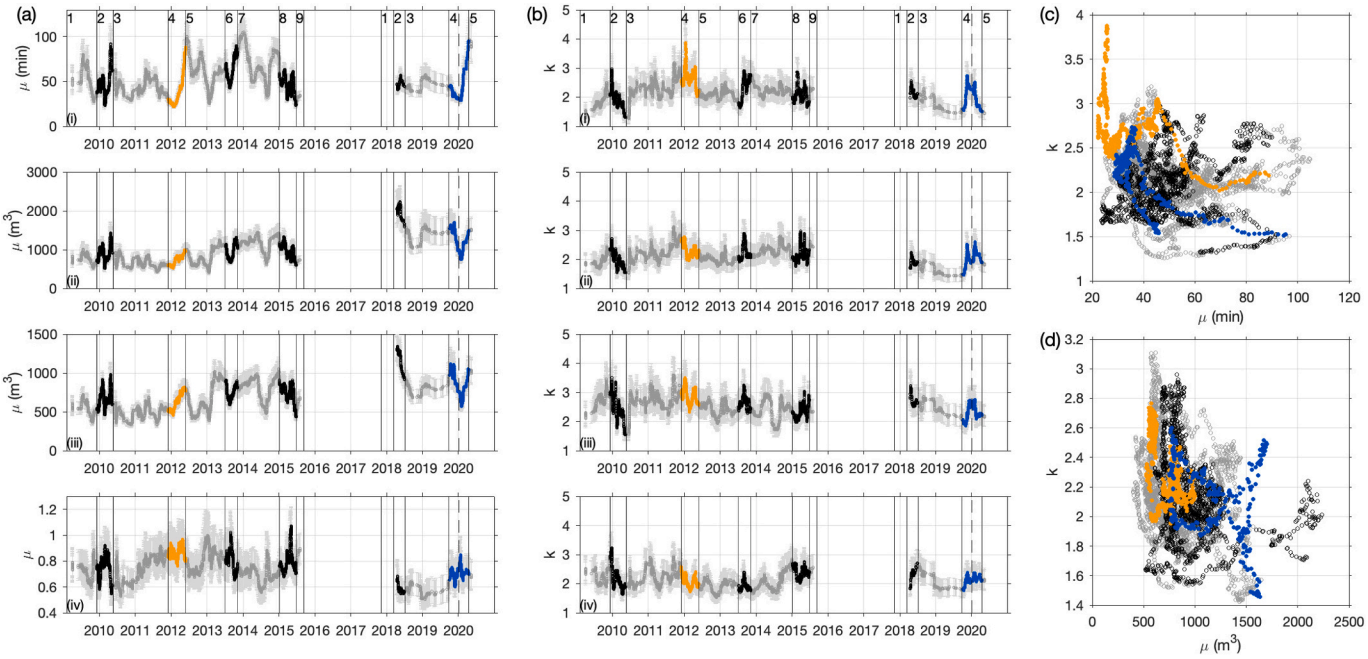


Fig. 6. Temporal variation in the model parameters of the log-logistic distribution.
(a) The scale parameter μ and (b) the shape parameter k estimated from every 100 events. Gray error bars indicate 95 % confidence intervals. Panels (i), (ii), (iii), and (iv) represent the duration of inflation, inflation volume, deflation volume, and deflation ratio, respectively. Orange and blue symbols represent the model parameters estimated from every 100 events in Episode 4 of Showa and Minami-dake craters. Black symbols are those of the even-numbered episodes and gray ones are those of the other periods. (c), (d) Correlation of the model parameters of the log-logistic distribution, μ and k. Orange, blue, black, and gray symbols represent model parameters estimated for every 100 events occurring during the same period as shown in (a) and (b). (c) Duration of inflation, and (d) inflation volume. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

The temporal changes in the model parameters of the log-logistic distribution reflect changes in volcanic activity. Iguchi et al. (2019) reported that the GPS baseline length extended significantly three times during the active period of Showa crater (October 2009–May 2010, October 2011–March 2012, and January 2015–June 2015). These were interpreted as an inflation of the main magma chamber located 10 km deep and a magma intrusion into the region beneath volcanic edifice (1–4 km deep) from the main magma chamber. The volume of intruded

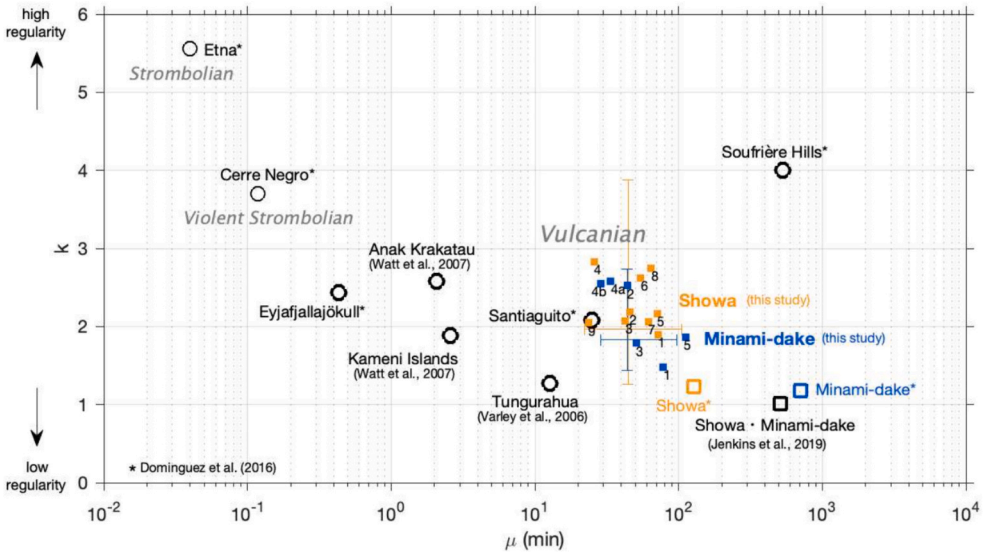


Fig. 7. Comparison of model parameters of the log-logistic distribution (μ and k).
Solid squares represent the model parameters estimated from the sub-datasets of the duration of inflation (orange: Showa crater, blue: Minami-dake crater). The numbers next to the markers indicate the dataset numbers. The error bars show the range from the minimum to maximum values when the parameters were estimated every 100 events (Fig. 6a(i) and b(i)). Open squares represent the model parameters of previous studies dealing with the repose intervals of Vulcanian explosions at Sakurajima volcano. The parameters indicated by the black square were estimated using the published catalog of Jenkins et al. (2019). Open circles with thick lines represent model parameters estimated from the repose intervals of Vulcanian explosions at other volcanoes. The open circles with thin lines are model parameters estimated from the repose intervals of Strombolian and Violent Strombolian explosions. References have been added near the markers in the figure. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

magma from the deep chamber was estimated as $\sim 2 \times 10^6$ m^3 ($\sim 1 \times 10^4$ m^3/day) in October 2011–March 2012 (Hotta et al., 2016). Iguchi et al. (2022b) also noted a large change in areal strain from September 2019 to February 2020 during the active period of Minami-dake crater, indicating a magma intrusion beneath the crater (2–6 km deep). These intrusion periods correspond to episodes 2, 4, and 8 of Showa crater and episode 4 of Minami-dake crater defined in this study. Considering that shape parameter k increased during these episodes, it is suggested that the increase in magma supply from the deep region would shift the phase of volcanic activities, which is one of the factors that increase the regularity of discrete explosions occurring in the shallow regions. Dominguez et al. (2016) indicated that k is related to the complexity of the system. As the system becomes more complex, more factors influence the conduit process, or the interaction among the factors has nonlinear effects. This decreases the regularity of the explosions. Based on this idea, the factors controlling an explosion may be more complicated when the magma supply is low than when it is high. Because a long time is required for magma ascent to a shallow region when the magma supply is low, the effects of degassing and crystallization may be greater.

4.2. Short-term evaluation of explosion probability

The probability of an explosion can be calculated using the probability distributions estimated above. First, we evaluated the timing of the explosion based on the PDF of the duration of inflation using a survival analysis approach. This approach is often applied to long-term forecasting of eruptive cycles using databases of repose intervals (e.g., Bebbington and Lai, 1996; Garcia-Aristizabal et al., 2012). This is the first study to use a database of strain change and real-time strain monitoring.

In the case that t minutes have elapsed since inflation started, the probability that an explosion occurs within Δt minutes is described as Eq. (7).

$$P(t, t+\Delta t) = \frac{\int_t^{t+\Delta t} f(x)dx}{S(t)} = \frac{\int_t^{t+\Delta t} f(x)dx}{1 - F(t)}, \tag{7}$$

$$F(t) = \int_0^t f(x)dx = \frac{1}{1 + (t/\mu)^{-k}}, \tag{8}$$

where $f(t)$ is the PDF, and $F(t)$ is the cumulative distribution function (CDF). In the case of a log-logistic distribution, the CDF is written using a simple analytical expression, as shown in Eq. (8) (Fig. 8a). $S(t)$ is the survival function ($\equiv 1 - F(t)$), which describes the probability that an object still survives at time t (i.e., an explosion has not occurred yet at time t). That is to say, $P(t, t+\Delta t)$ means the conditional probability, which is the probability that an explosion occurs in a specific time window, given that no explosion has occurred until the current time. The integration range of the numerators in Eq. (7) can be modified by considering the time window of interest. The survival function in the denominator of Eq. (7) decreases with time from the start of inflation (Fig. 8a).

Fig. 8b shows the probability of explosions every 10-min window. Here we used the PDF of the log-logistic distribution for the duration of inflation of the Minami-dake crater (Fig. 3b(i)) as an example, considering that the recent activity of Sakurajima volcano has been mainly at the Minami-dake crater. However, it is more appropriate to use the PDF estimated from the dataset of events occurring in the eruptive episode of interest. As time passed, the probability of the explosion changed (Fig. 8b). At the beginning of inflation, an explosion is expected to occur with a typical length of the duration of inflation which is a mode of the PDF. If inflation continues, an explosion is expected to immediately occur.

In addition, the deflation volume related to the eruptive mass can be predicted using real-time strain monitoring. The inflation volume was predicted from the extrapolated strain change, assuming that the inflation rate was sustained (Fig. 9). We assumed several cases of deflation volume: 1/5, 1/2, 1, 2, and 5 times of inflation volume. The probability of each case was obtained from the PDF of the deflation ratio (Fig. 3b (iv), Supplementary material 7). Thus, by multiplying the probabilities of these cases by the probability of explosions in each time window, the probability of explosions was described as a matrix consisting of the timing and the deflation volume (Fig. 9). Since the inflation rate

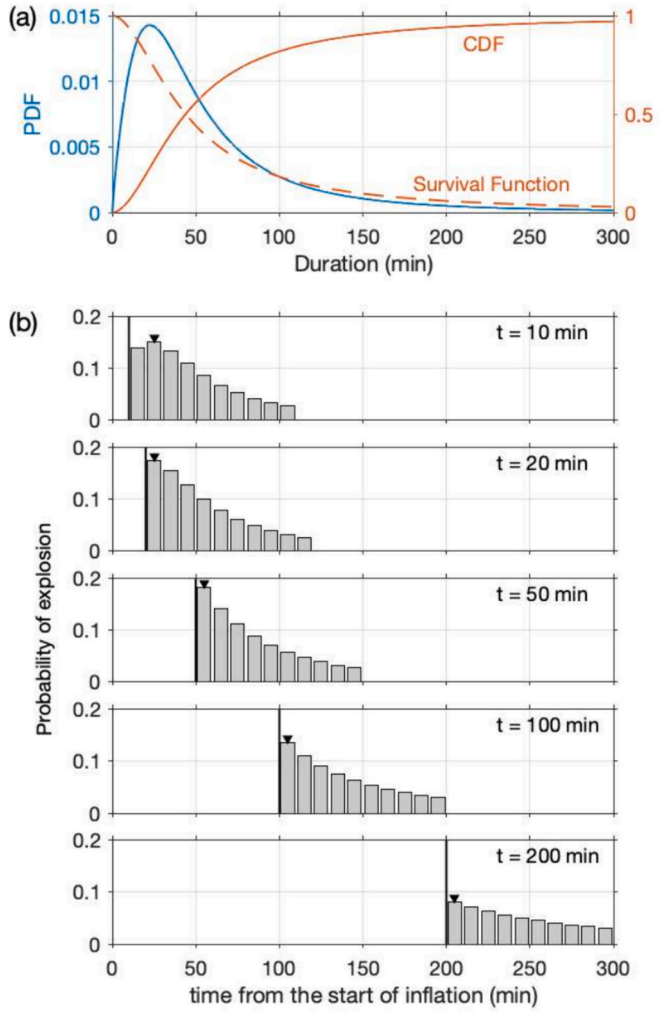


Fig. 8. Evaluation of the timing of explosions.
(a) Probability density function (PDF) of the log-logistic distribution estimated from events that occurred at Minami-dake crater (left axis; same as Fig. 3b(i)). Solid and dashed lines of the right axis are the cumulative distribution function (CDF) and survival function calculated from PDF, respectively.
(b) Probabilities of explosions for every 10-min window (up to 100 min later). Black vertical lines indicate the current time *t*. Black triangles indicate the time windows with the highest probability at each time *t*.

changed as time passed, the extrapolated strain was sequentially updated. Therefore, the deflation volume forecast was updated with time as well. As shown in Fig. 9 (the explosion on September 23, 2022), the explosion occurred in a scenario with a relatively high probability.

Unlike the use of repose intervals, this method can evaluate the size of explosions based on monitored data. Thus, this approach could be helpful for other volcanoes in which cyclic ground deformation has been observed during eruptions, such as the Soufrière Hills (Voight et al., 1998), St. Helens (Anderson et al., 2010), and Etna (Currenti and Bonaccorso, 2019). For Sakurajima volcano, the deflation volume can be empirically converted into ash weight and/or ash emission rate (Iguchi, 2016). The ash emission rate can also be empirically converted into the maximum plume height, which can be used as an input for numerical simulations of ashfall (Iguchi et al., 2022a). Therefore, our estimated deflation volume can contribute to probabilistic forecasting of the area or amount of ashfall (Jenkins et al., 2012; Bear-Crozier et al., 2016). Since the purpose of this research is a statistical survey of ground deformation, we only proposed the method to forecast here, but our next step will be to verify this method.

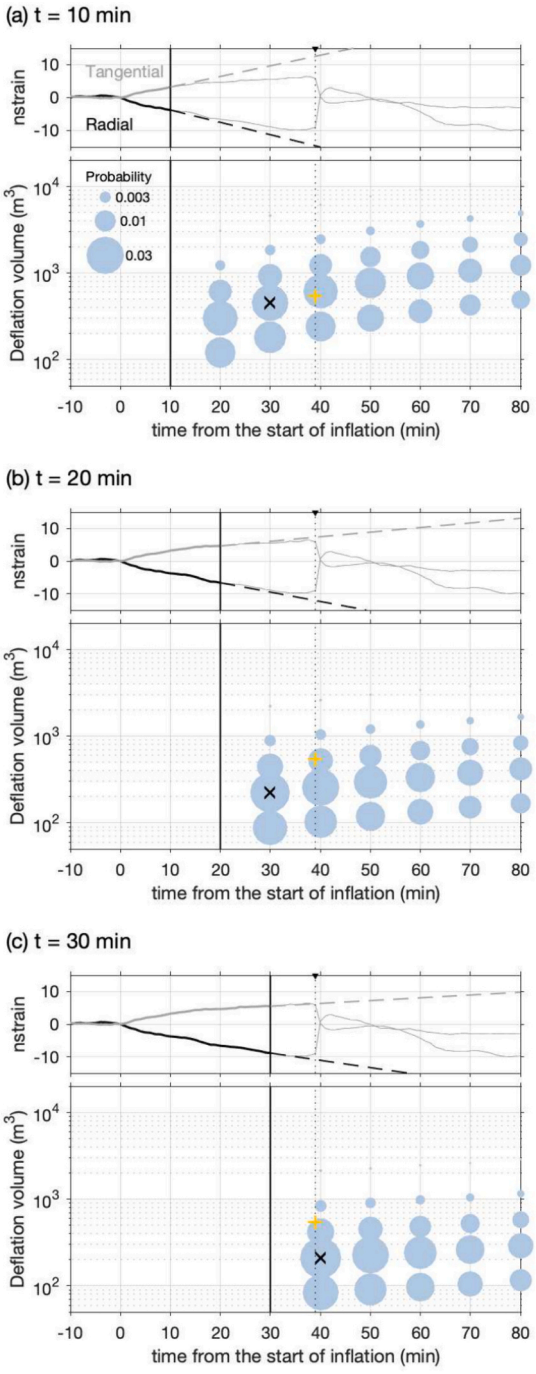


Fig. 9. Evaluation of the timing and size of explosions.
The upper panel shows the strain records of the explosion that occurred at 23:05 JST (UTC + 9) on September 23, 2022. The black vertical line represents the current time t, and the black triangle represents the time of the explosion. The strain data recorded up to the current time are shown on the left side of the vertical line (black: radial strain; gray: tangential strain). On the right side of the line, the dashed lines represent the extrapolated strain data, considering the average inflation rate in the last 10 min. Thin lines are actual strain data. The lower panel shows the probability of explosions at the expected time (horizontal axis) and deflation volume (vertical axis). The bubble size reflects the probability. Black crosses indicate the most likely scenarios. The yellow crosses indicate the actual timing and size of the explosion. (a) Current time $t = 10$ min from the start of inflation. (b) $t = 20$ min, (c) $t = 30$ min. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

4.3. Considerations for the practical forecast of explosions

For a more practical forecast, we must consider following four cases:

- **[A]** Precursory inflation is observed, and **[B]** an explosion occurs.
- **[A]** Precursory inflation is observed, and **[b]** an explosion does not occur (or a small eruption occurs).
- **[a]** No clear precursory inflation is observed, and **[B]** an explosion occurs.
- **[a]** No clear precursory inflation is observed, and **[b]** an explosion does not occur.

This study dealt with only the first case, in which precursory inflation was observed, and an explosion occurred (case [AB]). We need to consider other cases to rigorously assess explosion probability.

However, for Sakurajima, when no clear precursory inflation was observed (cases [aB] and [ab]), explosions were less likely to occur. Explosions without clear precursory inflation account for less than 10 % of all listed explosions. In addition, the explosions in these cases were relatively small. Even if we ignore such cases, they do not pose a large impact on the assessment of hazards. Therefore, it is considered relatively safe in the absence of precursory inflation. The ability to show relatively safe periods is a major advantage of our method over the use of repose intervals. The probability of explosions does not increase if time has passed since the last eruption but no inflation has begun.

The case [Ab] where precursory inflation is observed, but an explosion does not occur (or a small eruption occurs), sometimes exists in the Sakurajima volcano. This study focuses on "Explosions" that generate large infrasound signals (>10 Pa). Other activities at Sakurajima volcano include "Eruptions" and "NEDs (Non-Eruptive Deflations)." Eruptions were defined as non-explosive events which have small infrasound signals (<10 Pa) with visible plumes (>1000-m height). NEDs are deflation events accounted as neither "Explosions" nor "Eruptions" and usually provide only a slight risk of ash hazards because they mainly emit gas (Iguchi et al., 2022b). Note that "Explosions" and "Eruptions" with capital letters are the conventional classifications of the Vulcanian activities at Sakurajima volcano, which are different from the usual meanings of explosions and eruptions (Iguchi et al., 2022b). The strain records for the three types of activities were so similar that it was difficult to distinguish them before the events (Figs. 10a–d). Therefore, an important factor in the forecast is the proportion of hazardous events (i.e., Explosions) and non-hazardous events (i.e., NEDs) when ground deformation is observed. It may be possible to empirically evaluate the probability of each type using the duration of inflation. For example, in February 2020, when NEDs frequently occurred, the durations of inflation of NED tend to be shorter than those of "Explosions" and "Eruptions" (Figs. 10e and f). Therefore, it may be necessary to evaluate the probability of NEDs, particularly if the duration of inflation is short.

Our proposed method can be useful when we evaluate a case [AB] in which precursory inflation is observed, and an explosion occurs. However, some improvements are required to obtain more realistic forecasts. Although the log-logistic distribution explains the data distribution well, it is difficult to specifically evaluate the timing of an explosion when inflation persists for a long period. For Sakurajima volcano, long-period inflation is usually followed by a large explosion (e.g., Hotta and Iguchi, 2021), which increases the need for a more accurate forecast. Nevertheless, as the inflation continues, the distribution of the probability of explosion timing flattens, and the maximum value decreases as shown in Fig. 8b. This issue is due to the characteristics of the log-logistic distribution; therefore, it is impossible to solve as long as a log-logistic distribution is used. Garcia-Aristizabal et al. (2012) suggested using the Brownian passage-time distribution (inverse Gaussian distribution) as an alternative. It may also be helpful to focus on the inflation rate of ground deformation. It is known that pre-eruptive inflation at Sakurajima volcano does not continue monotonically but rather slightly stagnates or contracts just before an explosion (Iguchi et al., 2013). By detecting such a stagnation in the inflation rate, it is possible to more accurately forecast the timing of explosions.

The inflation rate of the ground deformation is strongly related to the timing of the explosions. Iguchi et al. (2019) proposed that the inflation rate (i.e., the magma intrusion rate) prior to eruption is a key parameter for controlling the eruptive style of the Sakurajima volcano. We divided all events of the Minami-dake crater into three groups depending on the inflation rate: < 10 m^3/min, 10–20 m^3/min, and > 20 m^3/min. Explosions with higher inflation rates tended to have shorter inflation durations (Fig. 11a). In addition, the number of the high-inflation-rate events increased during even-numbered episodes. All four variables (duration of inflation, inflation volume, deflation volume, and deflation ratio) in these three groups followed a log-logistic distribution (Table 3;

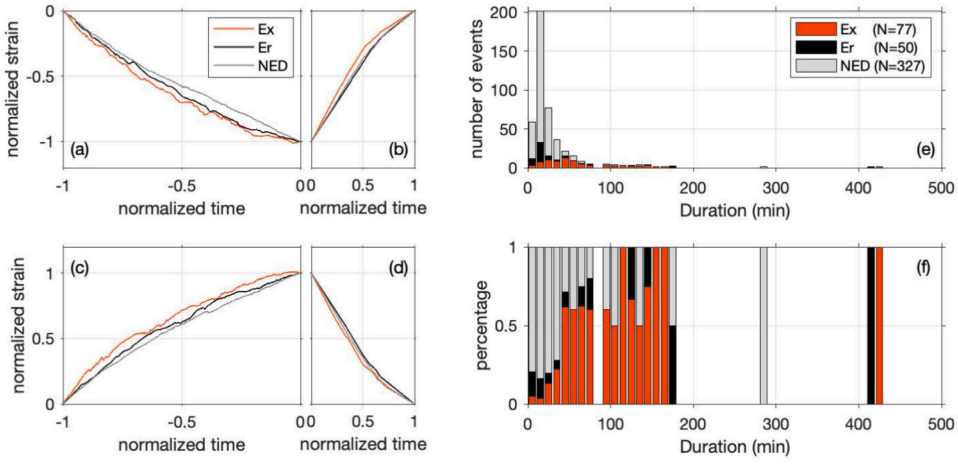


Fig. 10. Strain data and the duration of inflation depending on the event types.
(a)–(d) Normalized strain data of events that occurred in February 2020. The horizontal axis shows normalized time, where time − 1 indicates the start of inflation, time 0 is the event occurrence, and time 1 is the end of deflation. The vertical axis shows the normalized strain amplitude adjusted to make the amplitude at the start of deformation zero and to make the amplitude at the start of deflation one. Red, black, and gray lines are the stacked strain changes of Explosions, Eruptions, and NEDs, respectively. The stacked strain was calculated by taking the median of the normalized amplitudes of all events at each normalized time. (a) Radial strain in the inflation phase. (b) Radial strain in the deflation phase. (c) Tangential strain in the inflation phase. (d) Tangential strain in the deflation phase. (e) and (f) Distribution of the duration of inflation for all types of events in February 2020. (e) The number of events. (f) Percentages of the three types for each bin of (e). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

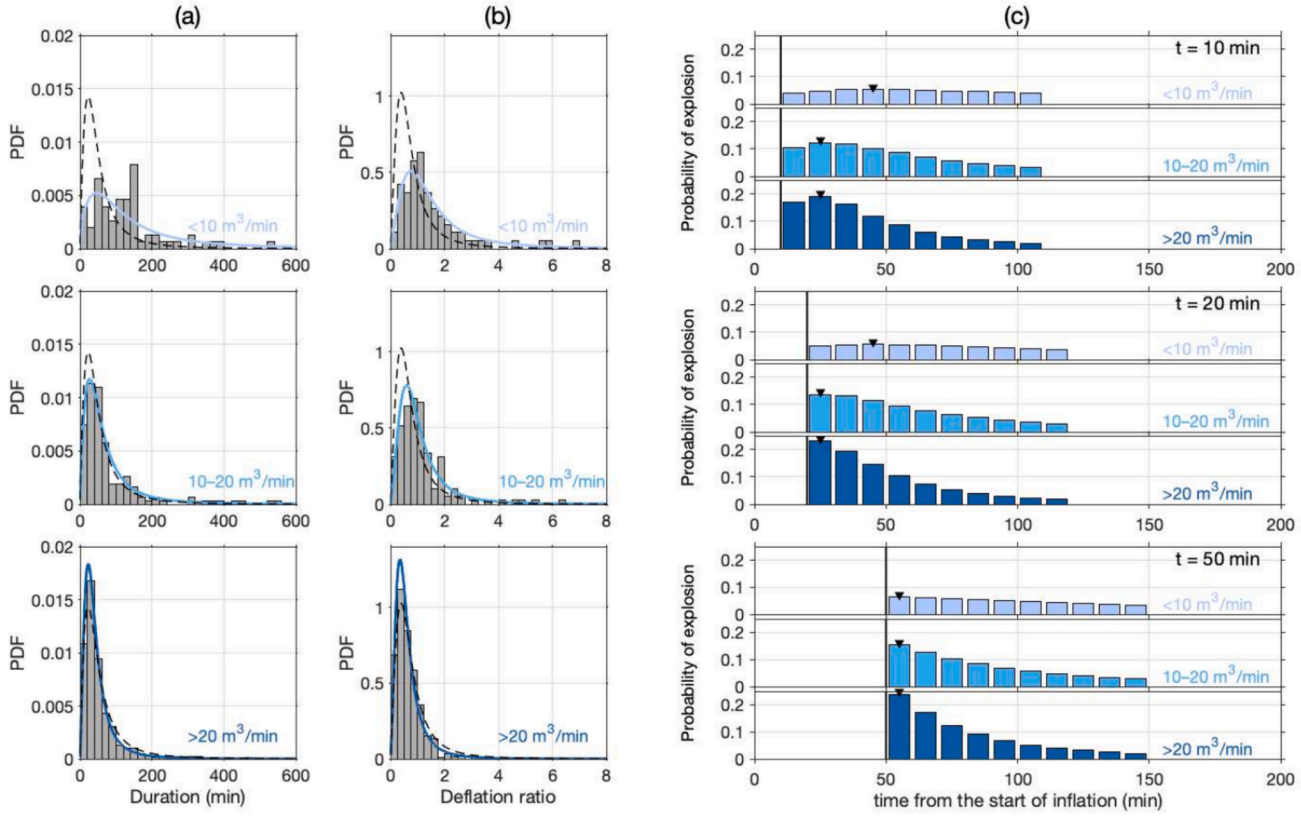


Fig. 11. Grouping based on inflation rate.
(a) The distributions of the duration of inflation depending on inflation rate (<10 m^3/min, 10–20 m^3/min, and > 20 m^3/min) for the explosions that occurred at Minami-dake crater. Dashed lines are the log-logistic distribution estimated using all events (same as Fig. 3b(i)). Colored lines are the log-logistic distributions estimated using events with inflation rate of interest. (b) The distributions of the deflation ratio depending on inflation rate. (c) The probabilities of explosions at t = 10, 20, and 50 min considering inflation rate. The format is the same as in Fig. 8b.

Table 3
Model parameters of the log-logistic distribution depending on inflation rate.

		Duration		Inflation volume		Deflation volume		Deflation ratio	
Inflation rate	Number	μ	k	μ	k	μ	k	μ	k
<10 m^3/min	76	116.7	1.58	752.3	1.76	1002	1.90	1.28	2.04
10–20 m^3/min	155	53.07	1.79	777.2	1.79	731.2	2.39	0.89	2.28
>20 m^3/min	479	36.39	2.13	1749	1.99	922.9	2.39	0.53	2.28
all	710	43.92	1.83	1355	1.76	883.4	2.30	0.65	2.08

Fig. 11a, b). Therefore, we can obtain the probabilities of explosion timing by considering the inflation rate (Fig. 11c). In the case of an explosion with a low inflation rate, the probabilities were broadly distributed, indicating that it may take a long time to reach the explosion. This would lead to a more realistic eruption forecast by choosing a stochastic model based on the inflation rate obtained from real-time data.

5. Conclusions

This study investigated strain records associated with Vulcanian explosions at Sakurajima volcano aiming to conduct a probabilistic forecast of the timing and size of the explosions. Statistical analysis of the duration of inflation, inflation volume, deflation volume, and deflation ratio revealed that a log-logistic distribution can be an appropriate model to explain the distributions of the four variables. The log-logistic distribution is a suitable model to describe a shallow conduit system in which pressurization and depressurization coexist. The balance between these processes controls the timescale and/or magnitude of ground deformation related to Vulcanian explosions. The shape parameter k of the log-logistic distribution quantifies the regularity of the explosions, which may depend on the magma supply from the deep region and the complexity of the system.

The major advantage of using strain data is that they allow the size of the explosion to be evaluated based on the monitoring data. We suggest a methodology to constrain the probability of the timing and size of an imminent explosion using an estimated stochastic renewal model and real-time strain monitoring. This approach provides a probability matrix consisting of the timing and the deflation volume for likely scenarios and updates the probabilities to reflect the strain data obtained in real time. When we conduct numerical simulations of ashfalls considering the obtained probability matrix, it is possible to more quantitatively evaluate the potential ash hazards (e.g., Jenkins et al., 2012). Considering that the distribution of the duration of inflation and the deformation volume varied with the activity level of the volcano, we believe that obtaining an appropriate probability density function according to the activity level and using it to calculate probabilities of the next explosion will lead to a more practical forecast. It would also be necessary to take

into account the ground deformation without explosions to prevent overestimation of explosion probability. In addition, the constraints on the timing of explosions can become more reliable by utilizing the inflation rate.

CRediT authorship contribution statement

Kyoka Ishii: Writing – original draft, Methodology, Investigation, Formal analysis. **Masato Iguchi:** Writing – review & editing, Supervision, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Acknowledgments

We thank the staff of the Sakurajima Volcano Research Center for supporting this study. Two reviewers, M. Ripepe and L. Dominguez, provided many constructive comments to improve our manuscript. This study was conducted under the Integrated Program for Next Generation Volcano Research and Human Resource Development, supported by Ministry of Education, Culture, Sports, Science and Technology (Japan). The ground deformation data were provided by the Osumi Office of River and National Highway, Kyushu Regional Development Bureau, Ministry of Land, Infrastructure, Tranport and Turism (Japan).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jvolgeores.2024.108185.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Trans. Autom. Control 19 (6), 716–723. https://doi.org/10.1109/TAC.1974.1100705.

Anderson, K., Lisowski, M., Segall, P., 2010. Cyclic ground tilt associated with the 2004–2008 eruption of Mount St. Helens. J. Geophys. Res. Solid Earth 115 (B11201). https://doi.org/10.1029/2009JB007102.

Bear-Crozier, A.N., Miller, V., Newey, V., Horspool, N., Weber, R., 2016. Probabilistic Volcanic Ash Hazard Analysis (PVAHA) I: development of the VAPAH tool for emulating multi-scale volcanic ash fall analysis. J. Appl. Volcanol. 5, 3. https://doi.org/10.1186/s13617-016-0043-4.

Bebbington, M.S., Lai, C.D., 1996. Statistical analysis of New Zealand volcanic occurrence data. J. Volcanol. Geotherm. Res. 74 (1–2), 101–110. https://doi.org/10.1016/S0377-0273(96)00050-9.

Bonaccorso, A., Currenti, G., Linde, A., Sacks, S., Sicali, A., 2020. Advances in understanding intrusive, explosive and effusive processes as revealed by the borehole dilatometer network on Mt. Etna Volcano. Front. Earth Sci. 7, 357. https://doi.org/10.3389/feart.2019.00357.

Bonaccorso, A., Carleo, L., Currenti, G., Bilotta, G., Cappello, A., Ganci, G., 2023. A new approach for real-time erupted volume estimation from high-precision strain detection validated by satellite topographic monitoring. Geophys. Res. Lett. 50 (21), e2023GL105424 https://doi.org/10.1029/2023GL105424.

Connor, C.B., Sparks, R., Mason, R., Bonadonna, C., Young, S., 2003. Exploring links between physical and probabilistic models of volcanic eruptions: the Soufrière Hills Volcano, Montserrat. Geophys. Res. Lett. 30 (13), 1701. https://doi.org/10.1029/2003GL017384.

Cox, D.R., Oakes, D., 1984. Analysis of Survival Data. Chapman and Hall/CRC.

Currenti, G., Bonaccorso, A., 2019. Cyclic magma recharge pulses detected by high-precision strainmeter data: the case of 2017 inter-eruptive activity at Etna volcano. Sci. Rep. 9, 7553. https://doi.org/10.1038/s41598-019-44066-w.

Delaney, P., McTigue, D., 1994. Volume of magma accumulation or withdrawal estimated from surface uplift or subsidence, with application to the 1960 collapse of Kilauea Volcano. Bull. Volcanol. 56, 417–424. https://doi.org/10.1007/BF00302823.

Diller, K., Clarke, A., Voight, B., Neri, A., 2006. Mechanisms of conduit plug formation: Implications for vulcanian explosions. Geophys. Res. Lett. 33 (L20302) https://doi.org/10.1029/2006GL027391.

Dominguez, L., Pioli, L., Bonadonna, C., Connor, C.B., Andronico, D., Harris, A., Ripepe, M., 2016. Quantifying unsteadiness and dynamics of pulsatory volcanic activity. Earth Planet. Sci. Lett. 444, 160–168. https://doi.org/10.1016/j.epsl.2016.03.048.

Fries, A., Dominguez, L., Jarvis, P.A., Pistolesi, M., Manrique, N., Aguilar, R., Valdivia, D., Rossi, E., Pollastri, S., Horwell, C.J., et al., 2023. The post-2016 long-lasting Vulcanian activity of Sabancaya volcano (Peru) and associated aeolian remobilisation of volcanic ash. J. Volcanol. Geotherm. Res. 441, 107876 https://doi.org/10.1016/j.jvolgeores.2023.107876.

Garcia-Aristizabal, A., Marzocchi, W., Fujita, E., 2012. A Brownian model for recurrent volcanic eruptions: an application to Miyakejima volcano (Japan). Bull. Volcanol. 74, 545–558. https://doi.org/10.1007/s00445-011-0542-4.

Genco, R., Ripepe, M., 2010. Inflation-deflation cycles revealed by tilt and seismic records at Stromboli volcano. Geophys. Res. Lett. 37 (12) https://doi.org/10.1029/2010GL042925.

Hotta, K., Iguchi, M., 2021. Tilt and strain change during the explosion at Minami-dake, Sakurajima, on November 13, 2017. Earth Planets Space 73, 70. https://doi.org/10.1186/s40623-021-01392-6.

Hotta, K., Iguchi, M., Ohkura, T., Yamamoto, K., 2016. Multiple-pressure-source model for ground inflation during the period of high explosivity at Sakurajima volcano, Japan-Combination analysis of continuous GNSS, tilt and strain data. J. Volcanol. Geotherm. Res. 310, 12–25. https://doi.org/10.1016/j.jvolgeores.2015.11.017.

Iguchi, M., 2016. Method for real-time evaluation of discharge rate of volcanic ash –Case study on intermittent eruptions at the Sakurajima volcano, Japan. J. Disaster Res. 11 (1), 4–14. https://doi.org/10.20965/jdr.2016.p0004.

Iguchi, M., Yakiwara, H., Tameguri, T., Hendrasto, M., Hirabayashi, J., 2008. Mechanism of explosive eruption revealed by geophysical observations at the Sakurajima, Suwanosejima and Semeru volcanoes. J. Volcanol. Geotherm. Res. 178 (1), 1–9. https://doi.org/10.1016/j.jvolgeores.2007.10.010.

Iguchi, M., Tameguri, T., Ohta, Y., Ueki, S., Nakao, S., 2013. Characteristics of volcanic activity at Sakurajima volcano's Showa crater during the period 2006 to 2011. Bull. Volc. Soc. Jpn. 58 (1), 115–135. https://doi.org/10.18940/kazan.58.1_115.

Iguchi, M., Tameguri, T., Hirabayashi, J., Nakamichi, H., 2019. Forecasting volcanic eruption of Sakurajima volcano based on magma intrusion rate. Bull. Volc. Soc. Jpn. 64 (2), 33–51. https://doi.org/10.18940/kazan.64.2_33.

Iguchi, M., Nakamichi, H., Takishita, K., Poulidis, A.P., 2022a. Continuously operable simulator and forecasting the deposition of volcanic ash from prolonged eruptions at Sakurajima volcano, Japan. J. Disaster Res. 17 (5), 805–817. https://doi.org/10.20965/jdr.2022.p0805.

Iguchi, M., Yamada, T., Tameguri, T., 2022b. Sequence of volcanic activity of Sakurajima volcano, Japan, as revealed by non-eruptive deflation. Front. Earth Sci. 10, 727909 https://doi.org/10.3389/feart.2022.727909.

Ishihara, K., 1990. Pressure sources and induced ground deformation associated with explosive eruptions at an andesitic volcano: Sakurajima volcano, Japan. In: Ryan, M. (Ed.), Magma Transport and Storage. John Wileys & Sons, pp. 335–356.

Jenkins, S., Magill, C., McAneney, J., Blong, R., 2012. Regional ash fall hazard I: a probabilistic assessment methodology. Bull. Volcanol. 74, 1699–1712. https://doi.org/10.1007/s00445-012-0627-8.

Jenkins, S.F., Goldstein, H., Bebbington, M.S., Sparks, R.S.J., Koyaguchi, T., 2019. Forecasting explosion repose intervals with a non-parametric Bayesian survival model: application to Sakura-jima volcano, Japan. J. Volcanol. Geotherm. Res. 381, 44–56. https://doi.org/10.1016/j.jvolgeores.2019.04.008.

Kamo, K., Ishihara, K., 1989. A preliminary experiment on automated judgement of the stage of eruptive activity using tiltmeter records at Sakurajima, Japan. In: Latter, J. H. (Ed.), Volcanic Hazards: Assessment and Monitoring. Springer, pp. 585–598.

Le, H.V., Murata, T., Iguchi, M., 2020. Can eruptions be predicted? Short-term prediction of volcanic eruptions via attention-based long short-term memory. Proc. AAAI Conf. Artif. Intell. 34 (08), 13320–13325. https://doi.org/10.1609/aaai.v34i08.7043.

Lyons, J.J., Waite, G.P., Ichihara, M., Lees, J.M., 2012. Tilt prior to explosions and the effect of topography on ultra-long-period seismic records at Fuego volcano, Guatemala. Geophys. Res. Lett. 39 (8), L08305. https://doi.org/10.1029/2012GL051184.

Marzocchi, W., Bebbington, M.S., 2012. Probabilistic eruption forecasting at short and long time scales. Bull. Volcanol. 74, 1777–1805. https://doi.org/10.1007/s00445-012-0633-x.

Mogi, K., 1958. Relations between the eruptions of various volcanoes and the deformations of the ground surfaces around them. Bull. Earthq. Res. Inst. 36, 99–134.

Nakada, S., Shimizu, H., Ohta, K., 1999. Overview of the 1990–1995 eruption at Unzen volcano. J. Volcanol. Geotherm. Res. 89 (1–4), 1–22. https://doi.org/10.1016/S0377-0273(98)00118-8.

Nishimura, T., Iguchi, M., Kawaguchi, R., Surono, Hendrasto M., Rosadi, U., 2012. Inflations prior to vulcanian eruptions and gas bursts detected by tilt observations at semeru volcano, Indonesia. Bull. Volcanol. 74, 903–911. https://doi.org/10.1007/s00445-012-0579-z.

Ripepe, M., Lacanna, G., Pistolesi, M., Silengo, M.C., Aiuppa, A., Laiolo, M., Massimetti, F., Innocenti, L., Della Schiava, M., Bitetto, M., et al., 2021. Ground deformation reveals the scale-invariant conduit dynamics driving explosive basaltic eruptions. Nat. Commun. 12 (1), 1683. https://doi.org/10.1038/s41467-021-21722-2.

Sparks, R.S.J., 2003. Forecasting volcanic eruptions. Earth Planet. Sci. Lett. 210 (1–2), 1–15. https://doi.org/10.1016/S0012-821X(03)00124-9.

Swanson, D., Casadevall, T., Dzurisin, D., Malone, S., Newhall, C., Weaver, C., 1983. Predicting eruptions at Mount St. Helens, June 1980 through December 1982. Science 221 (4618), 1369–1376. https://doi.org/10.1126/science.221.4618.1369.

Taddeucci, J., Edmonds, M., Houghton, B., James, M.R., Vergniolle, S., 2015. Hawaiian and strombolian eruptions. In: The Encyclopedia of Volcanoes. Elsevier, pp. 485–503.

Tamura, Y., Sato, T., Ooe, M., Ishiguro, M., 1991. A procedure for tidal analysis with a Bayesian information criterion. Geophys. J. Int. 104 (3), 507–516. https://doi.org/10.1111/j.1365-246X.1991.tb05697.x.

Udagawa, S., Imanaka, H., Koyaguchi, T., Takayasu, H., 1999. Statistical analysis of occurrence of eruptions at Sakurajima Volcano. Proceedings of the Volcanological Society of Japan, p. 32.

Varley, N., Johnson, J., Ruiz, M., Reyes, G., Martin, K., 2006. Applying statistical analysis to understanding the dynamics of volcanic explosions. In: Mader, H.M., Coles, S.G., Connor, C.B., Connor, L.J. (Eds.), Statistics in Volcanology. Geological Society, pp. 57–76.

Verhulst, P., 1838. Notice sur la loi que la population suit dans son accroisement. Corr. Math. Phys. X 113–121.

Voight, B., Hoblitt, R.P., Clarke, A.B., Lockhart, A.B., Miller, A.D., Lynch, L., McMahon, J., 1998. Remarkable cyclic ground deformation monitored in real-time on Montserrat, and its use in eruption forecasting. Geophys. Res. Lett. 25 (18), 3405–3408. https://doi.org/10.1029/98GL01160.

Watt, S.F.L., Mather, T.A., Pyle, D.M., 2007. Vulcanian explosion cycles: patterns and predictability. Geology 35 (9), 839–842. https://doi.org/10.1130/G23562A.

Weibull, W., 1951. A statistical distribution function of wide applicability. J. Appl. Mech. 18, 293–297. https://doi.org/10.1115/1.4010337.

Yokoo, A., Iguchi, M., Tameguri, T., Yamamoto, K., 2013. Processes prior to outbursts of Vulcanian eruption at Showa crater of Sakurajima volcano. Bull. Volc. Soc. Jpn. 58 (1), 163–181. https://doi.org/10.18940/kazan.58.1_163.