

# Geophysical Research Letters<sup>®</sup>



## RESEARCH LETTER

10.1029/2021GL096715

## Forecasting Eruptions at Poorly Known Volcanoes Using Analogs and Multivariate Renewal Processes

Ting Wang<sup>1</sup> , Mark Bebbington<sup>2</sup> , Shane Cronin<sup>3</sup> , and Joel Carman<sup>1</sup>

<sup>1</sup>Department of Mathematics and Statistics, University of Otago, Dunedin, New Zealand, <sup>2</sup>School of Mathematical and Computational Sciences, and School of Agriculture and Environment, Massey University, Palmerston North, New Zealand, <sup>3</sup>School of Environment, University of Auckland, Auckland, New Zealand

### Key Points:

- We create a statistical tool to identify potential analogue volcanoes based on repose-time characteristics from world-wide datasets
- Analogs selected by this tool significantly reduce forecasting uncertainties compared to using analogs selected by geological similarity
- We apply this tool to forecasting explosive eruptions at Tongariro volcano in New Zealand

### Supporting Information:

Supporting Information may be found in the online version of this article.

### Correspondence to:

T. Wang,  
ting.wang@otago.ac.nz

### Citation:

Wang, T., Bebbington, M., Cronin, S., & Carman, J. (2022). Forecasting eruptions at poorly known volcanoes using analogs and multivariate renewal processes. *Geophysical Research Letters*, 49, e2021GL096715. <https://doi.org/10.1029/2021GL096715>

Received 23 OCT 2021

Accepted 1 JUN 2022

**Abstract** Forecasting future destructive eruptions from re-awakening volcanoes remains a challenge, mainly due to a lack of previous event data. This sparks a search for similar volcanoes to provide additional information, especially those with better compiled and understood event records. However, we show that some of the most obviously geologically comparable volcanoes have differing statistical occurrence patterns. Using such matches produces large forecasting uncertainties. We created a statistical tool to identify and test the compatibility of potential analogue volcanoes based on repose-time characteristics from world-wide datasets. Selecting analogue volcanoes with compatible behavior for factors being forecast, such as repose time, significantly reduces forecasting uncertainties. Applying this tool to Tongariro volcano (NZ), there is a 5% probability for a Volcanic Explosivity Index (VEI)  $\geq 3$  explosive eruption in the next 50 years. Using pre-historic geological records of a smaller available set of analogs, we forecast a 1% probability of a VEI  $\geq 4$  eruption in the next 50 years.

**Plain Language Summary** Probabilistic forecasts of future destructive eruptions from re-awakening volcanoes can inform decision-making to help mitigate volcanic risk. However, a lack of data describing previous events at many volcanoes hinders robust forecasts. Often, the only available forecasts are those based on expert judgment and knowledge of other similar but better known volcanoes. We show that eruption forecasting based only on comparisons of geological criteria produces large uncertainties. We created an alternative statistical tool to identify and test the compatibility of potentially analogous volcanoes based on the time between successive eruptions from individual volcanoes worldwide. Selecting volcanoes with analogous statistical features in terms of the time breaks between successive eruptions significantly reduces the uncertainty of eruption forecasts. We applied this tool to forecasting explosive eruptions at NZ Tongariro volcano using two sets of analogs we identified and found a 5% probability of a Volcanic Explosivity Index (VEI)  $\geq 3$  eruption and 1% probability of a VEI  $\geq 4$  eruption in the next 50 years.

## 1. Introduction

Eruptions from re-awakening volcanoes can threaten life and infrastructure, and disrupt the economy, particularly after long periods of dormancy (e.g., Soufrière Hills, Montserrat, Weidner et al., 2001; Pinatubo, Philippines, Newhall, & Punongbayan, 1996). In the vast majority of cases, there is limited amount of data available for either geological or historical evidence of past volcanism (Loughlin et al., 2015). This makes robust estimation of eruptive hazard from individual volcanoes very challenging and limits the performance of probabilistic forecasting models of any type (e.g., Bebbington, 2014; Marzocchi & Bebbington, 2012; Sheldrake, 2014; Tierz et al., 2019).

An alternative way to forecast future activity of volcanoes with sparse geological records is to create a proxy eruption record, by examining analogue volcanoes, and extrapolating their behavior to the volcano of interest (Marzocchi et al., 2004; Newhall & Pallister, 2015; Rodado et al., 2011; Sheldrake, 2014; Tierz et al., 2019, 2020; Whelley et al., 2015). Analogue volcanoes should ideally have reliable eruption histories. More conservatively, they could also be restricted to those with well-matching magma compositions, tectonic settings and eruption styles to a target volcano (Tierz et al., 2019). Event tree models can incorporate information from analogue volcanoes as expert experience (e.g., Marzocchi et al., 2004; Neri et al., 2008; Newhall & Pallister, 2015), but these results are sensitive to the experts involved and are likely biased toward conservative geological-matching (Flandoli et al., 2011), due to the tendency to emphasize geology as a major criterion, and the general scarcity of well observed volcanological events. The definition of an analogue volcano depends on the goal of the study.

© 2022. The Authors.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs License](https://creativecommons.org/licenses/by-nc-nd/4.0/), which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

If the research interest is in forecasting the occurrence times of future eruptions, then the distribution of repose times is a key factor for identifying analogs.

In this study, we show how hierarchical renewal processes for multiple volcanoes can be used to select analogue volcanoes that display similar statistical features—specifically in eruption onset times. These may or may not share close geophysical and geological parallels to the target volcano. A renewal process is a statistical model for times between events, such as repose times, based on the assumption that they are independent and identically distributed. It has been extensively used to model eruption records (e.g., Bebbington & Lai, 1996b; De la Cruz-Reyna & Carrasco-Nunez, 2002; Marzocchi & Bebbington, 2012; Turner et al., 2008). The process is defined by its hazard function, which describes, at any instant, the expected number of eruptions per unit time. In addition to the geological problem of no two volcanoes being identical, there are serious problems with analogue volcanoes selected too closely based on geophysical and geological constraints if they show widely divergent statistical behavior in their eruption record (Sheldrake, 2014). In order to make robust forecasts using analogue volcanoes, it is necessary to identify groups of volcanoes that share similar properties in their statistical eruption-record features. The most useful analogs are those that have longer, or more complete eruption records than the target volcano. Similarity between multiple renewal processes representing each volcano can be modeled by specifying the same family of hazard functions, but allowing model parameters to vary between volcanoes. Using these tools we can select analogue volcanoes with behavior similar in respect of the quantity we want to forecast (usually, the length of the current repose).

## 2. Tongariro Volcano and Candidate Analogs

Tongariro is an andesitic volcano located in the southern Taupo Volcanic Zone, North Island, New Zealand. It has grown over the past c. 275 kyr as a complex of spatially distinct but overlapping and concurrently active eruptive centers (Hobden et al., 1999). The most notable and recently active vents are Ngauruhoe, Red Crater and TeMaari (Scott & Potter, 2014).

On 6 August 2012,  $a < 0.001 \text{ km}^3$  explosive eruption occurred from Tongariro in New Zealand (Pardo et al., 2014). The long repose since the previous eruption in 1926 was in marked contrast to the earlier frequent eruptions (Gregg, 1960). Past studies on the likelihood of future eruptions of New Zealand volcanoes have included national studies (Bebbington et al., 2018; Hurst & Smith, 2010; Stirling & Wilson, 2002) and those on individual volcanoes with well-described records of past activity, such as Ruapehu and Ngauruhoe (Bebbington & Lai, 1996b), Taranaki (Green et al., 2013; Wang & Bebbington, 2012), and Auckland (Bebbington & Cronin, 2011). Not considering the active parasitic cone Ngauruhoe (which has its own separate eruption dynamics (Moebis et al., 2011)), the eruption record of Tongariro has too few events to make a robust hazard evaluation.

The majority of historical eruptions at Tongariro occurred between 1846 and 1926 (Scott & Potter, 2014). These show a maximum Volcanic Explosivity Index (VEI) of 3. VEI is an ordinal measurement of eruption intensity derived from the available information, including its explosive volume, the eruption-column height, and the eruption duration (Newhall & Self, 1982). The scale of the index is logarithmic for VEIs over 2, with one unit corresponding to an approximate tenfold increase in pyroclastic eruptive volume. In addition to the small historical events, the geological record from the volcano includes several VEI 4 eruptions (Moebis et al., 2011), along with a few far larger VEI 5 eruptions (without historical event equivalents) between 11,000 years BP and 17,000 years BP (Heinrich et al., 2020; Shane et al., 2008; Topping, 1973). The known geological record is biased to recording pyroclastic eruptions, with timing of effusive events poorly constrained in time. For eruptions  $> 17,000$  years BP, the geological record is most complete for the largest explosive events producing the thickest tephra deposits.

The set of analogue volcanoes was selected broadly and empirically by eruption records as “empirical analogs” (Table 1). This considers a set of volcanoes from the Global Volcanism Program (GVP) catalog that may share similar statistical occurrence pattern to Tongariro. The potential analogue volcanoes were selected based on volcanological expertise to satisfy at least two of the three constraints: composite volcanoes, erupting basaltic to andesite compositions and sited in convergent margins. Candidate analogs also require good eruption records and a range of eruption sizes similar to the target (e.g., Tierz et al., 2019). Note that in our statistical formulation, that borrows strength using Bayesian methods, analogs need not have longer records than the target volcano provided that the behavior is similar to that of the target volcano. Table 1 also reports the percentage of “better analogs” that the VOLCANO ANALOGS SEARCH (VOLCANS) method (Tierz et al., 2019) identifies for Tongariro,

**Table 1**  
*Potential Analogue Volcanoes for Tongariro From the GVP Catalog*

Volcano name	Country	VT	MRT	TS	$N_{h,i}$	$N_{3,i}$	Order	$Y$	$Z$	$p$ (%)
Tongariro	NZ	SV, PC	A/BA	CC, SZ	10	6	1			
<b>Asamayama</b>	<b>Japan</b>	<b>C, LD</b>	<b>A/BA, D</b>	<b>CC, SZ</b>	117	21	2	0.081	0.813	1
<b>Avachinsky</b>	<b>Russia</b>	<b>SV, CD, LD</b>	<b>A/BA, B/PB</b>	<b>CC, SZ</b>	19	42	3	0.000	0.239	6
Azumayama	Japan	SV, S, LD, PC	A/BA, B/PB	CC, SZ	9	8	4	0.084	0.040	8
Bezymianny	Russia	SV, CD, LD	A/BA, D	CC, SZ	52	38	5	0.107	0.036	5
<b>Calbuco</b>	<b>Chile</b>	<b>SV, LD</b>	<b>A/BA</b>	<b>CC, SZ</b>	14	30	6	0.012	0.797	16
<b>Chokaisan</b>	<b>Japan</b>	<b>SV, CD, LD</b>	<b>A/BA, B/PB</b>	<b>CC, SZ</b>	11	7	7	0.062	0.559	16
<b>Colima</b>	<b>Mexico</b>	<b>SV, CD, LD</b>	<b>A/BA</b>	<b>CC, SZ</b>	52	46	8	0.002	0.870	20
Fujisan	Japan	SV, PC	B/PB	CC, SZ	15	38	9	0.000	0.088	12
<b>Gamalama</b>	<b>Indonesia</b>	<b>SV, M</b>	<b>A/BA</b>	<b>OC, SZ</b>	66	12	10	0.000	0.784	24
<b>Guagua Pichincha</b>	<b>Ecuador</b>	<b>SV, CD, LD</b>	<b>A/BA, D</b>	<b>CC, SZ</b>	20	18	11	0.023	0.266	14
<b>Kuchinoerabujima</b>	<b>Japan</b>	<b>SV, PC</b>	<b>A/BA</b>	<b>OC, SZ</b>	15	15	12	0.285	0.545	6
<b>Miyakejima</b>	<b>Japan</b>	<b>SV, CD, PC</b>	<b>B/PB, A/BA</b>	<b>OC, SZ</b>	26	24	13	0.011	0.728	13
<b>Nevado del Ruiz</b>	<b>Colombia</b>	<b>SV, CD, LD, PC</b>	<b>A/BA, D</b>	<b>CC, SZ</b>	13	12	14	0.192	0.906	21
<b>Pico de Orizaba</b>	<b>Mexico</b>	<b>SV, CD, LD</b>	<b>A/BA, D</b>	<b>CC, SZ</b>	6	17	15	0.036	0.260	28
Popocatepetl	Mexico	SV, CD, PC, LD	A/BA, D, B/PB	CC, SZ	32	12	16	0.059	0.051	9
<b>Puyehue-Cordón</b>	<b>Chile</b>	<b>SV, CD, FV, PC, LD</b>	<b>A/BA, D, B/PB, R</b>	<b>CC, SZ</b>	11	18	17	0.012	0.385	4
Sheveluch	Russia	SV, CD, LD	A/BA, B/PB	CC, SZ	20	91	18	0.000	0.127	3
<b>Suwanosejima</b>	<b>Japan</b>	<b>SV, PC, FV</b>	<b>A/BA</b>	<b>OC, SZ</b>	23	5	19	0.011	0.780	14
<b>Tokachidake</b>	<b>Japan</b>	<b>SV, LD, PC</b>	<b>A/BA, B/PB, D</b>	<b>CC, SZ</b>	17	12	20	0.021	0.388	5
<b>Tolbachik</b>	<b>Russia</b>	<b>S, CD, PC</b>	<b>B/PB, A/BA</b>	<b>CC, SZ</b>	32	28	21	0.010	0.807	1
<b>Tungurahua</b>	<b>Ecuador</b>	<b>SV</b>	<b>A/BA, D</b>	<b>CC, SZ</b>	15	25	22	0.043	0.476	23
<b>Villarrica</b>	<b>Chile</b>	<b>SV, CD, PC, FV</b>	<b>B/PB, A/BA, R</b>	<b>CC, SZ</b>	74	11	23	0.208	0.483	16

*Note.* Notation: Volcano Type (VT), CD = Caldera, C = Complex, FV = Fissure Vents, LD = Lava Domes, M = Maar. PC = Pyroclastic Cones, S = Shield, SV = Stratovolcano(es); Major Rock Type (MRT), A = Andesite, BA = Basaltic Andesite, B = Basalt, D = Dacite, PB = Picro-Basalt, R = Rhyolite; Tectonic Setting (TS), CC = Continental crust, OC = Oceanic crust, SZ = Subduction zone;  $N_{h,i}$ : number of historical observations;  $N_{3,i}$ : number of confirmed eruptions with minimum VEI 3. The columns for  $Y$  and  $Z$  give the estimated proportion that the MCMC samples of the posterior of  $Y_i$  and  $Z_i$ ;  $i = 2, \dots, 23$  overlap with that of the target volcano (i.e.,  $Y_1$  or  $Z_1$ ) using model M3; and  $p$  is the percentage of better analogs that PyVOLCANS (Tierz et al., 2021) identifies for Tongariro, compared to each of the empirical analogs.

compared to each of the empirical analogs identified based on volcanological expertise. We used the open-source code PyVOLCANS with equal weight for each of the five volcanological criteria: tectonic setting, rock geochemistry, volcano morphology, eruption size, and eruption style (Tierz et al., 2021).

Renewal processes were used to identify a subset of volcanoes from the empirical analogs that we named “statistical analogs” for Tongariro. We calculated forecasts from these statistical analogs for  $VEI \geq 3$  events. This also serves as a test of the method. These statistical analogs were identified using the Holocene records from GVP. In Section 5, we also consider two analogs that have much longer (Quaternary) records than that in the GVP catalog. For the ease of discussion, we call the former statistical analogs with Holocene records and the latter statistical analogs with Quaternary records wherever clarification is needed.

### 3. Empirical Analogos

Recurrences of volcanic eruptions are often modeled as renewal processes if there is no evidence of systematic, or statistical correlation of the intervals between consecutive eruptions. We use a Weibull renewal process with hazard rate function at time  $t$

$$\lambda(t) = \alpha(t - s)^{\alpha-1} / \beta^\alpha, \quad (1)$$

as the base to study analogue volcanoes, with shape parameter  $\alpha > 0$  controlling the degree of clustering/periodicity, scale parameter  $\beta > 0$  controlling the mean time between events, and  $s < t$  being the occurrence time of the most recent event (Bebbington & Lai, 1996a). This has been shown to produce consistent forecasts for many volcanoes (e.g., De la Cruz-Reyna & Carrasco-Nunez, 2002; Marzocchi & Bebbington, 2012).

In a set of  $M$  volcanoes, the scale (or shape) parameter of the hazard rate function in Equation 1 is multiplied by a random variable  $Y_i$  (or  $Z_i$ ) ( $i = 1, \dots, M$ ) to capture the similarities between candidate volcanoes. This allows us to use volcanoes with similar clustering properties but different activity levels as potential analogs. In particular, the first model we consider, M1, allows some variation in the activity level between volcanoes, and has a hazard rate function

$$\lambda_{1,i}(t) = \frac{Y_i \alpha (t - s)^{\alpha - 1}}{\beta^\alpha}, \quad (2)$$

for the  $i$ th volcano, where  $i = 1, \dots, M$ , and  $Y_i$  is a random variable taking positive values in the real line. If  $Y_i$  do not have statistically different distributions for different  $i$ , then the set of volcanoes are considered as being good analogs in terms of activity levels. These types of models have been used to study survival data in medicine (Aalen & Husebye, 1991; Hougaard, 1986).

The second model M2 allows some variation in clustering behavior between volcanoes, and has a hazard rate function

$$\lambda_{2,i}(t) = \frac{\alpha (t - s)^{Z_i \alpha - 1}}{\beta^\alpha} \quad (3)$$

where statistically insignificant difference in  $Z_i$ 's will indicate that the various individual volcanoes share similar shapes and thus can be considered good analogs in terms of clustering behavior. The third model M3 with a hazard rate function

$$\lambda_{3,i}(t) = \frac{Y_i \alpha (t - s)^{Z_i \alpha - 1}}{\beta^\alpha}, \quad (4)$$

uses two random variables  $Y_i$  and  $Z_i$  taking positive values in the real line to distinguish both differences in the activity level and clustering behavior between volcanoes. If  $Z_1 = Z_2 = \dots = Z_M$  in distribution, then we consider all such candidate volcanoes form a set of analogue volcanoes. Note that it is not necessary to require  $Y_1 = Y_2 = \dots = Y_M$  in distribution. Significantly different shape parameters could imply that some volcanoes have periodic occurrence patterns while others have clustered events. These different behaviors produce very different hazard forecasts. However, significantly different scale parameters only imply variation in the activity level among the set of volcanoes. For example, analogue volcanoes with much more data than the target volcano within the same time period would have different activity levels. The activity levels can be rescaled by using  $Y_i$  to make these volcanoes useful analogs. When very different timescales are involved in the identification of common clustering/periodicity, they indicate a correspondingly different tempo in the “magmatic” renewal processes at those volcanoes. The statistical model effectively maps these volcanoes onto each other for forecasting purposes through a time re-scaling.

These models also produce Weibull hazard rate functions, but the shapes of the hazard rate functions (increasing/decreasing) depend on the combinations of  $Z$ ,  $Y$ , and the parameters  $\alpha$  and  $\beta$ . Using analogue volcanoes can produce more robust estimates of  $\alpha$  and  $\beta$  in the intensity in Equation 1. Our aim is to use the three models defined above to analyze the eruption times from  $M$  individual volcanoes that share similar geological properties and check if they can be considered as statistical analogs (volcanoes sharing similar statistical features in their occurrence behavior). To do this, we assume that  $Y_i$  and  $Z_i$  follow prior distributions such that they all have mean 1 to avoid identifiability problems, that is,

$$Y_i \sim \text{Gamma}(1/\sigma_Y^2, 1/\sigma_Y^2), \quad Z_i \sim \text{Gamma}(1/\sigma_Z^2, 1/\sigma_Z^2), \quad (5)$$

where  $\sigma_Y$  and  $\sigma_Z$  are the standard deviations of  $Y_i$  and  $Z_i$ , respectively. If the posteriors of  $Z_i$  (or  $Y_i$ ,  $i = 1, \dots, M$ ) overlap significantly and the posterior of  $\sigma_Z$  (or  $\sigma_Y$ ) has mass close to 0, then we can consider these volcanoes as

analogs. We now have a hierarchical model, and we will use Bayesian methods to estimate the parameters. More details about parameter estimation using the Markov Chain Monte Carlo (MCMC) algorithm are in Appendix B.

Models M1, M2, and M3 are used to test for the best analogs using all confirmed  $VEI \geq 3$  eruptions of the set of empirical analogs selected from the GVP catalog (Table 1). This is limited by the GVP containing only the Holocene (last ~12,000 years) record (which is also the case for Tongariro). We assume that the records of  $VEI \geq 3$  eruptions are more complete than the records with smaller eruptions (e.g., Mead & Magill, 2014) and an eruption without a specified VEI is assumed to be at least VEI 3. This is in accordance with Newhall and Self (1982) where explosive eruptions are a minimum of VEI 2, and all VEI 1–4 before a region-specific date are upgraded by one VEI class. Thus inference using the records of  $VEI \geq 3$  observations of the empirical analogs (up to September 2018) provides robust statistical analogs in the Holocene eruption record that are most useful in forecasting eruptions of moderate sizes. Duplicate ages in the eruption record (i.e., multiple events on the same date) are removed. The plot for VEI versus time for each volcano is shown in Figure 1a–1w, and the histogram of VEIs for each volcano is in the Supporting Information S1.

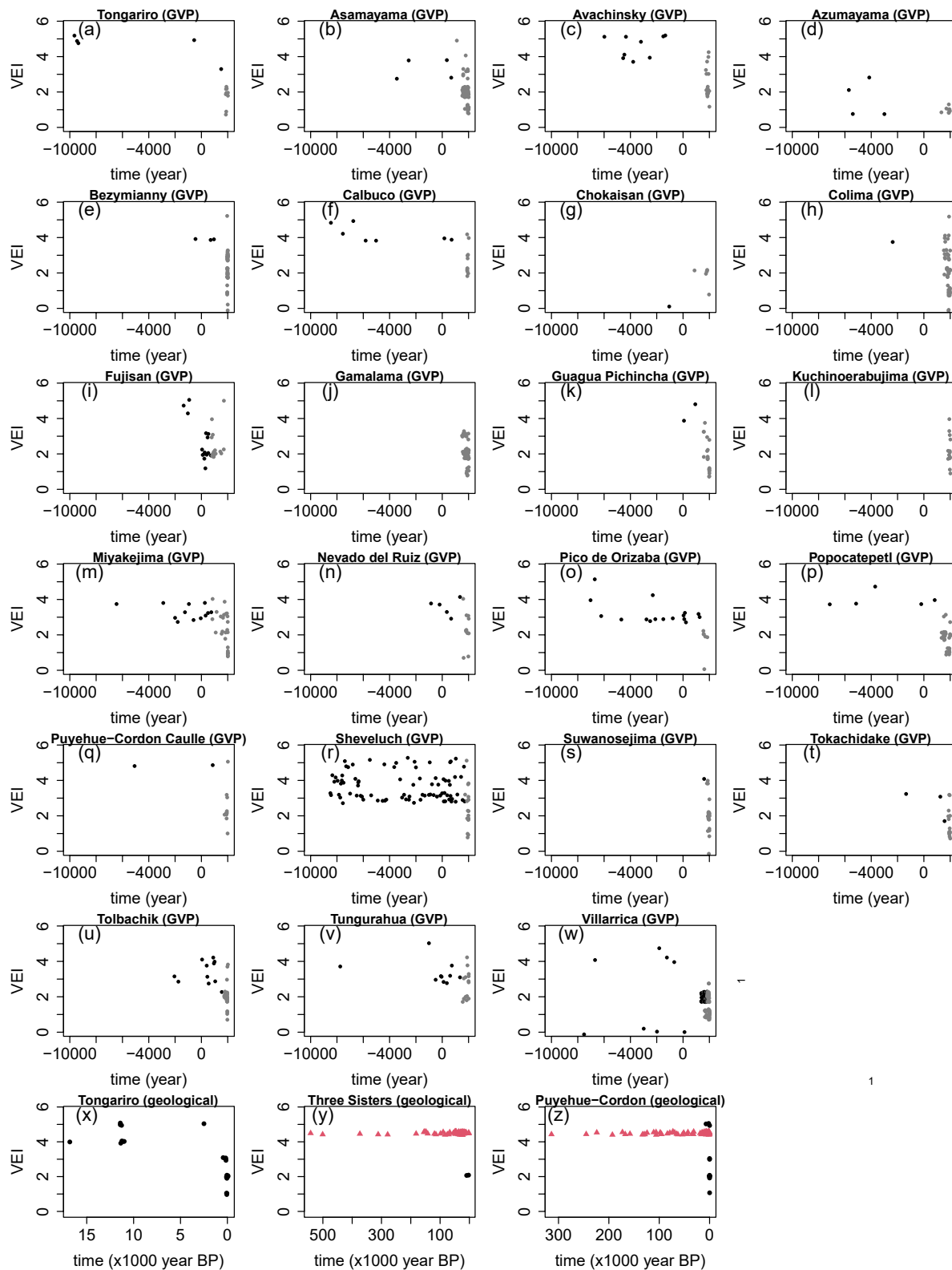
The posterior densities of  $Z_i$  from some volcanoes do not overlap well with that from Tongariro (Figure 2c, plots for models M1 and M2 in Supporting Information S1). This is confirmed by the posterior densities of  $\sigma_z$ , the standard deviation of  $Z_i$  that describe the overall variation in  $Z_i$ , having mass away from 0 (with 95% highest posterior density [HPD] interval [0.200, 0.471] not containing 0) (Figure 2d). An HPD interval gives the shortest interval that contains the most probable values of the variable. We mainly focus on model M3 because it performs the best (see further discussion in Section 4). The results suggest that the variation in  $Z_i$  is significantly greater than 0, which indicates significant differences in the shape of the hazard function among at least some of the candidate volcanoes in Table 1. Note that the variation in  $Y_i$  is also significantly greater than 0 with 95% HPD interval (0.468, 0.968) away from 0 (Figures 2a and 2b). As discussed above, this variation is allowed as  $Y_i$  serves as a scaling factor for the activity levels between volcanoes.

To refine differences in shape or scale of hazard functions from Tongariro, we calculated the proportion that the MCMC samples of the posterior of  $Z_i$  (or  $Y_i$ ),  $i = 2, \dots, 23$  overlap with Tongariro values (i.e.,  $Z_i$  or  $Y_i$ ). For each volcano  $i$ , if this proportion for the shape parameter in model M3 is greater than 0.2, then the  $i$ th volcano is a good analogue (Table 1). We use the value 0.2 to be deliberately conservative and to ensure that the variation in  $Z_i$  for the set of analogue volcanoes is not significantly different from 0, that is, to ensure that there isn't any significant difference in the shape of the hazard functions for all the chosen analogue volcanoes. We then examined the residuals, which are the differences between the observations and model prediction. Residuals reveal whether there is systematic, that is, non-random, behavior that is not explained by the model, and help to understand how well the model fits all the analogs by this cut-off value. Models with poorly behaved residuals fail to capture the essential features of the data. The estimated proportions in Table 1 suggest that 17 volcanoes, Asamayama, Avachinsky, Calbuco, Chokaisan, Colima, Gamalama, Guagua Pichincha, Kuchinoerabujima, Miyakejima, Nevado del Ruiz, Pico de Orizaba, Puyehue-Cordón Caulle, Suwanosejima, Tokachidake, Tolbachik, Tungurahua, Villarrica, are suitable analogs of Tongariro. We call this set of records “statistical analogs with Holocene records.”

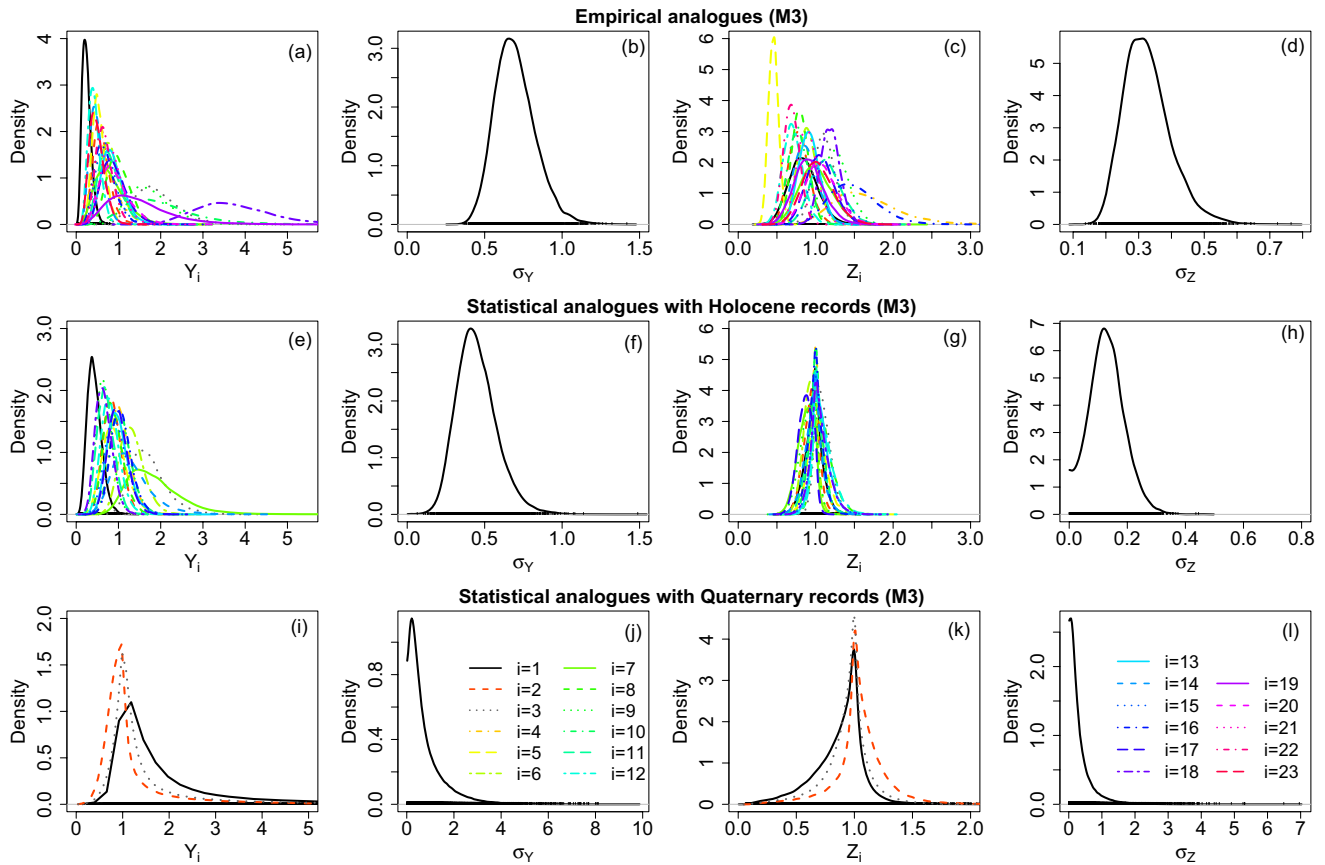
This set of statistical analogs are all classified as “stratovolcano(es)” in GVP, except for Asamayama and Tolbachik, located at convergent margins and erupt mainly basalt to andesite compositions. Tongariro has 6  $VEI \geq 3$  eruptions. These statistical analogs have between 5 and 46  $VEI \geq 3$  eruptions, with Colima having the largest number, 46  $VEI \geq 3$  eruptions. The posterior distributions of  $Z_i$  from fitting model M3 to Tongariro and the 17 statistical analogue volcanoes largely overlap with each other (Figure 2g). The posterior for  $\sigma_z$  that describes the overall variation in  $Z_i$  now has mass close to 0, with 95% HPD interval (0, 0.233) including 0 (Figure 2h). It suggests that there is no significant difference in the shape of the hazard functions of the 17 statistical analogue volcanoes to that of Tongariro. The variation in  $Y_i$  is still significantly greater than 0 with 95% HPD interval (0.218, 0.720) away from 0 (Figures 2e and 2f), which does not affect our choice of the statistical analogs as  $Y_i$  serves as a scaling factor for the activity levels between volcanoes.

#### 4. Forecast Using Statistical Analogs With Holocene Records

The GVP catalog reports the last eruption of  $\geq VEI 3$  from Tongariro at CE1500. In order to be consistent, the GVP record is used for Tongariro in this section for matching and forecasting. Considering eruptions with  $VEI \geq 3$  from the statistical analogs in the Holocene eruption record and assuming a minimum VEI 3 for non-specified

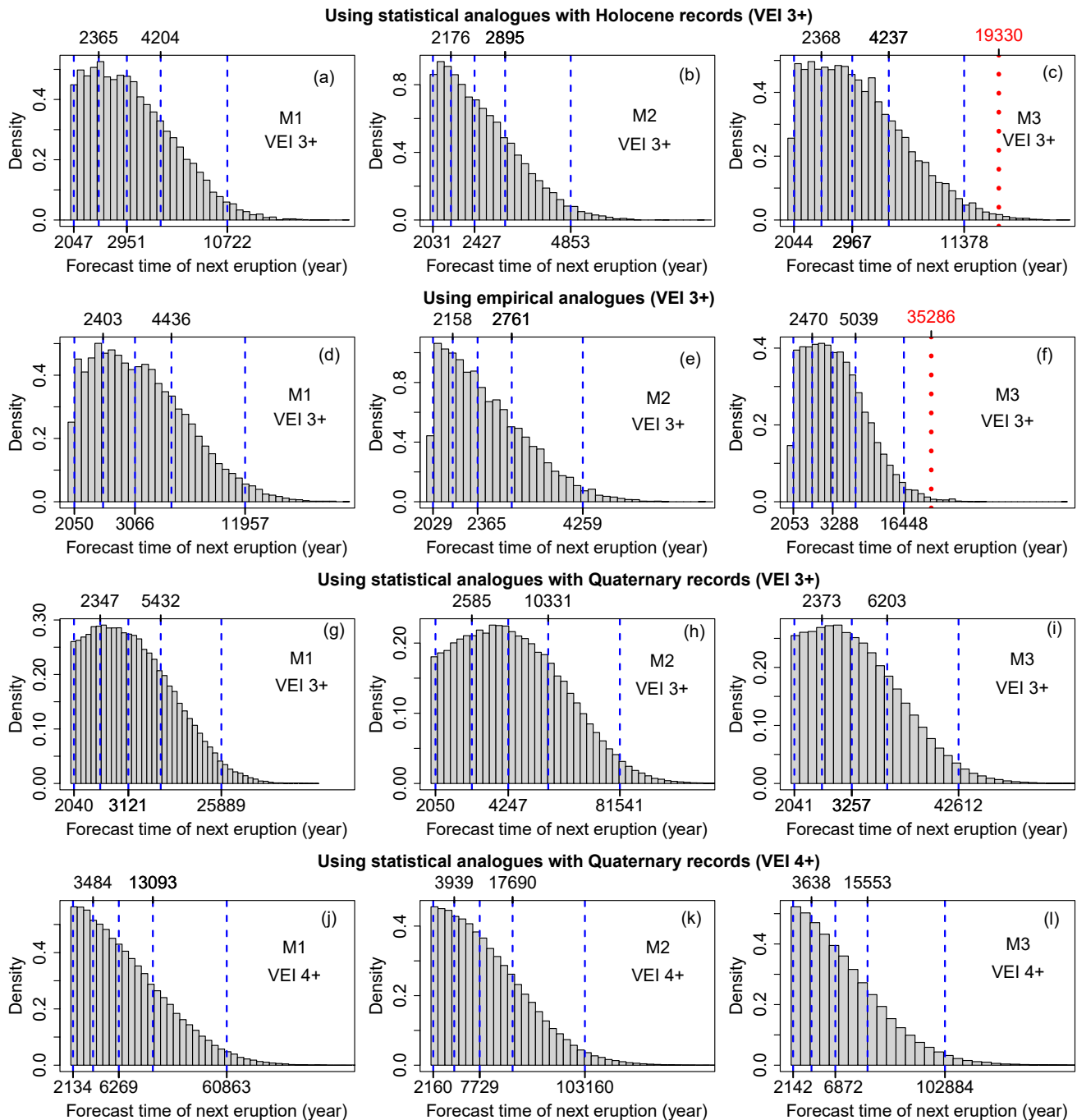


**Figure 1.** Volcanic Explosivity Index (VEI) versus time plot. Random noise is added to each VEI value to make all observations visible in the plot. Those without any VEI values given in the record (in red), were arbitrarily assigned a VEI 4.5 to show them in the plot. Historical observations as defined by Global Volcanism Program are shown in gray.



**Figure 2.** The posterior distributions of  $Y_i$ ,  $\sigma_Y$ ,  $Z_i$ , and  $\sigma_Z$ . (a–d) Model M3 fitted to the Global Volcanism Program (GVP) records of Volcanic Explosivity Index (VEI)  $\geq 3$  eruptions from Tongariro and the 22 empirical analogs. The order of the volcanoes in our data analysis and in (a–d) is: (1) Tongariro, (2) Asamayama, (3) Avachinsky, (4) Azumayama, (5) Bezymianny, (6) Calbuco, (7) Chokaisan, (8) Colima, (9) Fujisan, (10) Gamalama, (11) Guagua Pichincha, (12) Kuchinoerabujima, (13) Miyakejima, (14) Nevado del Ruiz, (15) Pico de Orizaba, (16) Popocatepetl, (17) Puyehue-Cordón Caulle, (18) Sheveluch, (19) Suwanosejima, (20) Tokachidake, (21) Tolbachik, (22) Tungurahua, (23) Villarrica. (e–h) Model M3 fitted to the GVP records of VEI  $\geq 3$  eruptions from Tongariro and the 17 statistical analogs with Holocene records. The order of the volcanoes in this data analysis and in (e–h) is: (1) Tongariro, (2) Asamayama, (3) Avachinsky, (4) Calbuco, (5) Chokaisan, (6) Colima, (7) Gamalama, (8) Guagua Pichincha, (9) Kuchinoerabujima, (10) Miyakejima, (11) Nevado del Ruiz, (12) Pico de Orizaba, (13) Puyehue-Cordón Caulle, (14) Suwanosejima, (15) Tokachidake, (16) Tolbachik, (17) Tungurahua, (18) Villarrica. (i–l) Model M3 fitted to Tongariro, Three Sisters and Puyehue-Cordón Caulle (the two statistical analogs with Quaternary records) eruptions with minimum VEI 3.

eruptions, we forecast the next interevent time (i.e., the time from CE1500 to the next eruption) using the MCMC algorithm by conditioning on quiescence since 1500 until the censored time September 2018. The credible intervals are calculated for the interevent time to the next eruption (recognizing no  $\geq$ VEI 3 events between 1500 and September 2018). A 95% credible interval means that the probability that the interevent time to the next eruption falls within this interval is 0.95. The credible intervals of the forecast time of the next VEI 3+ eruption (Figures 3a–3c) are obtained by adding 1,500 years. To evaluate the different models, we examined the residuals. The residual analysis (Supporting Information S1) revealed that model M3 performed the best, and model M2 performed the worst. One can use the Watanabe-Akaike information criterion (WAIC, Watanabe (2010)) to choose the best model. A model with the smallest WAIC among a set of candidate models is considered the best. Here, model M3 has the smallest WAIC (−41.37) and model M2 has the largest (−27.29). We thus adopt the forecast from model M3. The posterior distribution of model M3 suggests that there is a 5% probability that an eruption of VEI  $\geq 3$  will occur at Tongariro in the next 50 years, and a 50% probability before CE2967.



**Figure 3.** The posterior distribution of the forecast time of the next eruption from Tongariro. (a–c) Forecast for eruptions with Volcanic Explosivity Index (VEI)  $\geq 3$  using the statistical analogs with Holocene records. (d–f) Forecast for eruptions with VEI  $\geq 3$  using the empirical analogs. (g–i) Forecast for eruptions with VEI  $\geq 3$  using the statistical analogs with Quaternary records. (j–l) Forecast for eruptions with VEI  $\geq 4$  using the statistical analogs with Quaternary records. The blue dashed lines indicate the quantiles of the posterior distributions, from left to right in each plot, 2.5%, 25%, 50%, 75%, 97.5%. The red dotted lines indicate the 99.5% quantiles of the posterior distributions.

### 5. Forecast Using Statistical Analogs With Quaternary Records

To forecast eruptions of greater size (i.e., VEI 4+), a catalog combining historical and longer geological records is needed. The eruption records in the GVP catalog may record only 0.1% to 0.05% of a volcano’s geological history. Some long-lived volcanoes may not have large eruptions in the Holocene GVP record. Therefore, we

explored analogue volcanoes that have longer geological records of VEI 4+ eruptions that have published records in the literature. For Tongariro, the longer record includes 28 eruptions since ~17,000 years BP, with a maximum of VEI 5 and minimum VEI 1 (Figure 1x). This record augmented the GVP catalog by adding additional eruptions discovered in the form of distinct (3+ cm) tephra in a sedimentological investigation of the Tongariro Volcanic Center, and assigned to Red Crater (Moebis, 2011).

Here we consider statistical analogs with Quaternary records that were chosen via expert judgment based on geological properties. The number of available statistical analogs with long Quaternary eruption records is very small. Seeking these statistical analogs for volcanoes relies on matching, as best as possible, eruptive compositions, local tectonic setting, volcano morphology, and environmental conditions. This must also consider that volcanoes may undergo evolutionary change, requiring matching volcanoes in similar states of development and/or volcanoes that may plausibly experience similar changes in state. Even accounting for many differences between neighboring volcanoes along the same volcanic arc (or continental margin), the most common volcano-types are easily matched. Some are more difficult, such as Tongariro, because the volcano has a complex, distributed form comprising at least 14 known vents along an active tectonic graben (Gómez-Vasconcelos et al., 2017), and has erupted magmatic compositions spanning basalt to dacite (Hobden et al., 1999; Moebis et al., 2011). Few clear statistical analogs with Quaternary records exist, but two systems have longer periods of eruption record and a close geological/tectonic similarity: Three Sisters in the Cascade Arc in Oregon, USA, and Puyehue-Cordón Caulle in the Chilean portion of the Andean arc. These continental margin volcanoes erupt the same magma compositions as Tongariro, as well as being in a similar local tectonic setting of extension (Bacon, 1985; Guffanti & Weaver, 1988; Singer et al., 2008). Eruption records of the three volcanoes are presented in Figures 1x–1z and similarities between them are outlined in Table A1 in Appendix A. If a VEI is not specified, it was assumed to be a VEI  $\geq 4$  event. This is in accordance with the LaMEVE database (Croweller et al., 2012)—for any eruption in geological times to leave traces for dating in recent history, the tephra deposit must have been large enough to survive erosion. Duplicate ages are removed from the records of these analogue volcanoes.

The same three models are used to check if the three volcanoes, Tongariro, Three Sisters and Puyehue-Cordón Caulle, have similar statistical features. The posterior distributions of  $Y_i$  and  $Z_i$  in Figures 2i–2l suggest that there is no significant difference in the shape or scale of the hazard function among the three volcanoes. This is also confirmed by the posteriors for  $\sigma_Y$  and  $\sigma_Z$  (the standard deviations of  $Y_i$  and  $Z_i$ ) having mass close to 0, suggesting that the overall variation in  $Y_i$  and  $Z_i$  is not significantly different from 0. Thus, the statistical eruption-record features of Three Sisters and Puyehue-Cordón Caulle are not significantly different from those of Tongariro, and they are good analogs.

To provide forecasts for future Tongariro eruptions ( $\geq$ VEI 4), support from the statistical analogs with Quaternary records is added by using each of models M1, M2 and M3. The last eruption of VEI  $\geq 4$  from Tongariro occurred in 2,500 years BP. By considering only the VEI 4+ record for the three analogue volcanoes, (including the events with unspecified VEI), we obtained the credible intervals of the forecast time of the next eruption with VEI  $\geq 4$  from Tongariro (Figures 3j–3l). The residual analysis (Supporting Information S1) revealed that Model M3 performed the best, and model M2 performed the worst. We thus adopt the forecast from model M3. The posterior distributions for the forecast time of the next VEI  $\geq 4$  eruption (Figure 3i) suggest a 1% probability of an eruption of VEI  $\geq 4$  in the next 50 years, and a 50% probability before CE6872.

We also used the statistical analogs with Quaternary records to forecast VEI 3+ events to compare with the forecast from the statistical analogs with Holocene records. The last eruption of VEI  $\geq 3$  from Tongariro in the augmented catalog is CE1798 (Moebis, 2011), in contrast to the GVP record as described above. Considering only VEI  $\geq 3$  events of the statistical analogs with Quaternary records, and the best-performing Model M3, the posterior distribution from Tongariro (Figure 3f) suggests that there is a 5% probability of an eruption of VEI  $\geq 3$  at Tongariro in the next 50 years, and a 50% probability before CE3257.

Models M1 and M3 provide wider 95% credible intervals than the equivalent models obtained via the statistical analogs with Holocene records. Model M2 produces very narrow overall credible intervals for the statistical analogs with Holocene records. Comparing the residuals from M1, M2, and M3 (Supporting Information S1) shows that M3 best explains the data. We note that these analogs with Quaternary records have much longer records and larger eruptions than the Holocene analogs, and are more homogeneous in those respects. Although the analogs with Quaternary records produce forecasts with higher uncertainties (longer tails) for VEI 3+ events,

we can use them to forecast VEI 4+ events, which is not feasible with the Holocene eruption records. The longer tails could be due to the fact that some VEI 3 eruptions are missing from the older Quaternary records, or that there are fewer possible statistical analogs with Quaternary records. A recent study used either 50 or 100 analogs from the GVP catalog to reduce the uncertainty of forecast for Aluto (Central Ethiopia) (Tierz et al., 2020), but these analogs could not be tested in terms of their statistical similarity in repose time distribution, given that such datasets of repose times were not available for Aluto. One outstanding issue related to the use of sets of analogue volcanoes to forecast eruption onset and/or size relates to better understanding the difference between “volcano-specific analogs” and “global analogs” (Tierz, 2020; Tierz et al., 2020).

It is difficult to assess the relative reliability of results from the statistical analogs with Quaternary versus Holocene eruption records. Both types of analogs performed their function and considerably reduced the forecast uncertainty. Simply fitting the Weibull model in Equation 1 to only the target volcano, using the Quaternary record produces a 50% and 95% credible intervals of (2,387, 6733) and (2,041, 70,027) years, respectively, for the forecast time of next eruption. Using Tongariro's GVP Holocene record, these become (3,249, 19,100) and (2,098, 925,837) years for the forecast time of next eruption. Both of these results are wider intervals than the equivalent forecasts from either the statistical analogs with Quaternary or Holocene records (Figure 3). The credible intervals obtained by using the GVP record are particularly wide and thus imply much larger uncertainty, which is due to the fact that there are only 6 VEI  $\geq 3$  eruptions in the GVP record. This is the main reason why analogue volcanoes are needed/sought to assess volcanic hazard at Tongariro (and many other volcanoes worldwide). Further, fitting model M3 to VEI 3+ eruptions from the empirical analogs with Holocene records (i.e., all the candidates in Table 1) results in a forecast for Tongariro having a 95% credible interval that is 1.5 times as wide as the equivalent obtained from its subset of statistical analogs (Figures 3c and 3f). By selecting a subset of empirical analogs with a closer distribution of repose times to the one shown by Tongariro (i.e., the statistical analogs), we expect this reduction in the variability of the forecast time of the next eruption.

## 6. Discussion

The records for Tongariro and the empirical analogue volcanoes may be incomplete with eruptions missing for geological/preservation reasons. We chose the empirical analogs based on volcanological expertise using these potentially incomplete and time-inhomogeneous records to include proxies for both the actual behavior and the observation pattern in the analogs. The missing data could still hinder our understanding of the volcano. Geological records could be more useful in this case as they include a greater number of large eruptions that are less likely to be missing. As tested by the empirical and statistical analogs with Holocene records, our method provides a good framework for modeling and borrowing strength from a set of statistical analogs to forecast eruptions from a target volcano. A possible future extension to the method could employ a trend-renewal process to directly model the trend that is either caused by missing data, or actual time-inhomogeneity (Bebbington, 2010).

The results from both the statistical analogs with Quaternary and Holocene records show that varying the scale parameter of the hazard rate function between the volcanoes is important for capturing the main features of the eruption records (i.e., M1 and M3 performed much better than M2). In other words, the clustering behavior of the volcano is more important for selecting analogs than the activity level. This is also practically important, because a target volcano may have a low activity rate, and the analogs should have more recorded eruptions, or a higher activity rate. Model M3 performs slightly better than M1 based on the residual analysis. This means that more robust forecasts can be produced by incorporating a certain amount of variation in clustering behavior, provided there is enough variability in the activity level.

When modeling and forecasting using both the statistical analogs with Quaternary and Holocene records, we assumed that the records of these analogue volcanoes are complete for eruptions with VEI  $\geq 3$ . This assumption could potentially underestimate the hazard if some VEI  $\geq 3$  eruptions were missing from the records. While this important issue has been studied (e.g., Mead & Magill, 2014; Wang & Bebbington, 2012, 2013; Wang et al., 2020), the solutions all require either a long homogeneous record, or a very large set of volcanoes. Previous studies on missing data make simple assumptions, such as the global large volcanic eruptions following a Poisson process, or the probabilities of an eruption missing from a record dependent only on time (or region, time and VEI). It is impossible to justify these assumptions for all the volcanoes considered in this study, but more geological constraints to incorporate possible missing data should be pursued. However, the structure of the hierarchical

Bayesian model we use in this study and the strength that we borrow from the analogue volcanoes reduce the influence of missing events. To demonstrate this, we carried out a sensitivity analysis exploring the influence of missing data in the eruption records on the forecasts (see Supporting Information S1). The simulation study suggests that the hierarchical Bayesian model greatly reduces the influence of missing events on hazard forecasts with up to 40% missing data from the records of all the volcanoes (both target and analogue volcanoes).

In this paper, we selected candidate volcanoes (empirical analogs) based on their sharing similar geological feature—satisfying at least two of the three constraints: composite volcanoes, erupting basaltic to andesite compositions and sited in convergent margins. We also used eruptions with minimum VEI 3, that is, moderately-explosive eruptions. By choosing this way, we are minimizing the potential influence of not directly modeling VEI on the forecasts. As discussed by Sheldrake et al. (2016, 2017), eruptions of the same size or type can occur at two volcanoes, but the processes leading to the events can be very different. For the same volcano, eruptions of different VEIs have different occurrence rates. Future research can consider incorporating the size and type of eruptions in the model. However, this may be difficult given that many geological records do not have VEI information provided.

## 7. Conclusion

We can improve hazard forecasts at single volcanoes with poor or incomplete eruption records by applying a multivariate Weibull renewal process that borrows strength from analogue volcanoes. These volcanic systems have similar geological/tectonic and statistical features to the target volcano, but longer, or more complete eruption records. Whether a set of candidate volcanoes can be considered “good” analogs in terms of eruption occurrence can be tested using Bayesian inference. Using this model we forecast future VEI 3+ or 4+ eruptions from Tongariro volcano (New Zealand), by identifying geologically reasonable analogs from the GVP catalog. In general terms, what is needed is a means to ‘seed’ analogue sets as the statistical procedure then eliminates “outlier” volcanoes. The VOLCANS method (Tierz et al., 2019) could certainly be one such potential seeding system. Expert knowledge is another such method.

We identified a subset of 17 volcanoes with compatible statistical features and produced forecast for VEI 3+ events. Just using the two most closely geologically similar analogs (Three Sisters and Puyehue-Cordón Caulle), forecasts have long tails, so that the credible intervals are very wide, suggesting large uncertainties. By applying the 17 identified statistical analogue volcanoes, forecasts show much narrower 50% and 95% credible intervals, and thus lower uncertainties.

Due to low numbers of large events in historical records, only expert-matched analogs in the Quaternary eruption record can produce reliable forecasts for VEI 4+ events, particularly when using a reduced number of analogue volcanoes. The method proposed in this study can be readily applied to other volcanoes with scarce record of past eruptions, or to validate and compare with other data-driven and expert-elicitation approaches for choosing analogue volcanoes for hazard assessment and/or emergency planning purposes. Note that application of this method requires that at least some observations of eruption dates are available from the target volcano.

## Appendix A: Data for Statistical Analogs With Quaternary Records

Table A1 shows the characteristics for the statistical analogs of Tongariro with Quaternary records. The eruption record for Three Sisters has 71 eruptions since 541,800 years BP, but with few having volume estimates (Fierstein et al., 2011; Schmidt & Grunder, 2009). The event record for Puyehue-Cordón Caulle includes 82 eruptions since 314,000 years BP, with volume estimates available for most eruptions (Lara et al., 2006; Singer et al., 2008).

**Table A1**  
*Characteristics for the Three Volcanoes, Tongariro, Puyehue-Cordón Caulle and Three Sisters*

Volcano name	Country	VT	MRT	TS	$N_{3,i}$	$N_{4,i}$	Order
Tongariro	NZ	SV, PC	A/BA	CC, SZ	14	9	1
Three Sisters	USA	C, PC, S, FV	A/BA, B/PB, D	CC, SZ	69	69	2
Puyehue-Cordón Caulle	Chile	SV, CD, FV, PC, LD	A/BA, D, B/PB, R	CC, SZ	74	72	3

*Note.* Notation: Volcano Type (VT), CD = Caldera, C = Complex, FV = Fissure Vents, LD = Lava Domes, M = Maar. PC = Pyroclastic Cones, S = Shield, SV = Stratovolcano(es); Major Rock Type (MRT), A = Andesite, BA = Basaltic Andesite, B = Basalt, D = Dacite, PB = Picro-Basalt, R = Rhyolite; Tectonic Setting (TS), CC = Continental crust, OC = Oceanic crust, SZ = Subduction zone;  $N_{3,i}$ : number of eruptions with minimum VEI 3;  $N_{4,i}$ : number of eruptions with minimum VEI 4.

## Appendix B: Parameter Estimation for Models M1, M2, and M3

Given the observed interevent times  $x_{i1}, \dots, x_{iN_i}$  for the  $i$ th individual with the final censored time  $s_i$ , the log-likelihood of each interevent time for model  $j$  is

$$L_{jk}(\alpha, \beta) = \log(\lambda_{j,i}(x_{ik})) - \int_0^{x_{ik}} \lambda_{j,i}(t) dt, \quad (B1)$$

where  $k = 1, \dots, N_i$ ; and the log-likelihood of the censored time of the  $i$ th individual for model  $j$  is

$$L_{js_i}(\alpha, \beta) = - \int_0^{s_i} \lambda_{j,i}(t) dt. \quad (B2)$$

The explicit form of the log-likelihood function for model M1 is

$$L_{1k}(\alpha, \beta) = \log(Y_i) + \log(\alpha) + (\alpha - 1)\log(x_{ik}) - \alpha \log(\beta) - \frac{Y_i x_{ik}^\alpha}{\beta^\alpha}, \quad L_{1s_i}(\alpha, \beta) = - \frac{Y_i s_i^\alpha}{\beta^\alpha}; \quad (B3)$$

for model M2 it is

$$L_{2k}(\alpha, \beta) = \log(\alpha) + (Z_i \alpha - 1) \log(x_{ik}) - \alpha \log(\beta) - \frac{x_{ik}^{Z_i \alpha}}{Z_i \beta^\alpha}, \quad L_{2s_i}(\alpha, \beta) = - \frac{s_i^{Z_i \alpha}}{Z_i \beta^\alpha}; \quad (B4)$$

and for model M3 it is

$$L_{3k}(\alpha, \beta) = \log(Y_i \alpha) + (Z_i \alpha - 1) \log(x_{ik}) - \alpha \log(\beta) - \frac{Y_i x_{ik}^{Z_i \alpha}}{Z_i \beta^\alpha}, \quad L_{3s_i}(\alpha, \beta) = - \frac{Y_i s_i^{Z_i \alpha}}{Z_i \beta^\alpha}. \quad (B5)$$

We use Bayesian methods to estimate the parameters. An MCMC algorithm generates samples from the joint posterior distribution of  $\alpha, \beta, Y_i, Z_i, \sigma_Y$ , and  $\sigma_Z$  given the interevent times  $x_{i1}, \dots, x_{iN_i}$  and the final censored times  $s_i$  for  $i = 1, \dots, M$  volcanoes, using software JAGS and the R2jags package in R (Su & Yajima, 2012). We use three chains, half-normal priors for  $\alpha$  and  $\beta$ , and weakly informative half- $t$  prior distributions for the variance parameters  $\sigma_Y$  and  $\sigma_Z$  (Gelman, 2006),

$$\begin{aligned} \alpha &\sim N(0, 100^2) T(0, ), & \beta &\sim N(0, 100^2) T(0, ), \\ \sigma_Y &\sim dt(0, 0.04, 3) T(0, ), & \sigma_Z &\sim dt(0, 0.04, 3) T(0, ). \end{aligned}$$

A renewal process has independent and identically distributed interevent times. To test whether the model is a good fit to the data, we calculate the posterior predictive distribution of  $\bar{F}_n(x_{ij}) - \hat{F}(x_{ij})$  ( $i = 1, \dots, M$ ;  $j = 1, 2, \dots, N_i$ ) and check if this is centered around 0, where  $\bar{F}_n(x_{ij})$  is the empirical cumulative distribution of the interevent times,  $\hat{F}(x_{ij})$  is the estimated cumulative distribution of the interevent times from one of the models M1, M2, M3, and  $N_i$  is the number of interevent times from the  $i$ th volcano. If the distribution of this difference is centered around zero, then the model is considered to have captured the main features of the data.

## Data Availability Statement

Datasets for the empirical analogues with Holocene records (from GVP) may be downloaded from <https://doi.org/10.5281/zenodo.6484912> under the branch ‘‘HoloceneRecords’’, which were obtained from the Global Volcanism Program, Smithsonian Institution <https://volcano.si.edu/>. Datasets for the statistical analogues with Quaternary records are available via <https://doi.org/10.5281/zenodo.6484912> under the branch ‘‘QuaternaryRecords.’’ For the Quaternary records, the eruption record for Three Sisters is from Fierstein et al. (2011) (<https://doi.org/10.1016/j.jvolgeores.2011.06.003>) and Schmidt and Grunder (2009) (<https://doi.org/10.1130/B26442.1>), and the record for Puyehue-Cordón Caulle is from Lara et al. (2006) (<https://doi.org/10.1016/j.jvolgeores.2006.04.010>) and Singer et al. (2008) (<https://doi.org/10.1130/B26276.1>). All the R code is available via <https://doi.org/10.5281/zenodo.6484912>.

### Acknowledgments

This work was funded by the National Hazards Research Platform (Grant No. 2015-MAU-01-NHRP), MBIE Endeavour Fund (Research Programme, Grant No. UOAX1913) the Resilience National Science Challenge: Volcanism theme (Grant No. GNS-RNC047). We thank Tom Sheldrake, Pablo Tierz, and an anonymous reviewer for their constructive comments that greatly improved this manuscript. We thank Marco Brenna for helpful discussions and the New Zealand eScience Infrastructure (NeSI) for providing access to the supercomputers. Open access publishing facilitated by University of Otago, as part of the Wiley – University of Otago agreement via the Council of Australian University Librarians.

### References

- Aalen, O. O., & Husebye, E. (1991). Statistical analysis of repeated events forming renewal processes. *Statistics in Medicine*, *10*(8), 1227–1240. <https://doi.org/10.1002/sim.4780100806>
- Bacon, C. R. (1985). Implications of silicic vent patterns for the presence of large crustal magma chambers. *Journal of Geophysical Research*, *90*(B13), 11243–11252. <https://doi.org/10.1029/jb090ib13p11243>
- Bebbington, M. S. (2010). Trends and clustering in the onsets of volcanic eruptions. *Journal of Geophysical Research*, *115*(B1), B01203. <https://doi.org/10.1029/2009jb006581>
- Bebbington, M. S. (2014). Long-term forecasting of volcanic explosivity. *Geophysical Journal International*, *197*(3), 1500–1515. <https://doi.org/10.1093/gji/ggu078>
- Bebbington, M. S., & Cronin, S. J. (2011). Spatio-temporal hazard estimation in the Auckland Volcanic Field, New Zealand, with a new event-order model. *Bulletin of Volcanology*, *73*(1), 55–72. <https://doi.org/10.1007/s00445-010-0403-6>
- Bebbington, M. S., & Lai, C. D. (1996a). On nonhomogeneous models for volcanic eruptions. *Mathematical Geology*, *28*(5), 585–600. <https://doi.org/10.1007/bf02066102>
- Bebbington, M. S., & Lai, C. D. (1996b). Statistical analysis of New Zealand volcanic occurrence data. *Journal of Volcanology and Geothermal Research*, *74*(1–2), 101–110. [https://doi.org/10.1016/s0377-0273\(96\)00050-9](https://doi.org/10.1016/s0377-0273(96)00050-9)
- Bebbington, M. S., Stirling, M., Cronin, S., Wang, T., & Jolly, G. (2018). National-level long-term eruption forecasts by expert elicitation. *Bulletin of Volcanology*, *80*(6), 56. <https://doi.org/10.1007/s00445-018-1230-4>
- Crowther, H. S., Arora, B., Brown, S. K., Cottrell, E., Deligne, N. I., Guerrero, N. O., et al. (2012). Global database on large magnitude explosive volcanic eruptions (LaMEVE). *Journal of Applied Volcanology*, *1*(1), 1–13. <https://doi.org/10.1186/2191-5040-1-4>
- De la Cruz-Reyna, S., & Carrasco-Nunez, G. (2002). Probabilistic hazard analysis of Citlaltepeltl (Pico de Orizaba) Volcano, eastern Mexican Volcanic Belt. *Journal of Volcanology and Geothermal Research*, *113*(1–2), 307–318. [https://doi.org/10.1016/s0377-0273\(01\)00263-3](https://doi.org/10.1016/s0377-0273(01)00263-3)
- Fierstein, J., Hildreth, W., & Calvert, A. T. (2011). Eruptive history of South sister, Oregon Cascades. *Journal of Volcanology and Geothermal Research*, *207*(3–4), 145–179. <https://doi.org/10.1016/j.jvolgeores.2011.06.003>
- Flandoli, F., Giorgi, E., Aspinall, W. P., & Neri, A. (2011). Comparison of a new expert elicitation model with the Classical Model, equal weights and single experts, using a cross-validation technique. *Reliability Engineering & System Safety*, *96*(10), 1292–1310. <https://doi.org/10.1016/j.res.2011.05.012>
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, *1*, 515–533.
- Gómez-Vasconcelos, M. G., Villamor, P., Cronin, S. J., Procter, J. N., Palmer, A. S., Townsend, D., & Leonard, G. (2017). Crustal extension in the Tongariro Graben, New Zealand: Insights into volcano-tectonic interactions and active deformation in a young continental rift. *The Geological Society of America Bulletin*, *129*(9–10), 1085–1099. <https://doi.org/10.1130/b31657.1>
- Green, R. M., Bebbington, M. S., Cronin, S. J., & Jones, G. (2013). Geochemical precursors for eruption repose length. *Geophysical Journal International*, *193*(2), 855–873. <https://doi.org/10.1093/gji/ggt044>
- Gregg, D. R. (1960). The geology of Tongariro subdivision. *New Zealand Geological Survey Bulletin*, *40*, 152.
- Guffanti, M., & Weaver, C. S. (1988). Distribution of late Cenozoic volcanic vents in the Cascade Range: Volcanic arc segmentation and regional tectonic considerations. *Journal of Geophysical Research*, *93*(B6), 6513–6529. <https://doi.org/10.1029/jb093ib06p06513>
- Heinrich, M., Cronin, S. J., & Pardo, N. (2020). Understanding multi-vent Plinian eruptions at Mt. Tongariro Volcanic complex, New Zealand. *Bulletin of Volcanology*, *82*(3), 30. <https://doi.org/10.1007/s00445-020-1369-7>
- Hobden, B. J., Houghton, B. F., Davidson, J. P., & Weaver, S. D. (1999). Small and short-lived magma batches at composite volcanoes: Time windows at Tongariro volcano, New Zealand. *Journal of the Geological Society*, *156*(5), 865–868. <https://doi.org/10.1144/gsjgs.156.5.0865>
- Hougaard, P. (1986). A class of multivariate failure time distributions. *Biometrika*, *73*(3), 671–678. <https://doi.org/10.2307/2336531>
- Hurst, T., & Smith, W. (2010). Volcanic ashfall in New Zealand—Probabilistic hazard modelling for multiple sources. *New Zealand Journal of Geology and Geophysics*, *53*, 1–14. <https://doi.org/10.1080/00288301003631129>
- Lara, L. E., Moreno, H., Naranjo, J. A., Matthews, S., & Pérez de Arce, C. (2006). Magmatic evolution of the Puyehue-cordón Caulle Volcanic Complex (40°S), Southern Andean Volcanic Zone: From shield to unusual rhyolitic fissure volcanism. *Journal of Volcanology and Geothermal Research*, *157*(4), 343–366. <https://doi.org/10.1016/j.jvolgeores.2006.04.010>
- Loughlin, S. C., Vye-Brown, C., Sparks, R., Brown, S. K., Barclay, J., Calder, E., et al. (2015). An introduction to global volcanic hazard and risk. In S. C. Loughlin, S. Sparks, S. K. Brown, S. F. Jenkins, & C. Vye-Brown (Eds.), *Global volcanic hazards and risk* (pp. 1–80). Cambridge University Press. <https://doi.org/10.1017/CBO9781316276273.003>
- Marzocchi, W., & Bebbington, M. S. (2012). Probabilistic eruption forecasting at short and long time scales. *Bulletin of Volcanology*, *74*(8), 1777–1805. <https://doi.org/10.1007/s00445-012-0633-x>
- Marzocchi, W., Sandri, L., Gasparini, P., Newhall, C., & Boschi, E. (2004). Quantifying probabilities of volcanic events: The example of volcanic hazard at Mount Vesuvius. *Journal of Geophysical Research*, *109*(B11), B11201. <https://doi.org/10.1029/2004jb003155>
- Mead, S., & Magill, C. (2014). Determining change points in data completeness for the Holocene eruption record. *Bulletin of Volcanology*, *76*(11), 1–14. <https://doi.org/10.1007/s00445-014-0874-y>
- Moebis, A. (2011). *Understanding the Holocene explosive eruption record of the Tongariro Volcanic Centre, New Zealand* (Doctoral Dissertation). Massey University. Retrieved from <http://hdl.handle.net/10179/2200>
- Moebis, A., Cronin, S. J., Neall, V. E., & Smith, I. E. M. (2011). Unravelling a complex volcanic history from fine-grained, intricate Holocene ash sequences at the Tongariro Volcanic Centre, New Zealand. *Quaternary International*, *246*(1–2), 352–363. <https://doi.org/10.1016/j.quaint.2011.05.035>
- Neri, A., Aspinall, W. P., Cioni, R., Bertagnini, A., Baxter, P. J., Zuccaro, G., et al. (2008). Developing an event tree for probabilistic hazard and risk assessment at Vesuvius. *Journal of Volcanology and Geothermal Research*, *178*(3), 397–415. <https://doi.org/10.1016/j.jvolgeores.2008.05.014>
- Newhall, C. G., & Pallister, J. S. (2015). Using multiple data sets to populate probabilistic volcanic event trees. In *Volcanic hazards, risks and disasters* (pp. 203–232). Elsevier.
- Newhall, C. G., & Punongbayan, R. (1996). *Fire and mud: Eruptions and lahars of Mount Pinatubo, Philippines*. Philippine Institute of Volcanology and Seismology Quezon City.
- Newhall, C. G., & Self, S. (1982). The Volcanic Explosivity Index (VEI): An estimate of explosive magnitude for historical volcanism. *Journal of Geophysical Research*, *87*(C2), 1231–1238. <https://doi.org/10.1029/jc087ic02p01231>
- Pardo, N., Cronin, S. J., Németh, K., Brenna, M., Schipper, C. I., Breard, E., et al. (2014). Perils in distinguishing phreatic from phreatomagmatic ash; insights into the eruption mechanisms of the 6 August 2012 Mt. Tongariro eruption, New Zealand. *Journal of Volcanology and Geothermal Research*, *286*, 397–414. <https://doi.org/10.1016/j.jvolgeores.2014.03.001>

- Rodado, A., Bebbington, M. S., Noble, A., Cronin, S., & Jolly, G. (2011). On selection of analogue volcanoes. *Mathematical Geosciences*, 43(5), 505–519. <https://doi.org/10.1007/s11004-011-9345-6>
- Schmidt, M. E., & Grunder, A. L. (2009). The evolution of North Sister: A volcano shaped by extension and ice in the central Oregon Cascade Arc. *The Geological Society of America Bulletin*, 121(5–6), 643–662. <https://doi.org/10.1130/b26442.1>
- Scott, B. J., & Potter, S. H. (2014). Aspects of historical eruptive activity and volcanic unrest at Mt. Tongariro, New Zealand: 1846–2013. *Journal of Volcanology and Geothermal Research*, 286, 263–276. <https://doi.org/10.1016/j.jvolgeores.2014.04.003>
- Shane, P., Doyle, L. R., & Nairn, I. A. (2008). Heterogeneous andesite-dacite ejecta in 26–16.6 ka pyroclastic deposits of Tongariro Volcano, New Zealand: The product of multiple magma-mixing events. *Bulletin of Volcanology*, 70(4), 517–536. <https://doi.org/10.1007/s00445-007-0152-3>
- Sheldrake, T. E. (2014). Long-term forecasting of eruption hazards: A hierarchical approach to merge analogous eruptive histories. *Journal of Volcanology and Geothermal Research*, 286, 15–23. <https://doi.org/10.1016/j.jvolgeores.2014.08.021>
- Sheldrake, T. E., Aspinall, W. P., Odbert, H. M., Wadge, G., & Sparks, R. S. J. (2017). Understanding causality and uncertainty in volcanic observations: An example of forecasting eruptive activity on Soufrière Hills Volcano, Montserrat. *Journal of Volcanology and Geothermal Research*, 341, 287–300. <https://doi.org/10.1016/j.jvolgeores.2017.06.007>
- Sheldrake, T. E., Sparks, R. S. J., Cashman, K. V., Wadge, G., & Aspinall, W. P. (2016). Similarities and differences in the historical records of lava dome-building volcanoes: Implications for understanding magmatic processes and eruption forecasting. *Earth-Science Reviews*, 160, 240–263. <https://doi.org/10.1016/j.earscirev.2016.07.013>
- Singer, B. S., Jicha, B. R., Harper, M. A., Naranjo, J. A., Lara, L. E., & Moreno-Roa, H. (2008). Eruptive history, geochronology, and magmatic evolution of the Puyehue-Cordón Caulle volcanic complex, Chile. *GSA Bulletin*, 120(5–6), 599–618. <https://doi.org/10.1130/b26276.1>
- Stirling, M. W., & Wilson, C. J. N. (2002). Development of a volcanic hazard model for New Zealand: First approaches from the methods of probabilistic seismic hazard analysis. *Bulletin of the New Zealand Society for Earthquake Engineering*, 35(4), 266–277. <https://doi.org/10.5459/bnzsee.35.4.266-277>
- Su, Y. S., & Yajima, M. (2012). *R2jags: A package for running jags from R. R package version 0.03-08*.
- Tierz, P. (2020). Long-term probabilistic volcanic hazard assessment using open and non-open data: Observations and current issues. *Frontiers of Earth Science*, 8, 257. <https://doi.org/10.3389/feart.2020.00257>
- Tierz, P., Christodoulou, V., Stevenson, J. A., & Loughlin, S. C. (2021). PyVOLCANS: A python package to flexibly explore similarities and differences between volcanic systems. *Journal of Open Source Software*, 6(68), 3649. <https://doi.org/10.21105/joss.03649>
- Tierz, P., Clarke, B., Calder, E. S., Dessalegn, F., Lewi, E., Yirgu, G., et al. (2020). Event trees and epistemic uncertainty in long-term volcanic hazard assessment of rift volcanoes: The example of Aluto (Central Ethiopia). *Geochemistry, Geophysics, Geosystems*, 21(10), e2020GC009219. <https://doi.org/10.1029/2020gc009219>
- Tierz, P., Loughlin, S. C., & Calder, E. S. (2019). Volcans: An objective, structured and reproducible method for identifying sets of analogue volcanoes. *Bulletin of Volcanology*, 81(12), 76. <https://doi.org/10.1007/s00445-019-1336-3>
- Topping, W. W. (1973). Tephrostratigraphy and chronology of late quaternary eruptives from Tongariro Volcanic Centre, New Zealand. *New Zealand Journal of Geology and Geophysics*, 16(3), 397–423. <https://doi.org/10.1080/00288306.1973.10431368>
- Turner, M., Cronin, S., Bebbington, M. S., & Platz, T. (2008). Developing a probabilistic eruption forecast for dormant volcanoes: A case study from Mt. Taranaki, New Zealand. *Bulletin of Volcanology*, 70(4), 507–515. <https://doi.org/10.1007/s00445-007-0151-4>
- Wang, T., & Bebbington, M. S. (2012). Estimating the likelihood of an eruption from a volcano with missing onsets in its record. *Journal of Volcanology and Geothermal Research*, 243–244, 14–23. <https://doi.org/10.1016/j.jvolgeores.2012.06.032>
- Wang, T., & Bebbington, M. S. (2013). Robust estimation for the Weibull process applied to eruption records. *Mathematical Geosciences*, 45(7), 851–872. <https://doi.org/10.1007/s11004-013-9461-6>
- Wang, T., Schofield, M., Bebbington, M. S., & Kiyosugi, K. (2020). Bayesian modelling of marked point processes with incomplete records: Volcanic eruptions. *Journal of the Royal Statistical Society: Series C: Applied Statistics*, 69(1), 109–130. <https://doi.org/10.1111/rssc.12380>
- Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research*, 11, 3571–3594.
- Weidner, D. M., Laya, G. E., & Serano, J. A. (2001). World swordfish fisheries: An analysis of swordfish fisheries, market trends and trade patterns, past-present-future. *Volume IV. Latin America, Part B. Caribbean, Section, 4*, 1053–1145.
- Whelley, P. L., Newhall, C. G., & Bradley, K. E. (2015). The frequency of explosive volcanic eruptions in Southeast Asia. *Bulletin of Volcanology*, 77(1), 1–11. <https://doi.org/10.1007/s00445-014-0893-8>