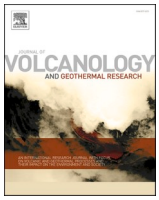
Journal of Volcanology and Geothermal Research 432 (2022) 107687

Contents lists available at ScienceDirect

Journal of Volcanology and Geothermal Research

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



# Development of a Bayesian event tree for short-term eruption onset forecasting at Taupō volcano

Emmy Scott [a,*], Mark Bebbington [b], Thomas Wilson [a], Ben Kennedy [a], Graham Leonard [c]

[a] University of Canterbury, Private Bag 4800, Christchurch, New Zealand
[b] Massey University, Private Bag 11222, Palmerston North 4442, New Zealand
[c] GNS Science, PO Box 30368 Lower Hutt, New Zealand

ARTICLE INFO

*Keywords:*
Eruption forecasting
Bayesian event tree
Expert elicitation
Caldera unrest
Taupō Volcanic Zone

ABSTRACT

Taupō volcano, located within the Taupō Volcanic Zone (TVZ) in the central North Island of Aotearoa-New Zealand, is one of the world's most active silicic caldera systems. Silicic calderas such as Taupō are capable of a broad and complex range of volcanological activity, ranging from minor unrest episodes to large destructive supereruptions. A critical tool for volcanic risk management is eruption forecasting. The Bayesian Event Tree for Eruption Forecasting (BET_EF) is one probabilistic eruption forecasting tool that can be used to produce short-term eruption forecasts for any volcano worldwide. A BET_EF model is developed for Taupō volcano, informed by geologic and historic data. Monitoring parameters for the model were obtained through a structured expert elicitation workshop with 30 of Aotearoa-New Zealand's volcanologists and volcano monitoring scientists. The eruption probabilities output by the BET_EF model for Taupō volcano's 17 recorded unrest episodes (between 1877 and 2019) were examined. We found time-inhomogeneity in the probabilities stemming from both the changes over time in the monitoring network around Taupō volcano and increasing level of past data (number of non-eruptive unrest episodes). We examine the former issue through the lens of the latest episodes, and the latter by re-running the episodes assuming knowledge of all 16 other episodes (calibration to 2021 data). The time variable monitoring network around Taupō volcano and parameter weights had a substantial impact on the estimated probabilities of magmatic unrest and eruption. We also note the need for improved monitoring and data processing at Taupō volcano, the existence of which would prompt updates and therefore refinements in the BET_EF model.

## 1. Introduction

Volcanic eruptions are destructive natural phenomena that can pose a constant threat to populations around the world. Eruption forecasting can be long- or short-term, or a combination of the two (Newhall and Hoblitt, 2002). Long-term forecasts use geologic and historical data to produce long-term probabilities and are often used for land-use planning and risk mitigation methods (Marzocchi and Bebbington, 2012) (e.g., land-use planning against lahars Tabayag, 2010; Pierson et al., 2014, and early warning systems Lavigne et al., 2000; Massey et al., 2010). Short-term forecasts are used in heightened unrest and eruption periods and typically use monitoring data (e.g., seismicity, deformation, gas) (Marzocchi and Bebbington, 2012; Marzocchi et al., 2008). If monitoring detects signs that may indicate a potential eruption, scientists can use them to provide short-term eruption forecasts. These forecasts should provide the likelihood, spatial extent, and timing of volcanic activity, to key stakeholders and decision-makers (Sparks, 2003).

Eruptions at caldera volcanoes are particularly challenging to forecast due to infrequent eruptions, much wider range of eruption magnitudes, and poorly understood episodes of volcanic unrest (Phillipson et al., 2013). Volcanic unrest is defined by Phillipson et al. (2013) as "the deviation from the background or baseline behaviour of a volcano towards a behaviour which is a cause for concern in the short-term because it may prelude an eruption". Unfortunately, unrest is often not clearly defined in the context of any particular volcano. Only well-studied volcanoes, if then, will have enough monitored data to determine what behaviour is above "background or baseline behaviour". Worldwide, silicic calderas that have not erupted for at least 100 years previously have experienced 60 unrest episodes, only 10 of which led to eruption (Newhall and Dzurisin, 1988). However, every eruption was

* Corresponding author.
*E-mail address:* emmy.scott@pg.canterbury.ac.nz (E. Scott).

https://doi.org/10.1016/j.jvolgeores.2022.107687
Received 14 June 2022; Received in revised form 9 September 2022; Accepted 30 September 2022
Available online 12 October 2022

preceded by an unrest episode (Acocella et al., 2015), although two of these may not have been observed (Sandri et al., 2017).

### 1.1. Taupō volcano

Taupō volcano in Aotearoa-New Zealand has a history of very large, explosive eruptions that impacted large areas (Wilson, 1993; Sutton et al., 1995; Wilson et al., 2006). A large volcanic eruption from Taupō volcano could pose a severe risk to populations, requiring the evacuation of large parts of the central North Island (e.g., Manville et al., 2009).

Taupō volcano is the southernmost caldera within the Taupō Volcanic Zone (TVZ), marking the southern boundary of the central TVZ. The volcano is most obviously reflected in the modern shape of Lake Taupō - Taupō-nui-a-Tia, which conceals the area of the Oruanui and Taupō collapse structures and young vent sites. The caldera structural margin is expressed by a large and clearly defined negative gravity anomaly (Bibby et al., 1995; Davy and Caldwell, 1998; Stagpoole et al., 2021).

The Oruanui eruption ejected $> 530$ km$^3$ DRE of magma ($> 1100$ km$^3$ of pyroclastic material) over several months (Davy and Caldwell, 1998; Wilson, 2001; Wilson et al., 2006). Caldera collapse occurred in the climactic Phase 10, which has now infilled and reflects the modern shape of Lake Taupō (Davy and Caldwell, 1998; Wilson, 2001; Stagpoole et al., 2021). The 232 AD Taupō eruption was the most recent known explosive eruption from the volcano, producing $> 35$km$^3$ DRE of magma and a further caldera collapse (Bibby et al., 1995; Davy and Caldwell, 1998). The 27 known eruptions between the 26.5 ka Oruanui and the 232 AD Taupō events varied widely in volume ($> 0.001 - 16.9$ km$^3$), with distinct eruption styles (degree of magma-water interaction), intra-eruptive time breaks, and potential hazards (Wilson, 1993; Barker et al., 2020). Fresh obsidian in the lithic fractions found in many of the smaller Holocene eruptives suggest that many of

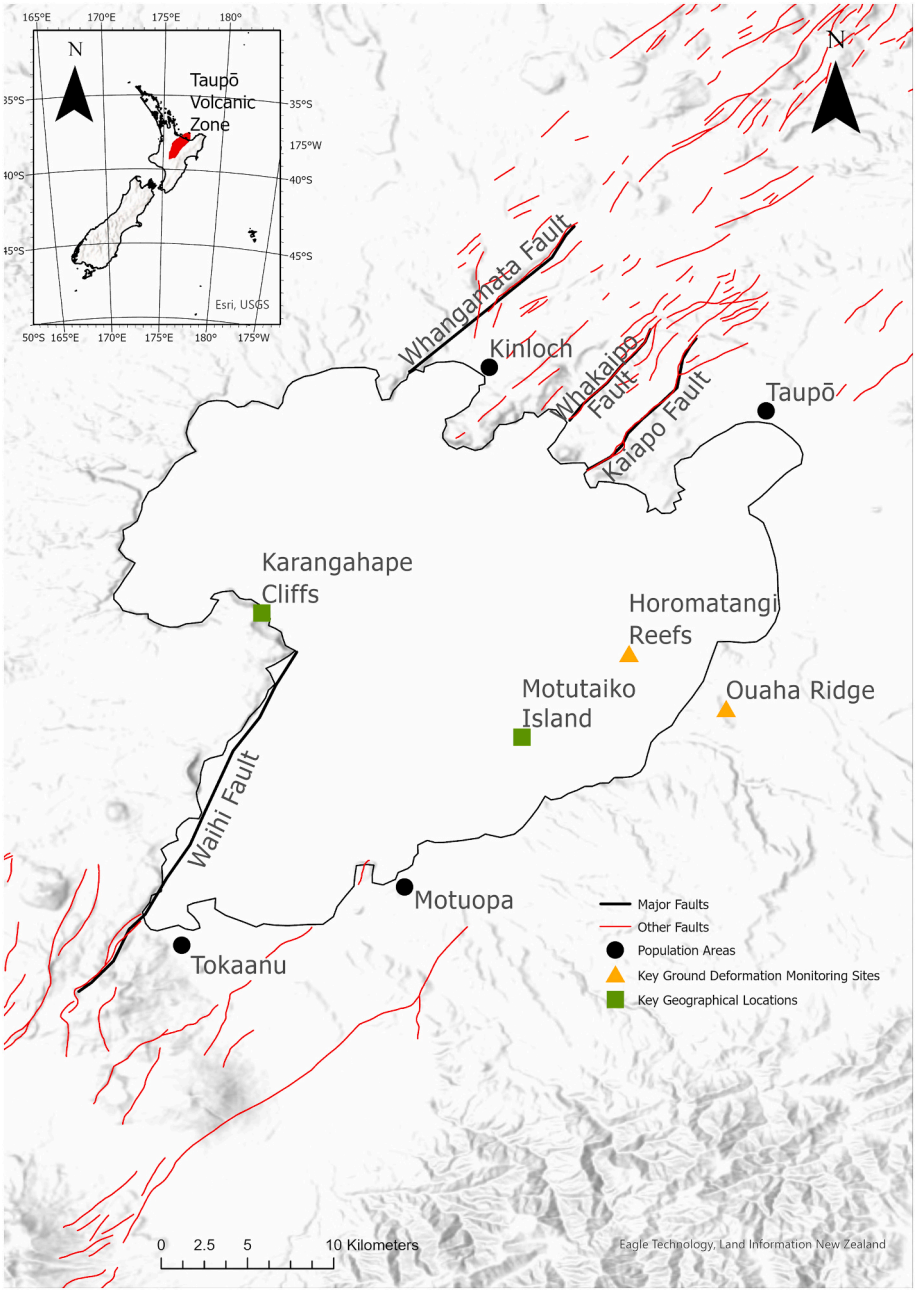



**Fig. 1.** Map of Taupō volcano highlighting key features and sites.

the explosive eruptions accompanied or followed by dome extrusion, which were later destroyed during the 232 AD eruption (Wilson, 1993).

The crystal mush system at Taupō volcano is reasonably well delineated in regards to its size, location, and physical state (Barker et al., 2020; Illsley-Kemp et al., 2021). It is believed that >200 $km^3$ to possibly 1000 $km^3$ of crystal mush exists between 5–8 km depth within the crust, underneath Lake Taupō (Barker et al., 2015). Calculations were derived from vent spacing and inferred crystallisation depths, but it is only a first-order calculation, and lateral migration of magma during unrest is possible. Geophysical evidence from Illsley-Kemp et al. (2021) suggests the modern magmatic reservoir covers a minimum area of 80 $km^2$, with a maximum area possibly exceeding 250 $km^2$.

Taupō volcano is surrounded by three geothermal areas and has one within the lake. The southernmost geothermal area is Tokaanu, located at the southern tip of the lake. Wairakei and Tauhara geothermal systems are located just north of Lake Taupō and are commercial geothermal power areas. The Wairakei/Tauhara geothermal systems are inferred to have been initiated around 60kya by magmatism linked to the NE-Dome system, and this maintains the systems thermal flux today (Rosenberg et al., 2020). The geothermal system within Lake Taupō is Horomatangi Reefs, which is located on the inferred Taupō eruption vent site and contains two hydrothermal vents - Te Hoata and Te Pupu (De Ronde et al., 2002) (Fig. 1). However, studies about Horomatangi Reefs are limited. While the NE-Domes and the Taupō magmatic systems are spatially close, they are considered to be two different systems (Sutton et al., 1995; Charlier et al., 2005; Wilson and Charlier, 2009). However, due to monitoring limitations around Taupō volcano, such as the absence of volcanic gas monitoring at the Wairakei/Tauhara geothermal fields (Potter et al., 2015), we are forced to consider Taupō volcano as a geographic area. This means that any volcanic phenomena near Lake Taupō, even if it may have originated within the NE-Dome system, is considered as unrest of Taupō volcano.

The modern system at Taupō volcano and the focus of current magmatism is considered to be in the area of the caldera collapse structure of the Taupō eruption (Barker et al., 2020). This is supported by the locations of the majority of the young Holocene vents under the eastern side of Lake Taupō, as well as a modern 200 MW geothermal system on the lake floor (Horomatangi Reefs) (Bibby et al., 1995; De Ronde et al., 2002). This area has also experienced clusters of seismicity associated with unrest episodes with potential magmatic origins (Barker et al., 2020).

## 2. Background

We first provide the rationale used to define unrest at Taupō volcano, followed by the origins of unrest and how this information is incorporated into the Bayesian Event Tree for Eruption Forecasting (BET_EF) (Marzocchi et al., 2004; Marzocchi et al., 2008).

### 2.1. Volcanic unrest and eruption

Volcanic unrest encompasses a variety of processes, including magmatic and non-magmatic processes. Magmatic unrest can be classed as “magma on the move”, particularly when magma is making a new pathway by breaking rocks or reusing an old pathway to intrude mush. Non-magmatic unrest is caused by either tectonic or hydrothermal processes without moving magma (Newhall and Dzurisin, 1988; Rouwet et al., 2014).

Magmatic eruptions involve the rise of magma toward the surface, producing juvenile magmatic ejecta. Other eruptions, such as phreatic and hydrothermal eruptions, do not directly involve magma. Phreatic eruptions are steam-driven explosions generated by water interacting with the heat from magma (Barberi et al., 1992; Stix and de Moor, 2018); Hydrothermal eruptions eject some solid material, whose energy derives solely from phase changes and heat loss in a convecting hot water or steam-dominated hydrothermal system (Browne and Lawless, 2001).

Tectonic unrest is usually due to regional strain, which would induce more stress into an area. Volcanoes in a delicate equilibrium can get disturbed by major tectonic earthquakes, inducing a state of volcanic unrest or even eruption (Manga and Brodsky, 2006; Bebbington and Marzocchi, 2011; Seropian et al., 2021; White and McCausland, 2016). The relationship between magma systems and tectonics can be very intertwined, as changes in the magmatic system may activate regional/local crustal faults and fractures. Conversely, tectonic stresses may trigger changes in the magmatic system (Sparks and Cashman, 2017).

Hydrothermal unrest is caused by magmatic fluids and heat interacting with country rock and sub-surface hydrologic systems such as groundwater (Newhall and Dzurisin, 1988). Both Newhall and Dzurisin (1988) and Acocella et al. (2015) found that felsic calderas with established hydrothermal systems tend to have higher rates of unrest. Acocella et al. (2015) found that hydrothermal systems can also buffer small mafic intrusions, explaining the higher rates of non-eruptive unrest at calderas with established hydrothermal systems (Newhall and Dzurisin, 1988). The interaction of heat and groundwater can cause changes in pore water pressure, which can cause uplift, tremors, long-period earthquakes, and phreatic eruptions, which are all behaviours of unrest (Newhall et al., 2001; Acocella et al., 2015). Overall, most volcanic unrest occurs in response to volcanic processes and/or regional tectonics that cause magma to interact with pre-existing rock and sub-surface fluids such as groundwater (Newhall and Dzurisin, 1988; Potter et al., 2012; Seropian et al., 2021).

Felsic calderas have been known to undergo multiple non-eruptive unrest episodes due to behaviours in the magmatic system. Behaviours such as magma intrusion from deeper feeder systems and magma migration through dykes and sills can both promote unrest through convection in the magma chambers (Pritchard et al., 2019). It has been suggested that mafic intrusions that recharge larger mush systems can be buffered, and repeated magma intrusions can explain multiple non-eruptive unrest episodes (Acocella et al., 2015). Repeated intrusions may slowly accumulate and increase the eruptability of the magma reservoir (Acocella et al., 2015; Sparks and Cashman, 2017; Sandri et al., 2017). Kennedy et al. (2018) reviewed developments in magma storage and transport concepts, observing that magma migration at calderas may involve considerable lateral transport through dykes or sills to a site of eruption.

There are recorded timescales at a silicic caldera of eruptive magma movement on the order of one meter per second, with an implied transit time from 5 km depth to the near-surface of about four hours, during the 2008 Chaitén Eruption (Castro and Dingwell, 2009). There are only a handful of eruptive volcanic unrest episodes recorded at felsic volcanoes, and most are small-scale, open vent eruptions (Acocella et al., 2015). Some notable eruptions from felsic calderas include the 1538 Monte Nuovo eruption from Campi Flegrei, Italy, the 1886 Tarawera eruption from Mt Tarawera in the Okataina Volcanic Zone, Aotearoa-New Zealand, and the 1994 eruption at Rabaul caldera, Papua New Guinea. While the 1886 Tarawera eruption was a basaltic eruption, Mt Tarawera has a co-located history of rhyolitic eruptions, including the second-youngest eruption in 1314 AD (Sahetapy-Engel et al., 2014).

### 2.2. Unrest signals

Volcanoes may exhibit different types of precursory signals, such as changes in the seismic activity, ground deformation and/or degassing pattern (Newhall and Dzurisin, 1988; Phillipson et al., 2013; Biggs et al., 2014). Distinguishing between magmatic, hydrothermal, and tectonic unrest is difficult, as all processes share similar monitoring signals (Constantinescu et al., 2016; Pritchard et al., 2019).

Seismicity typically evolves during magmatic unrest, from volcanic-tectonic (VT) earthquakes to shallow low frequency (LF) or hybrid earthquakes as magma shallows and gases exsolve. From there,

seismicity can evolve further into tremor (Pallister and McNutt, 2015). Previous studies (e.g., Newhall and Dzurisin, 1988; Acocella et al., 2015; Sandri et al., 2017; Gottsmann et al., 2019) found that eruptive unrest is characterised by rapid changes of the magmatic system, such as an increase in the number of earthquakes with shallowing hypocenters as magma rises through the crust. On the other hand, Newhall and Dzurisin (1988) and Gottsmann et al. (2019) also found that sudden (hours/days) declines in seismicity, after previously increasing seismic energy release, can also be a sign of an impending eruption (e.g., Endo et al., 1996).

During periods of volcanic unrest, magmatic and/or hydrothermal fluids can move within the crust, which can cause pressure changes in the surrounding rock, producing uplift (Acocella et al., 2015). Pritchard et al. (2019) found that deformation sources of more than 10 km deep are likely to indicate magmatic unrest, as hydrothermal systems are usually shallower than 10 km. However, most magma storage zones are also shallower than 10 km, somewhat reducing the utility of the findings. Newhall and Dzurisin (1988) found that fissuring and acceleration of uplift are two strong indicators of eruption, but only when observed with other unrest behaviours.

Sandri et al. (2017) and Gottsmann et al. (2019) noted that sudden increases in $SO_2$ and $CO_2$ can indicate an imminent eruption. The ratio of $CO_2$ and $SO_2$ is a common parameter in eruption forecasting, as increases in $SO_2$ are usually associated with shallow degassing (Oskarsson, 1984; Lee et al., 2018). However, Newhall and Dzurisin (1988) concluded that geochemical changes are not strong indicators of impending eruptions compared to seismicity and ground deformation.

Monitoring signals from previous rhyolitic caldera eruptions are almost non-existent, as there have only been two eruptions of this type in modern history: the 1912 Novarupta eruption in Alaska, USA (Adams et al., 2006) and the 2008 Chaitén eruption in Chile (Carn et al., 2009). The latter is the only one that had recorded pre-eruptive (although only observed in retrospect) unrest signatures. There was no scientific monitoring on Chaitén before the 2008 eruption, and unrest preceding the eruption was not noted (Carn et al., 2009). Retrospective analysis revealed precursory seismic activity detected by instruments up to 300 km away (Basualto et al., 2008). Earthquakes were felt only 24 h before eruption in Chaitén township (Castro and Dingwell, 2009).

### 2.3. Bayesian event trees

In the last few decades, eruption forecasts have been evolving (Sparks, 2003; Whitehead and Bebbington, 2021) from empirical pattern recognition towards forecasting based on models of underlying magma dynamics (e.g., dyke propagation Soosalu et al., 2005). Statistical and physical models are used to recognise geologic and monitored data patterns to inform an eruption forecast (Poland and Anderson, 2020). Event trees (e.g., Newhall and Hoblitt, 2002; Connor et al., 2001) are one of the main probabilistic tools used in volcanology today (Poland and Anderson, 2020), and have been applied in real eruption crises by the USGS VDAP (Volcano Disaster Assistance Program, (e.g., Newhall and Pallister, 2015), and in hindcasting (e.g., Sandri et al., 2009; Tierz et al., 2017; Tierz et al., 2020), and simulation exercises (e.g., Lindsay et al., 2010; Marzocchi et al., 2008). The structure of event trees emphasises the inherently probabilistic nature of volcanic systems, containing multiple possible outcomes (Newhall and Hoblitt, 2002). They thus provide a means to visualise complex volcanic processes and assign them probabilities. Event trees can assist scientists to summarise in a systematic and traceable manner key information needed by decision-makers to make informed decisions like evacuations (Marzocchi and Woo, 2007; Marzocchi and Woo, 2009; Newhall and Pallister, 2015). The flexible nature of event trees means that they can often be deployed readily in a crisis, but only if there is adequate volcanic monitoring (Newhall and Pallister, 2015).

A further development of event trees involved adding prior information (i.e., theoretical models, geologic data, and monitoring observations) using Bayes' theorem, creating a Bayesian Event Tree (BET) (Marzocchi et al., 2004; Marzocchi et al., 2008). In essence, BET does not rule out any possibilities, although some branches are not developed in detail, such as non-magmatic unrest, but instead shapes the probability distribution of the event depending on the available data, and the associated uncertainties (i.e., aleatoric and epistemic) (Marzocchi et al., 2008).

An event tree is essentially a branching graphical representation of events from a general starting state (i.e., the onset of unrest), which evolve into increasingly specific subsequent events (i.e., the presence of magma, the onset of eruption). The points where new branches are created are referred to as *nodes* (Newhall and Hoblitt, 2002; Marzocchi et al., 2004; Marzocchi et al., 2008). In BET_EF the nodes are as follows:

**Node 1:** there is either unrest, or no unrest, in the time interval $(t_0, t_0 + \tau)$, where $t_0$ is the present time, and $\tau$ is the forecast time window considered;

**Node 2:** the unrest is due to magma, or due to other causes not directly involving magma (e.g., hydrothermal or tectonic activity), given that unrest is detected;

**Node 3:** the magma will reach the surface (i.e., it will erupt), or it will not erupt, in the time interval $(t_0, t_0 + \tau)$, provided that the unrest has a magmatic origin

**Node 4:** the eruption will occur in a specific location, provided that there is an eruption;

**Node 5:** the eruption will be of a certain *size/style* (e.g., VEI), provided that there is an eruption in a certain location.

The probability of eruption is calculated by multiplying the probabilities at each node. Using a simplified formula the probability of eruption is given by:

$$P(E) = P(U) \times P(M|U) \times P(E|M) \quad (1)$$

where $U$ denotes 'unrest', $M$ 'magmatic unrest', $E$ 'eruption' and $P(E|M)$ $(= P(E|M, U)$ in BET_EF Marzocchi et al., 2008) is the probability of eruption given magmatic unrest.

The BET_EF framework has two components - *monitoring* and *non-monitoring*. Both components comprise a *prior distribution* and a *likelihood*. For the non-monitoring component, the prior distribution at each node describes general knowledge about that node. Generally, the prior distribution represents a 'best guess' probability, estimated without the data. This ' best guess' probability is assigned an 'equivalent number of data' ($\Lambda$) that weights the probability, reflecting our confidence in the estimate. The lower the weight, the less reliable the best-guess probability. Both the 'best guess' and 'equivalent number of data' are transformed by the code into the parameters of a Beta distribution that is used during calculations (Marzocchi et al., 2008 and Supplementary material).

Any change in the monitoring data updates the forecast, depending on the occurrence of relevant anomalies in volcanic activity. Within BET_EF, monitoring parameters can have either fuzzy or Boolean thresholds. Boolean parameters are binary; either they are anomalous or not. Fuzzy parameters are assigned upper and lower thresholds, generally by expert judgement. These upper and lower thresholds essentially define when a monitoring measure can be considered 'anomalous'. As setting parameter thresholds is also subjective, fuzzy logic is applied in the BET_EF model, which emulates the expert-like flexible judgement of the anomalous state of a monitoring parameter (Marzocchi et al., 2008).

In volcano monitoring, the use of one parameter is limiting. Volcanic systems can move to an anomalous state gradually, expressed through a wide range of behaviours involving different classes of monitoring. Hence weights may be assigned to monitoring parameters in Nodes 2 and 3. Weightings of 2 imply the parameter is a strong indicator for the node, and in calculations is the equivalent of two parameters with weight equal to 1. No weights are assigned to parameters in Node 1, as any positive result from one parameter indicates unrest, regardless of its

weight. Thus $P(U)$ in Eq. (1) is either 0 or 1. In the event of $P(U) = 0$, the probability of eruption is the long-term probability derived only from the eruption record. Hence we will use 'probability of eruption' as a shorthand for 'probability of eruption given unrest'.

At Nodes 2 and 3 the total (weight of) anomalous parameters $Z$ is converted into probability via the equation:

$$P = 1 - a\ \exp(-bZ) \tag{2}$$

Where $P$ is the conditional probability $P(M|U)$ (Node 2) or $P(E|M)$ (Node 3) and the parameters $a$ and $b$ have prior distributions $a \sim U(0.5,1)$ and $b \sim U(0,2)$. These priors can then be updated on the basis of, for example at Node 2, the number of observed unrest episodes that include magmatic unrest.

The forecast time window in which the probability of an event (i.e., magmatic unrest or eruption given magmatic unrest) is estimated is, to an extent, arbitrary. However, it must be set before considering the monitoring parameters, as these should be consistent with the window length chosen. Here we used the time window length of 30 days. Therefore, the probability of entering magmatic unrest that can potentially lead to an eruption is considered for the next 30 days, unless updated by a change in monitoring signals. The 'look back' period is the time horizon from which a particular monitoring data stream is considered informative for the forecast time window. This can vary according to the data concerned.

A BET_EF has been applied to several volcanoes around the world in simulated scenarios and retrospective studies, including a silicic caldera, stratovolcanoes, and a volcanic field (Marzocchi et al., 2008; Sandri et al., 2009; Lindsay et al., 2010; Sandri et al., 2012; Brancato et al., 2011; Brancato et al., 2012; Selva et al., 2012; Constantinescu et al., 2015; Tonini et al., 2016; Rinawati et al., 2018). In Aotearoa-New Zealand, BET_EF has only been used once, during emergency management exercise Rūaumoko for the Auckland Volcanic Field (AVF) (Lindsay et al., 2010). There is a general need to develop more probabilistic forecasting tools for Aotearoa-New Zealand's volcanoes especially during non-crisis periods, which then can be used during a crisis in an objective manner (e.g., Whitehead and Bebbington, 2021). Short-term probabilistic forecasting models are particularly needed at silicic calderas like Taupō volcano, as eruptions from these calderas are likely to pose a severe risk, and evacuations may be needed. Probabilistic forecasting models set up in times of repose, that are established through a consensus within the scientific community, are a good way to justify the probability output of forecasting models during and after volcanic crisis (Lindsay et al., 2010; Newhall and Pallister, 2015; Whitehead and Bebbington, 2021).

In the key literature summarising and analysing caldera unrest (Newhall and Dzurisin, 1988; Acocella et al., 2015; Sandri et al., 2017; Gottsmann et al., 2019) phreatic eruptions are consistently included within the 'eruption' term. However, as we are using BET_EF, we will only estimate the probability of magmatic eruptions, for which we treat phreatic activity is a possible precursor. The impacts of those events that never progress beyond phreatic activity (Newhall and Dzurisin, 1988) can be included through use of the BET_UNREST model (Tonini et al., 2016; Constantinescu et al., 2016), but this was considered a step too far for us to elicit information about.

For the Taupō volcano BET_EF model, only Nodes 1 to 3 were considered. We also only list median probability estimates. The BET_EF code also produces uncertainty bounds for these probabilities.

### 2.4. Taupō volcano unrest history

Taupō volcano has undergone 17 recorded periods of unrest in the 142 years between 1877 and 2019 (Potter et al., 2015; Illsley-Kemp et al., 2021). Potter et al. (2015) assembled a catalogue of all past unrest activity at Taupō volcano (excluding the 2019 unrest period), the dates and duration of which are listed in Table 1.

**Table 1**
Table of all unrest episodes from Taupō volcano from 1878 to 2019. Adapted from Potter et al. (2015).

| | Duration of Unrest (days) |
|---|---|
| **1.** 8 Apr 1877–23 Apr 1878 | 381 |
| **2.** 27 June–24 Sept 1880 | 90 |
| **3.** May–2 Sept 1895 | 125 |
| **4.** 20 March–17 Oct 1897 | 212 |
| **5.** 10 May 1922– Jan 1923 | 251 |
| **6.** Apr–Sept 1961 | 168 |
| **7.** 3–4 Mar 1964 | 2 |
| **8.** 7 Dec 1964–29 Jan 1965 | 54 |
| **9.** Dec 1967–30 Jan 1968 | 46 |
| **10.** Mar–14 Apr 1974 | 45 |
| **11.** 14–23 Feb 1975 | 10 |
| **12.** 30 Dec 1975 | 1 |
| **13.** 1 Feb 1983–6 Mar 1974 | 400 |
| **14.** Mar 1996– Mar 1999 | 1126 |
| **15.** Dec 1999–Jan 2001 | 549 |
| **16.** 30 Mar 2008–9 Feb 2010 | 682 |
| **17.** 17 Feb–9 Sept 2019 (deformation started Oct 2018) | 203 |

Here we break the history of Taupō volcano into two time periods; Non-instrumental (before 1985) and instrumental (1985 to present), based on a change in how GNS Science and later the GeoNet system detected and located earthquakes (Barker et al., 2020). Potter et al. (2015) identified non-instrumental unrest episodes from information in newspapers, scientific literature, circulars, reconnaissance reports, correspondence, felt reports, sparse earthquake catalogues (from the 1960's - 1980's), and workshop material. Hence the large uncertainty in the estimated duration of the non-instrumental unrest episodes. Lake levelling began in 1979 to measure deformation using Lake Taupō as a tilt meter (Otway, 1989; Otway et al., 2002).

Fig. 2 shows the comparison between the instrumental and non-instrumental unrest durations. A non-parametric Mann Whitney U test, for the difference in medians, between the durations of non-instrumental and instrumental unrest episodes gives a P-value of 0.015. This implies that the instrumental unrest episode durations come from a different distribution (they tend to be significantly longer) than the non-instrumental unrest episode durations at significance level $\alpha = 0.05$. This reflects the fact that monitoring equipment can detect the initial and subtle changes, such as unfelt earthquakes, that may go unobserved to humans. There is no significant relationship between quiescence length (between unrest episodes) and duration of unrest episodes.

#### 2.4.1. Non-instrumental unrest episodes (pre-1985)

Significant unrest episodes from the non-instrumental period include 1895 (Episode 3), 1922 (Episode 5), 1964–1965 (Episode 8), and 1983–1984 (Episode 13). The 1895 unrest episode is one of the most notable unrest episodes to have occurred at Taupō volcano, with societal effects such as self-evacuations, effects on tourism, and psychological impacts on residents such as loss of sleep and rumour (Potter et al., 2015). A sizeable M6-7.5 earthquake caused damage to Taupō township and was large enough to be felt in Wellington. Even though this event is thought to be tectonic in nature, an earthquake of this magnitude would have inevitably interacted with the Taupō magmatic system directly or through ground shaking (Barker et al., 2020), and could have triggered an eruption.

The 1922 unrest episode was one of the most severe and the most notable episode with intense seismic activity and dramatic ground deformation. The seismicity caused damage in Taupō township, public alarm, and self-evacuations, which impacted tourism locally, regionally, and nationally. Professor Marsden (Victoria College, Wellington) stated that "the chances of blow-up are one in six". Seismicity was accompanied by up to 3 m of subsidence at the northern lake shoreline, with 0.5 m displacement on the Kaiapo fault and displacement in the Whakaipo

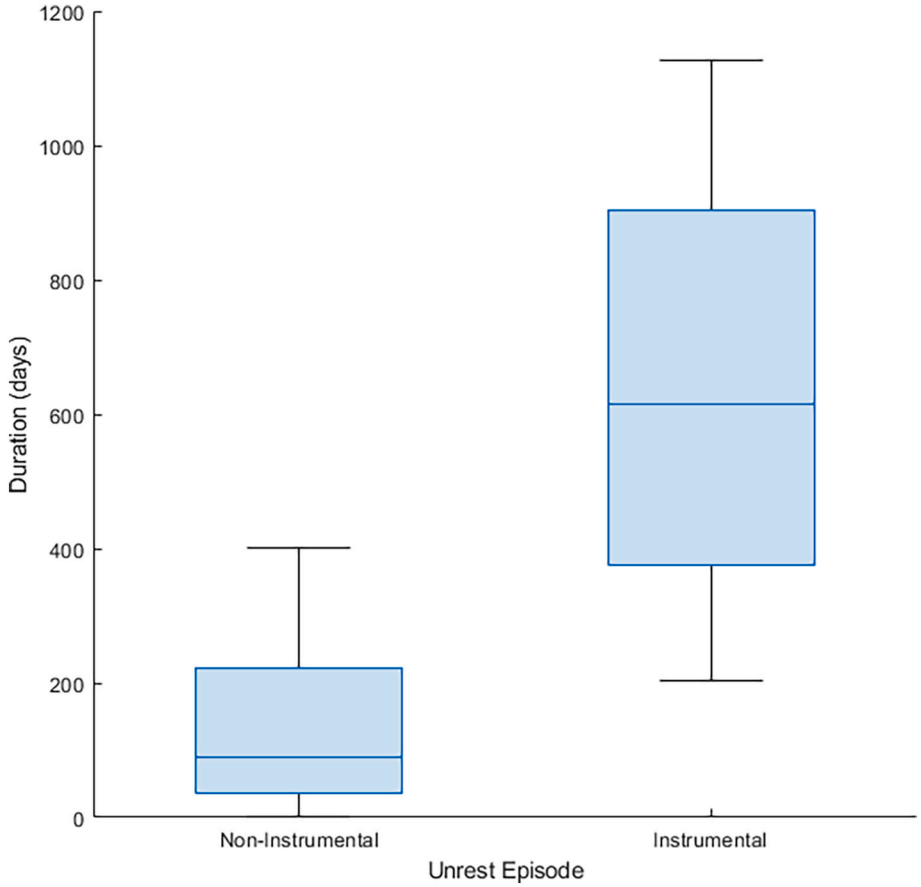



**Fig. 2.** Box Plot showing non-instrumental unrest durations versus instrumental unrest durations.

and Whangamata faults (Johnston et al., 2002; Barker et al., 2020) (Fig. 1). Although there is no clear evidence that this unrest episode was associated with the magmatic system, the relatively small magnitude main-shocks could not have induced such dramatic fault displacement (Barker et al., 2020), therefore, we treated this unrest episode as volcanic in nature.

In 1964–1965, many earthquakes were clustered around the Western Bay area; However, this area has been almost entirely aseismic in the Instrumental era, and thus there is uncertainty about the location of the swarm, as well as whether the seismic activity was directly related to magmatic activity (Barker et al., 2020). In the 1984–1985 unrest episode, lake levelling measurements showed that Kinloch was uplifted by 53 mm between October 1982 and June 1983 (Fig. 1). On 23 June 1983, the Kaiapo fault ruptured, causing 43 mm of subsidence to the west of the fault, with a 1.2 km fault scarp. Seismicity during this time migrated towards Horomatangi Reefs and Motutaiko Island, similarly to the 2008 and 2019 unrest episodes, which suggests that the unrest was likely magmatic (Barker et al., 2020).

### 2.4.2. *Instrumental unrest episodes (post-1985)*

Before the 1999 unrest episode, between 1985 and 1990, earthquake activity was mainly focused in the south of Taupō, near Motuoapa and Tokaanu, and was primarily rift-related faulting (Barker et al., 2020) (Fig. 1). In March 1987, a spatially defined earthquake swarm near Karangahape Cliffs was preceded by uplift at the shoreline (Otway and Sherburn, 1994; Barker et al., 2020) (Fig. 1). The swarm was ascribed to the poorly constrained Waihi Fault (Sherburn, 1992) (Fig. 1). Earthquake activity since this time, combined with the observed deformation and uplift, suggests it may have been influenced by the Taupō magmatic system (Barker et al., 2020). Despite this, this event was not included in the unrest catalogue for Taupō volcano, as it failed to meet volcanic unrest classification thresholds created by Potter et al. (2015).

During Unrest Episode 14 (March 1996 - March 1999) from 1997 to 1998 there was an increase in seismic activity, preceded by uplift at Horomatangi Reefs. The seismic activity was focused mainly between Karangahape Cliffs and Motutaiko Island, a pattern that is shared with later swarms in 2008 and 2019 (Barker et al., 2020) (Fig. 1).

In 2001 during Unrest Episode 15, there was significant earthquake activity to the north of Lake Taupō. Peltier et al. (2009) interpreted this as a slip on the Kaiapo fault (Fig. 1). However, no rupture was observed at the surface, and the earthquakes are not located accurately enough to definitively ascribe them to specific faults, although they were likely tectonic in nature (Barker et al., 2020). During the 2008–2010 unrest episode, there was significant uplift of between 40 and 50 mm observed at Horomatangi Reefs (Jolly et al., 2008) (Fig. 1). The unrest episode was thought to have been triggered by a tectonic slow-slip event, resulting in fluid-driven inflation (Fournier et al., 2013). However, it has not been analysed in detail (Barker et al., 2020).

The 2019 unrest episode is the most recent unrest episode to have occurred at Taupō volcano. This episode was characterised by earthquake activity that occurred in seven distinct swarms denoted swarm A to swarm G), each with spatially distinct clusters, from the 17$^{th}$ of February to the 6$^{th}$ of September 2019 (Illsley-Kemp et al., 2021). All seven earthquake swarms showed evidence of non-double-couple earthquakes i.e., earthquake mechanisms (opening and closing) different from that expected for slip on a fault plane. Ground deformation was also identified before and during the 2019 unrest episode. The first period of ground deformation started before the earthquake swarms, from October 2018 through to the beginning of swarm A. Horizontal deformation occurred at Horomatangi Reefs (6 mm) and Ouaha Ridge (4 mm). The second period of ground deformation occurred during the earthquake swarms, where 10 and 7 mm was

observed at Horomatangi Reefs and Ouaha Ride respectively, accompanied by 6 mm of vertical uplift at Horomatangi Reefs (Fig. 1). It has been suggested by Illsley-Kemp et al. (2021) that an inflating magma body caused the observed deformation, approximately 1 km NW of Horomatangi Reefs. This was considered to be due to a recharge event, where melt was extracted from the deep crustal mush into the shallow crustal mush. This melt was subsequently pushed outwards and inserted into a sill in the country rock. The seismic data outlines an aseismic zone inferred to be in the area of shallow crustal mush. However, there must be a melt fraction of $> 20\%$ for this area to be aseismic and inhibit seismic activity. The 2019 unrest episode was volcanic in nature and origin and is the only unrest episode to be positively identified as such.

It is important to note that both currently and historically, degassing or other geochemical signatures have not been actively monitored at Taupō volcano. The presence of the lake inhibits any degassing or geochemical measurements being taken from Horomatangi Reefs.

## 3. Methodology

In the absence of data, such as monitored eruptions, the key inputs in probabilistic models such as BET_EF are usually obtained via expert opinion (Aspinall and Cooke, 1998; Marzocchi et al., 2008; Marzocchi et al., 2012; Selva et al., 2012; Newhall and Pallister, 2015). Expert opinion can be elicited through a structured questioning process known as expert elicitation (Cooke, 1991; Colson and Cooke, 2018). Expert opinion is often needed in eruption forecasting due to the lack of appropriate or available information (Rowe and Wright, 2001; Marzocchi et al., 2012). It is also generally agreed that information obtained from a group of experts tends to be more coherent than individual experts (Marzocchi and Bebbington, 2012; Marzocchi et al., 2012). In the BET_EF model, expert elicitation requires experts to be directly queried about anomalies, during times of repose, rather than being asked to directly evaluate probabilities (Marzocchi et al., 2008). There are two main advantages to getting an expert opinion before a crises: 1) there is time to set rules and establish a consensus over them, reducing stress during an emergency when cognition can be biased by intense pressure; and 2) when a crisis does arise, forecasts can be made quickly and without debate, strictly as a function of the monitoring measures (Lindsay et al., 2010; Marzocchi et al., 2012).

### 3.1. BET_EF workshop

As Taupō volcano has no previous pre-eruptive monitored observations, an expert elicitation workshop was necessary to develop the BET_EF model. Invitations were issued to all appropriate Aotearoa-New Zealand volcano scientists, who were provided the information in the Background Section above through pre-workshop Zoom sessions and in hardcopy. Prior to the elicitation workshop, two to three experts (primarily monitoring) were pre-selected to form three panels (seismicity, geochemistry, or geodesy). These panel members were asked to create

**Table 2**
Table of seismic parameters defined by panel members versus the refined parameters at the end of the workshop.

| | Parameters - PANEL MEMBER (PM) INITIAL | Threshold | Weight | Parameters - END OF WORKSHOP | Threshold | Weight |
|---|---|---|---|---|---|---|
| **Node 1** | (PM 1) - Moderate magnitude earthquake(s) (>mag. 4) in or around the caldera | YES/NO | | 1. Moderate magnitude earthquake(s) (largest in the last month) (>mag. 4) in or around the caldera (shallower than 30 km) | >4–4.5 | |
| | (PM 1) - Increased rate of low magnitude (<mag. 4) earthquakes in or around the caldera | >100 per month | | 2. Increased rate of low magnitude earthquakes in or around the caldera {(shallower than 30 km, within last month) (>mag. 2) | >50–100 per month | |
| | (PM 2) - Many earthquakes N > x | | | 3. LP/VLP/Tremor (per month) | YES/NO | |
| | (PM 2) - Larger earthquake with magnitude M > y (where y is what we see all the time in the last 20 years) | | | | | |
| **Node 2** | (PM 1) - Non-double-couple earthquakes in or around the caldera | YES/NO | 2 | 1. Non-double-couple earthquakes (within last month) | YES/NO | 2 |
| | (PM 1) - Migrating earthquake hypocentres (vertically or laterally) | YES/NO | 1 | 2. Migrating earthquake hypocentres (vertically or laterally) (within last month) | YES/NO | 1 |
| | (PM 1) - LF earthquakes/tremor within mush zone | YES/NO | 2 | 3. LP/VLP/Tremor (within last month) | >1–10 | 1 |
| | (PM 2) - Earthquake swarm (from b-values or Omori decay) | YES/NO | | 4. Amplitude of largest tremor | X | X |
| | (PM 2) - LP/VLP/Tremor (from duration and f-content, non-double-couple) | YES/NO | | 5. Number of earthquakes within 1.5 magnitude of largest in swarm within one month (from b-values or Omori decay) (as opposed to mainshock/aftershock sequence) | >50–100 | 1 |
| | (PM 2) - Increasing rate of number of events | | | 6. Earthquakes (<30 km depth) in last day | >5–100 | 1 |
| **Node 3** | (PM 1) - Large earthquake(s) (>mag. 6) in or around the caldera | YES/NO | 1 | 1. Magnitude of largest earthquake in last month | >5.5–6.5 | 1 |
| | (PM 1) - Vertically migrating earthquakes within the shallow (<5 km depth) in or around the caldera | YES/NO | 2 | 2. Resolvable vertically migrating earthquakes within the shallow (<5 km depth) | YES/NO | 1 |
| | (PM 2) - Swarm or tremor is longer than z days | | | 3. Earthquakes (<30 km depth) in last day | >100–1000 | 1 |
| | (PM 2) - Upward movement of hypocentres | | | 4. Duration of current sustained tremor/ continuous saturation of earthquakes (hours) | >0.1–24 | 2 |
| | (PM 2) - Sudden change or rate of seismicity (sudden quiescence or esculation) | | | 5. Sudden change or rate of seismicity (sudden quiescence) | YES/NO | 1 |

their own initial set of parameters, thresholds, and weights. The purpose of this was to establish a somewhat crude set of parameters that would jump-start conversations within the expert elicitation workshop. Panel members were referred to past BET_EF papers, in particular that on Campi Flegrei (Selva et al., 2012) for example parameters, thresholds, and weights.

The Taupō volcano expert elicitation workshop started with presentations from all the panel members, one discipline at a time. After all the panel members from each discipline had explained their initial parameters, thresholds, and weights at the first BET_EF node, facilitated discussion occurred involving the rest of the workshop participants to refine the parameters presented by the panel members until a consensus was reached. This step was repeated to cover Node 2 and Node 3. An example of the refined parameters versus the panel member initial parameters can be found in Table 2. Words underlined represent the key differences and changes between panel member parameters and the parameters refined by the end of the workshop. The final stage of the workshop was a cross-discipline discussion to 'harmonise' parameters for all three disciplines. The probability of a new vent opening in a specific location (Node 4) within or outside the caldera is a complex issue. While a prior spatial likelihood was elicited by Bebbington et al. (2018), it is unknown to what extent seismicity might be used to update this prior during unrest, especially given the scale and composition of the mush system. Node 5 (eruption style/size) has also been explored (Bebbington, 2020), and again it is unclear how monitoring signals might be used to update this (Sandri et al., 2004; Lindsay et al., 2010).

The probability of eruption at Taupō volcano in the next time window (30 days) is (re) calculated when there is a change in the monitoring data, resetting the forecast window to the next 30 days. However, the probability is known not to be uniformly distributed across the forecast window (Lindsay et al., 2010; Bebbington and Zitikis, 2016). A further problem arises when trying to compare magmatic unrest and eruption probabilities for instrumental and non-instrumental unrest episodes as they have different sensitivities. So, in order to facilitate comparisons between the data-poor pre-instrumental and data-rich instrumental episodes, we represent the probability of eruption from an unrest episode as the peak monthly probability. We note that this is our best estimate in the case of the earlier simplest episodes, and represents a reasonable approximation for the more complex, as the abstraction works in opposite directions (we use a higher than average probability, but a shorter length of unrest). This also ameliorates somewhat the influence of the 'grainy' nature of the estimated probabilities, particularly in the pre-instrumental data, and the heterogeneity in Fig. 2, which seems to stem from a definition of unrest which cannot be disentangled from monitoring completeness.

As worded in the Taupō volcano BET_EF model, all seismic parameters based on counts have a condition that only earthquakes more than Magnitude 2 should be considered should the seismic monitoring network be enhanced by additional sensors. Otherwise the additional sensors would inflate the counts with the addition of smaller events which are assumed to be there anyway.

BET_EF has typically been developed at volcanoes where there are few, or no, observed unrest episodes. Hence the $a$ and $b$ prior distributions in Eq. (2) are usually not updated due to a lack of data to provide a posterior. Tonini et al. (2016) actually use fixed scalar values in an attempt to decrease the calculated probabilities for well-monitored volcanoes. The relative wealth of Taupō volcano unrest data available to us presents a unique opportunity to examine the effect in the BET_EF of the gradual accumulation of (non-eruptive) unrest episodes. Hence, the Taupō BET_EF model will be used to run through each historic unrest episode as if calibrated to their respective occurrence time - Suite 1. In this suite, in 1922 there would have been four unrest episodes without eruption to update the priors). However, in order to validate the BET_EF, we need to calculate how likely it is that we would have seen this many unrest episodes without an eruption, and to do this we need to 'renormalise' each estimated probability of eruption to the baseline of 17 observed non-eruptive unrests. Hence we will also consider each unrest episode as if it were to occur at the time of writing this paper (in 2021) (i.e., we use 17 unrest episodes with no eruption between 1877 and 2019 to update the priors). We refer to these as the "2021 Calibrated BET_EF" model, or Suite 2.

## 4. Results

We will now describe the final elicited BET_EF settings for Taupō Volcano, and then run each of the historical unrest episodes through the resulting lens.

### 4.1. Taupō volcano BET_EF settings

The definitions of background behaviour and unrest are inescapably specific to each volcano and its monitoring network. For Taupō volcano these definitions also have a potentially high degree of observation bias given the relatively short time period during which the volcano has been monitored. The BET_EF code is set up by entering information on the past history (i.e., past data) of Taupō volcano, as well as a list of monitoring parameters and their thresholds expected in the event of magmatic unrest and an impending eruption. The results of the Taupō volcano expert elicitation workshop are summarised in Tables 3 and 4, and described in more detail in Appendix A. Table 3 refers to Suite 2 (the 2021 calibrated BET_EF model). When looking at Suite 1 (Retrospective BET_EF model), the prior probabilities remain the same, but the likelihood parameters (i.e., the length of the unrest catalogue, the number of unrest episodes, number of magmatic unrest episodes, and the number of observed eruptions) change for each eruption.

### 4.2. Unrest probabilities - retrospective BET_EF calibration (Suite 1)

Looking at the significant unrest episodes documented above, we can examine how the BET_EF performs retrospectively. The second identified unrest event at Taupō volcano occurred in 1895 and has the highest estimated probability of eruption (Fig. 3) among the non-instrumental unrest episodes at 51.7% (All probabilities are given as median estimates).

The 1922 unrest episode was so concerning that temporary seismometers were deployed. The estimated probability of eruption during the 1922 unrest episode peaked at 45%, exceeding the one in six estimate of Professor Marsden (Potter et al., 2015). As we can will see in Suite 2, this is due to the absence at the time of past data of unrest without eruption.

The BET_EF model did not recognise the 1880, 1964, and December 1978 unrest episodes as entering the unrest state, and consequently the long-term eruption probabilities were the default (<0.0001) per month estimates. However, this lack of recognition was on the basis of the pre-

**Table 3**
Background settings for a 2021 calibrated Taupō volcano BET_EF (Suite 2).

| Node | Prior Parameters | Likelihood Parameters |
|---|---|---|
| **Node 1: Unrest/No Unrest** | $\Theta_1 = 0.005$ | n1 = 1602 (length of unrest catalogue) |
| | $\Lambda_1 = 1$ (equivalent number of data) | y1 = 17 (episodes of unrest) |
| **Node 2: Magma/No Magma** | $\Theta_2 = 0.1$ | n2 = 17 (episodes of unrest) |
| | $\Lambda_2 = 1$ (equivalent number of data) | y2 = 2 (number of magmatic unrest) |
| **Node 3: Eruption/No Eruption** | $\Theta_3 = 0.1667$ | n3 = 2 (number of magmatic unrest) |
| | $\Lambda_3 = 1$ (equivalent number of data) | y3 = 0 (no observed eruptions) |

**Table 4**
Summary of the elicited monitoring BET_EF input information for Taupō volcano.

| | Parameters | Threshold | Weight |
|---|---|---|---|
| Node 1 | **1. Moderate magnitude EQ(s)** (largest in last month, >mag. 4) in or around the caldera (<30 km) | >4–4.5 month$^{-1}$ | – |
| | **2. Increased rate of low magnitude EQs** in or around the caldera (<30 km depth; >mag. 2) | >50–100 month$^{-1}$ | – |
| | **3. LP/VLP/Tremor** (last month) | YES/NO | – |
| | **4.** (Emission rate) **Appearance of gas emissions at surface AND/OR $CO_2/CH_4$ratios increasing from background** (~ 100) **AND/OR C/S ratio increasing from background** (>30) **AND/OR He & Ar isotopes changing toward mantle values** (>+1bg, <305) (last month) | YES/NO | – |
| | **5. Notable changes to temperature of any surface feature(s)** (last month) | YES/NO | – |
| | **6. Notable sustained surface deformation** (in three months) change as determined by geodesist | YES/NO | – |
| | **7. Long-term surface deformation (lake levelling)** subtracted from agreed multi-year trend (i.e., background) | >5–25 mm/yr | – |
| | **8. Notable deviation to geodesist in short-term deformation rate over multiple days** (horizontal and/or vertical), subtracted from agreed multi-year trend (i.e., background) | YES/NO | – |
| | **9. Visual change** - change in lake/spring/geothermal water appearance (last month) | YES/NO | – |
| Node 2 | **1. Non-Double-Couple EQ(s)** (last month) | YES/NO | 2 |
| | **2. Migrating EQ hypocentres** (vertically and/or laterally) (last month) | YES/NO | 1 |
| | **3. LP/VLP/Tremor** | >1–10 month$^{-1}$ | 1 |
| | **4. Amplitude of LP/VLP/Tremor is larger than co-located (nearby) earthquakes** (>mag. 2) (last month) | YES/NO | 1 |
| | **5. Number of EQs within 1.5 magnitude of largest in swarm** (from b-values or Omori decay) | >50–100 month$^{-1}$ | 1 |
| | **6. EQs** (<30 km, >mag. 2) | >5–100 day$^{-1}$ | 1 |
| | **7.** (Emission rate) **$CO_2/CH_4$ratios increasing** (>100) **AND/OR C/S ratios decreasing toward mantle values** (<30, controlled by depth of decoupling and scrubbing of S) **AND/OR & $^3He/^4He/^{40}Ar/^{36}Ar$shifting toward mantle values** (>5, <305) **AND/OR$d^{15}N$toward mantle value** | YES/NO | 2 |
| | **8. Temperature of any fumarole(s)/surface feature(s)** (> 100 °C) (last month) | YES/NO | 1 |
| | **9. Deformation** directly related to volume increase or pressure via modelling (inflation at depth) | YES/NO | 2 |
| | **10. Visual change** - low lake levels reported along nearby shoreline AND/OR increased bubbling in the lake AND/OR reports of unusual lake outflow/lake level changes in marinas (that have not been seen in decades) (last month) | YES/NO | 1 |
| Node 3 | **1. Magnitude of largest EQ** | >5.5–6.5 day$^{-1}$ | 1 |
| | **2.** Resolvable **vertically migrating earthquakes within the shallow crust** (<5 km depth) (last month) | YES/NO | 1 |
| | **3. EQs** (<30 km depth, >mag. 2) | >100–1000 day-1 | 1 |
| | **4. Duration of current sustained tremor/continuous saturation of EQs** | >0.1–24 h | 2 |
| | **5. Sudden change or rate of seismicity** (sudden quiescence) | YES/NO | 1 |
| | **6.** (Emission Rate) **$CO_2/CH_4$ratios at magmatic values** (>100,000) **AND/OR C/S ratios decreasing toward 'magmatic' values** (~5, **$SO_2$** emission increasing) **AND/OR$d^{15}N$at mantle values** (−3 to −5) **AND/OR He & Ar isotopes at mantle values** (7.5 & 295 respectively) (last month) | YES/NO | 2 |
| | **7. Temperature of any fumarole(s)/surface feature(s)** (>400 °C) (last month) | YES/NO | 1 |
| | **8. Acceleration in deformation rate** noted across the most recent data points | YES/NO | 2 |
| | **9. Migration on any 3D direction** of surface deformation (pattern consistent with magma migration) (last month) | YES/NO | 1 |
| | **10. Visual change** - phreatic activity AND/OR ground cracking AND/OR minor visual faulting (last month) | YES/NO | 1 |

instrumental 'monitoring network'. We consider it likely that the modern network would have generated enough additional data for these episodes to meet the Taupō volcano BET_EF definition of unrest.

For the instrumental unrest episodes, the median probability of magma being the source of unrest for each unrest episode varies drastically from 26.2% to 87.5%, between 1996–1999 and 2019 (Fig. 3). This difference in magmatic unrest probabilities principally stems from the parameter in Node 2 'Deformation directly related to volume increase or pressure via modelling'. The ground deformation observed in the 2008–2010 unrest episode did not activate this parameter, as the ground deformation is believed to have been be due to a slow slip event (Fournier et al., 2013).

### 4.3. Unrest probabilities - 2021 BET_EF calibration (Suite 2)

All the Taupō volcano unrest episodes were also run through a Taupō BET_EF model calibrated for June 2021, as shown in Fig. 4. Eruption probabilities differ between the 2021 calibrated BET_EF model (Suite 2) (Fig. 4) and and the retrospective calibrated BET_EF model (Suite 1) (Fig. 3). The 2021 calibrated BET_EF model (Suite 2) contains more observed outcome data that refines eruption probabilities (i.e., 17 unrest episodes and zero observed eruptions). In contrast, the retrospective BET_EF models (Suite 1) have progressively less past data as we go back in time (i.e., in 1922 there were only four unrest prior episodes with zero observed eruptions). For the same reason, the estimated probability of magmatic unrest for all the non-instrumental unrest episodes is lower in the 2021 calibrated Taupō BET_EF model compared to their retrospective Taupō BET_EF calibrations (Fig. 3).

For the 1922 unrest episode, the estimated probability of eruption (27.2%) is higher than the one in six eruption probability stated by Professor Marsden (Fig. 4). However, the monitoring network (or lack thereof) in 1922 is not comparable to the present one. If the 1922 unrest episode were to occur today, eruption probabilities would likely be different, as the seismic network has improved considerably over the last 100 years. This inhomogeneity in monitoring for the non-instrumental unrest events poses some constraints to resolving the probabilities.

## 5. Discussion

### 5.1. Unrest catalogues

The definition of unrest strongly affects the forecast probabilities

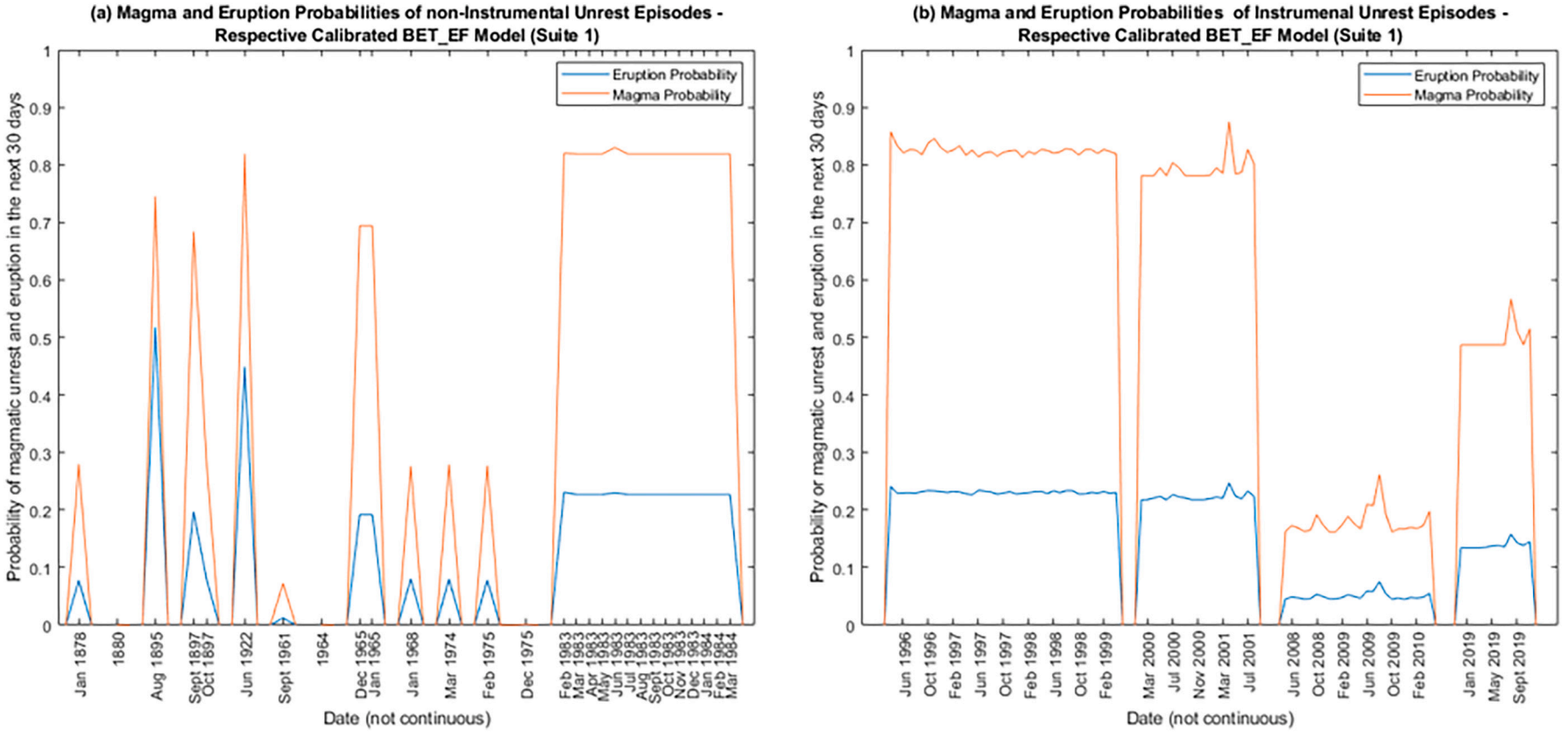



**Fig. 3.** Probabilities of magmatic unrest and eruption in the next 30 days from the Taupō volcano BET_EF calibrated to each unrest episode: (a) Non-Instrumental unrest and (b) Instrumental unrest. Note: date is not continuous, but rather represents each individual unrest episode.

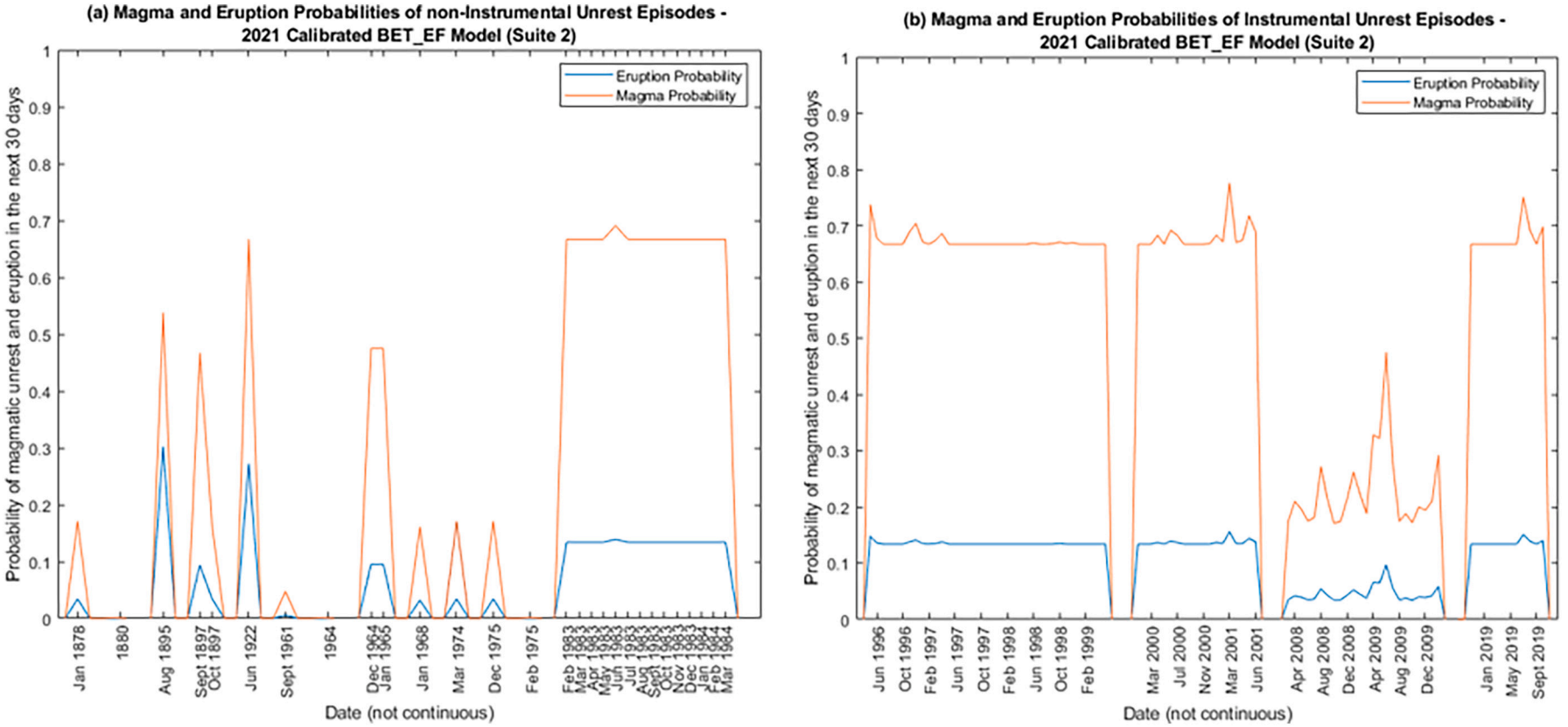



**Fig. 4.** Probabilities of magmatic unrest and eruption in the next 30 days, calibrated to a 2021 unrest episode: (a) Non-Instrumental unrest and (b) Instrumental unrest. Note: date is not continuous, but rather represents each individual unrest episode.

from BET_EF, and indeed from any short-term forecasting procedure. The design of BET_EF is deliberately conservative, with probabilities (Eq. (2)) considering only the number of anomalous parameters, with no reduction being made for monitored non-anomalous parameters. The definition of unrest will differ by volcano and monitoring system. Here we have defined unrest as relative to the 'quieter' level of activity over the past 140 years at Taupō.

Defining unrest is highly subjective, and must be in relation to the understood normal quiescent level of activity. In general, the best monitored and most active volcanoes have the best record of activity. However even, or especially, the best records are only a tiny fraction of the volcano's history, and the completeness and accuracy of any unrest catalogue decreases rapidly as we go back in time. Taupō is typical in this regard, with instrumental monitoring covering only a few decades of unrest, without any eruption. It is possible that 'unrest' might be a multi-level phenomenon, with low levels of activity (restlessness) lasting for centuries or longer, although the long-lived geothermal circulation system (Wilson and Rowland, 2016) argues that this is not necessarily the case at at Taupō.

### 5.2. Monitoring signals

The variety of possible monitoring signals can affect the interpretation of monitoring data. For example, geothermal exploration and the development of commercial geothermal fields at Wairakei/Tauhara may have been the cause of some of the unrest phenomena recorded at Taupō volcano. During the 1950s and 1960s, increased draw-off from production bores at the Wairakei Geothermal Fields caused hydrothermal eruptions and shallow earthquakes, which were not necessarily caused by magmatic processes (Thompson, 1960; Potter et al., 2015).

Other processes such as slow slip events, fault creep, and lake level changes (due to water inflow-outflow) can mimic or produce unrest behaviours that are not due to the movement of magma and may not interact with the magmatic system. This may mean that if these processes are (wrongly) identified to be magmatic in nature and activate monitoring parameters in the BET_EF model, the eruption probabilities may be overestimated. The expert elicitation meeting discussed this at length, and made every effort to ensure that the parameters in Nodes 2 and 3 are designed to be more indicative of magmatic unrest than other sources of unrest (i.e., tectonic and hydrothermal). However the elicited parameters may not account for all possible scenarios, which may particularly be the case for some of the non-instrumental unrest episodes.

An elaboration of the BET model, BET_UNREST, was developed by Sandri et al. (2017), specifically to include the forecasting of non-magmatic unrest. BET_UNREST differs from BET_EF in containing an additional branch to detail non-magmatic unrest outcomes. During the expert elicitation workshop, the issue of non-magmatic eruptions (i.e., phreatic and hydrothermal eruptions) was discussed, but no consensus was reached. It was decided that the BET_EF model would be more suitable option here as our aim is to forecast probabilities related to large magmatic events, rather than outcomes such as phreatic and hydrothermal eruptions.

### 5.3. Eruption probabilities

The estimated probabilities of magmatic unrest and eruption vary drastically between the differently calibrated BET_EF models. There are three reasons for these differences: 1) Prior data in BET_EF models, 2) Parameter weights, and 3) Monitoring data.

The thresholds elicited in the expert elicitation workshop were based upon the current (instrumental) monitoring set up at Taupō volcano. Because of this, the non-instrumental 'signals' (1877–1878 to 1996–1999) have to be translated to the current set up. For example, the lack of low magnitude seismicity in the non-instrumental period reduces the $Z$ score in Eq. (2). On the other hand, earlier unrest episodes have fewer past data, which raises their estimated eruption probability. This can be partially corrected for by running all 17 unrest episodes through a 2021 calibrated BET_EF model (Suite 2) - where it was found that eruption probabilities were lower (e.g., 1922 unrest episode went from 44.9% to 27.2%).

The magmatic unrest probabilities in particular, are sensitive to parameter weights. The ground deformation parameter in Node 2 (' Deformation directly related to volume increase or pressure') strongly influenced the eruption probabilities for the 1983–1984, 1996–1999, 1999–2001, and 2019 unrest episodes, due to its higher weight of 2. Even when daily seismic rates were low, magmatic unrest probabilities were still elevated due to this ground deformation node. For the 2008–2010 unrest episode, the estimated median eruption probabilities were lower (7.6% in 2008–2010 compared to 23%, 24.1%, 24.7% and 15.8% in 1983–1984, 1996–1999, 1999–2001, and 2019 respectively) than in other instrumental unrest episodes because the ground deformation node was not activated. Magmatic unrest probabilities for non-instrumental unrest episodes (excluding 1983–1984) were lower than instrumental unrest episodes as, by definition, any deformation could not be 'directly related' to volume increase or pressure.

Using the 2021 calibration allows us to compare eruption probabilities from instrumental and non-instrumental unrest episodes. The 1895 and 1922 unrest episodes have the highest estimated eruption probabilities, and are high even relative to the instrumental episodes. This is surprising considering no established monitoring network existed to record the 1895 and 1922 unrest episodes. It is very likely that if these events were to repeat themselves today, the eruption probabilities would be higher due to the more extensive and more accurate monitoring network. For example, ground deformation recordings (if any) would be more accurate, potentially allowing for ground deformation modelling (Parameter 9, Node 2). Seismic information may also be more accurate, with an increased number of recorded earthquakes (Parameter 6 and 3, Nodes 2 and 3) and the potential for moment tensors to be established (Parameter 1, Node 2). For example, assuming that if, in addition to what was recorded during the 1922 unrest episode, earthquake moment tensors and deformation due to inflation (Parameters 1 and 9, Node 2) were also observed, the probability of eruption (from the 2021 calibrated BET_EF model) would jump from 14.4% to 31.7%.

### 5.4. Validating the BET_EF model

A notable observation from the results is how high the estimated eruption probabilities are. The model likelihood of no eruption in the last 140 years (since records began) can be estimated using the retrospective calibrated model (Suite 2) as:

$$\prod_i (1-p_i) = (1-p_1) \times (1-p_2) \times (1-p_3) \times \ldots \times (1-p_{17}) = 0.042 \qquad (3)$$

where $p_i$ is the estimated probability of eruption from the $i$th episode.

Eq. (3) tells us that in the last 140 years, the actual outcome of not having an eruption had a likelihood, according to the BET_EF model, of only 4.2%. This is moderate evidence (at the 5% significance level) against the Taupō volcano BET_EF model, especially if the last 140 years are representative of the last 1,800 years since the last eruption. In particular it indicates that the BET_EF model is producing eruption probabilities that are unrealistically large.

One reason why may be due to the $a$ and $b$ priors within the BET_EF model (Eq. (2)). At Nodes 2 and 3 of the event tree, the anomalous monitoring parameters have a total weight of $Z$. This is converted into a probability through Eq. (2). In this equation, the parameters $a$ and $b$ are themselves random variables, with assigned priors, where $a$ is uniformly distributed between 0.5 and 1, and $b$ is uniformly distributed between 0 and 2. These priors have means of 0.75 and 1, respectively. In the absence of any past data, if there are no anomalous monitoring parameters ($Z = 0$), plugging these means into Eq. (2) yields a probability

of 0.25. So if no parameters are activated in Node 3, the probability of eruption is a quarter of the probability of magmatic unrest. The intent of the BET_EF is that the $a$ and $b$ priors should be updated with past data, specifically the number of non-eruptive monitored magmatic unrest episodes. With every prior magmatic unrest episode without eruption, the probability $P(E|M)$ will shrink as the posterior for $a$ and $b$ will move towards 1 (for $a$) and 2 (for $b$). A similar thing happens to $P(M|U)$, with each prior monitored unrest episode without evidence of magma

**Table 5**
BET_EF parameters for the first and final iterations of the Campi Flegrei model versus the Taupō volcano model. Yellow boxes indicate the most similar parameters between the three models.

| | CF First Iteration | CF Final Iteration | Taupō Volcano Final Iteration |
|---|---|---|---|
| **Node 1** | | | |
| 1. | # seismic events (per day); >1 - 85 | # VT EQ (per day); >5 - 15 | Moderate magnitude EQs (per month); >4 - 4.5 magnitude |
| 2. | Largest EQ magnitude (last month); >1.7 - 3.3 | #LP/VLP/ULP (per month); >2 - 10 | Increased rate of low magnitude EQs (per month); >50 - 100 |
| 3. | #LP events (last month); >0 | Cum. uplift (last 3 months); >2 - 6cm | LP/VLP/Tremor (last month); YES/NO |
| 4. | Cum. uplift (last 3 months); >1 - 20cm | Uplift rate (last 3 months); >0.7 - 1.3 cm/month | Appearance of gas emissions at surface AND/OR $CO_2/CH_4$ ratios increasing from background AND/OR C/S ratio increasing from background AND/OR He & Ar isotopes changing toward mantle values (last month); YES/NO |
| 5. | Uplift rate (last 3 months); >1 - 50 cm/month | New fractures (last 3 months); YES/NO | Notable changes to temperature of any surface feature(s) (last month); YES/NO |
| 6. | Positive gravimetric change (last 3 months); >36 - 180 µGal | Degassing structure extension or increase in flux (last month); YES/NO | Notable sustained surface deformation (in three months) change as determined by geodesist; YES/NO |
| 7. | Degassing structure extension or increase in flux (last month); >1.8 - 2.2 | Presence of acid gases (HF,HCl,$SO_2$)(last week); YES/NO | Long-term surface deformation (lake levelling) subtracted from agreed multi-year trend (i.e., background); >5 - 25mm/yr |
| 8. | Change in composition of gases (last month); YES/NO | Temperature at fumarole (Pisciarelli) (last month); >100 - 110 °C | Notable deviation to geodesist in short-term deformation rate over multiple days (horizontal and/or vertical), subtracted from agreed multi-year trend (i.e., background); YES/NO |
| 9. | $CO_2/CH_4$ ratio (last month); >0.2 | | Visual change - change in lake/spring/geothermal water appearance (last month); YES/NO |
| 10. | Change in temperature of gases; >10 - 15 °C | | |
| 11. | Change in thermal radiation of fractures; YES/NO | | |
| Node 2 | | | |
| 1. | # Seismic events (per day); >1 - 100 | # Deep VT EQ (every day); >2 - 20 (WEIGHT 0.9) | Non-double-couple EQ(s) (last month); YES/NO (WEIGHT 2) |
| 2. | Largest EQ magnitude (last month); >2.6 - 4.0 | # Deep LP EQ (every month); >3 - 20 (WEIGHT 0.5) | Migrating EQ hypocentres (vertically and/or laterally) (last month); YES/NO (WEIGHT 1) |
| 3. | #LP events (per month); >0 - 10 | #VLP/ULP (every month); >1 - 5 (WEIGHT 1) | LP/VLP/Tremor (last month); >1 - 10 (WEIGHT 1) |
| 4. | Maximum depth ( >6 events); >6 km | Presence of tremor (last month); YES/NO (WEIGHT 1) | Amplitude of LP/VLP/Tremor is larger than co-located (nearby) earthquakes (>mag. 2); YES/NO (WEIGHT 1) |
| 5. | Presence of CLVD; YES/NO | Presence of deep tremor (last month); YES/NO (WEIGHT 1) | Number of EQs within 1.5 magnitude of largest in swarm (from b-values or Omori decay) (last month); >50 - 100 (WEIGHT 1) |
| 6. | Cum. uplift (last 3 months); >1 - 100cm | Cum. uplift (last 3 months); >5 - 15cm (WEIGHT 1) | # EQs (per day); >5 - 100 (WEIGHT 1) |
| 7. | Positive gravimetric change (last 3 months); >100 - 270 µGal | New fractures (last 3 months); YES/NO (WEIGHT 0.2) | $CO_2/CH_4$ ratios increasing AND/OR C/S ratio decreasing toward mantle values AND/OR $^3He/^4He$ & $^{40}Ar/^{36}Ar$ shifting toward mantle values AND/OR $\mathbf{d}^{15}\mathbf{N}$ toward mantle value (last month); YES/NO (WEIGHT 2) |
| 8. | Presence of acid gases (HF, HCl, $SO^2$) (last week); YES/NO | Macroscopic variation on the deformation pattern (tens of m) (last 3 months); YES/NO (WEIGHT 1) | Temperature of any fumarole(s)/surface feature(s) (>100 °C) (last month); YES/NO (WEIGHT 1) |
| 9. | Change in temperature of gases (last month); >20 - 30 °C | Presence of acid gases (HF, HCl, $SO^2$) (last week); YES/NO (WEIGHT 1) | Deformation directly related to volume increase or pressure via modelling (inflation at depth); YES/NO (WEIGHT 2) |
| 10. | | Variation in magmatic component (last month); YES/NO (WEIGHT 0.1) | Visual change - low lake levels reported along nearby shoreline AND/OR increased bubbling in the lake AND/OR reports of unusual lake outflow/ lake level changes in marinas (that have not been seen in decades) (last month); YES/NO (WEIGHT 1) |
| Node 3 | | | |
| 1. | # seismic events (per day); >50 - 240 | Acceleration in # seismic events (last week); YES/NO (WEIGHT 1) | Magnitude of largest EQ (per day); >5.5 - 6.5 (WEIGHT 1) |
| 2. | Acceleration in # seismic events (last week); YES/NO | Acceleration in RSAM (last week); YES/NO (WEIGHT 0.7) | Resolvable vertically migrating earthquakes within the shallow crust (<5km depth) (last month); YES/NO (WEIGHT 1) |
| 3. | Presence of tremor (last month); YES/NO | Presence of tremor (last month); YES/NO (WEIGHT 1) | # EQs (per day); >100 - 1000 (WEIGHT 1) |
| 4. | Upward migration (EQs); YES/NO | Hypocenter dispersion (EQs) (last week); >1 - 3 km (WEIGHT 0.3) | Duration of current sustained tremor/continuous saturation of EQs (hours); >0.1 - 24 (WEIGHT 2) |
| 5. | Cum. uplift (last 3 months); >50 - 100cm | Macroscopic variation on the deformation pattern (tens of m) (last week); YES/NO (WEIGHT 1) | Sudden change or rate of seismicity (sudden quiescence); YES/NO (WEIGHT 1) |
| 6. | Acceleration in uplift; YES/NO | Migration if incremental maximum (m) (last week); YES/NO (WEIGHT 0.7) | CO2/CH4 ratios at magmatic values AND/OR C/S ratios decreasing toward "magmatic" values AND/OR $\mathbf{d}^{15}\mathbf{N}$ at mantle values AND/OR He & Ar isotopes at mantle values (last month); YES/NO (WEIGHT 2) |
| 7. | Macroscopic variation on the deformation pattern (tens of m) (last week); YES/NO | New fractures (last three months); YES/NO (WEIGHT 0.4) | Temperature of any fumarole(s)/surface feature(s) (>400 °C) (last month); YES/NO (WEIGHT 1) |
| 8. | New fractures (last 3 months); YES/NO | Presence of acid gases (HF, HCl, $SO_2$) (last week); YES/NO (WEIGHT 1) | Acceleration in deformation rate noted across the most recent data points; YES/NO (WEIGHT 2) |
| 9. | Degassing structure extension or increase in flux (last week); >5 - 50 | Phreatic activity (last week); YES/NO (WEIGHT 1) | Migration on any 3D direction of surface deformation (pattern consistent with magma migration) (last month); YES/NO (WEIGHT 1) |
| 10. | Change in temperature of gases (last week); >20 - 50 °C | | Visual change - phreatic activity AND/OR ground cracking AND/OR minor visual faulting (last month); YES/NO (WEIGHT 1) |
| 11. | Presence of acid gases (HF, HCl, $SO^2$) (last week); YES/NO | | |

reducing the posterior probability. In general, the eruption probability (Eq. (1)) drops because prior data modulates the $a$ and $b$ priors, updating them to reflect that given values of $Z$ have so far not resulted in an eruption.

For example, this could mean that the probable numbers of prior unrest episodes that may have involved magma is too low for both the 2019 calibrated Taupō volcano BET_EF model (one of the 16 prior unrest episodes was magmatic) and the 2021 calibrated Taupō volcano BET_EF model (2 of the 17 unrest episodes were magmatic). There could very well have been more than two magmatic unrest episodes in the unrest catalogue (Table 1). At present the probability of eruption from the processed 2019 unrest episode was 30.4%, assuming one of the 16 prior unrest episodes was magmatic. If instead we assume that there were five prior magmatic unrest episodes, the eruption probability drops to 13.7%. Assuming 10 prior magmatic unrest episodes reduces it further to 7.6%, a fourfold decrease. The uncertainty bounds of the eruption probabilities also narrow with the increased number of assumed prior magmatic unrest episodes.

To get a sense of where a realistic lower bound on the eruption probabilities might lie, let us suppose that unrest occurs roughly once per decade (as seen during the last three unrest episodes at Taupō volcano) and that the last eruption was about 1,800 years ago. Hence there could have been c. 180 unrest episodes without eruption since the last eruption. Further assuming that there are 0.118 magmatic unrest episodes per unrest episode (2 magmatic unrest episodes from 17 unrest episodes, Table 3), there would have been about 21 magmatic unrest episodes in the last 1,800 years (assuming 180 unrest episodes). Using this as prior data, we can recalculate the eruption probabilities for all the unrest episodes. These recalculated eruption probabilities were then input to Eq. (3), resulting in a increase in the model likelihood of no eruption in the last 140 years from 4.2% to 35%, assuming that the 1877–2021 period is representative of the activity between the Taupō eruption and 1877. A 35% model likelihood of no eruption in the last 140 years is less concerning from the viewpoint of model validation, and shows how more prior data within the BET_EF model can affect the resulting eruption probability estimates. It also suggests that more work is required to refine the model, a conclusion that applies generally to all short-term forecasting models such as BET_EF (Whitehead and Bebbington, 2021).

### 5.5. *Taupō volcano BET_EF Vs Campi Flegrei BET_EF*

Having considered internal validation of the Taupō BET_EF, we can also gauge how the Taupō volcano BET_EF model compares with the other BET_EF models. The Campi Flegrei BET_EF model by Selva et al. (2012) (BETEF_CF) is considered here to be the gold standard due to the updated form of the model, and the depth and breadth of the elicitation, and is also for a caldera volcano. Researchers held five expert elicitation workshops over five years, each time refining parameters and their thresholds. Table 5 compares the first and final BETEF_CF model iterations against the final Taupō volcano BET_EF model.

For Node 1 in Table 5, the first iteration and the final iteration for the BETEF_CF model contain 11 and eight parameters respectively, compared to Taupō volcanoes nine. Of the Taupō volcano BET_EF model's nine parameters, five are very similar to the first iteration. Four are similar to those in the final iteration of the BETEF_CF model. Node 1 parameters are particular to each volcano, representing the 'abnormal' behaviours above volcano-specific background behaviours. Overall, there is a good spread of seismic, geodetic, and geochemical parameters for all three BET_EF models.

The shared parameters (Node 2) in the first iteration of the BETEF_CF model are the 'number of seismic events', 'presence of CLVD (earthquake focal mechanism)' and 'the presence of acid gases'. The shared parameters in the final iteration of the BETEF_CF model are 'presence of acid gas', 'number of VLP/LP earthquakes', and 'presence of tremor'. Of these three shared parameters, only the 'presence of acid gases' was unchanged between the first and final BETEF_CF model iterations. It should be noted that the parameter 'presence of CLVD mechanisms' was not removed due to a lack of consensus among experts, but rather because the seismicity at Campi Flegrei (at least at that time) was too low to compute the CLVD mechanisms with reasonable accuracy (L. Sandri, personal communication). Overall, all three BET_EF models are skewed toward seismic parameters, with five out of nine parameters seismic in nature for the first iteration BETEF_CF model (55%), five out of ten for the last iteration of the BETEF_CF model (50%), and six out of ten for the Taupō volcano BET_EF model (60%). For the case of Taupō volcano, this is primarily due to the limited monitoring presence of geodetic and geochemical parameters, However, at Campi Flegrei, there is much more comprehensive monitoring.

For Node 3 of Table 5, Taupō volcano's ten parameters are comparable to both BETEF_CF models. However, of Taupō BET_EF models ten parameters, only seven are regularly monitored at Taupō volcano. Five of these parameters are seismic and visual and are monitored in real-time (Parameters 1, 3, 4, 5, and 10). Parameters 8 and 9 are geodetic in nature. They are also considered to be monitored in real time, but need to be processed so there is a time delay in data collection. Three parameters are not currently monitored at Taupō volcano (Parameters 2, 6, and 7). Overall, there is a reasonable level of similarity of parameters between the final BETEF_CF model versus the Taupō volcano BET_EF model, especially considering the monitoring challenges of Taupō volcano.

In terms of parameter weights, they differ considerably between the three models. In the first BETEF_CF iteration, no weights are specified, with only upper and lower thresholds reported (see Selva et al., 2012). Weights are specified in the final BETEF_CF iteration. However, they are fractional rather than integer, ranging between 0.1–1. Weights were identified from the estimates provided by experts through a weighted procedure. Lower and upper thresholds were selected as the 50th percentile of the corresponding distribution. This is effectively reducing the estimated probability of eruption from any combination of monitoring signals, thus compensating for the lack of prior data to restrict the $a$ and $b$ coefficients in Eq. (2). In contrast, the Taupō volcano BET_EF parameters have weights of either 1 or 2, as per the original BET_EF formulation of Marzocchi et al. (2008).

In Node 2 of the final iteration BETEF_CF model, the highest weighted parameters are '#VLP/ULP', 'Presence of tremor', 'Presence of deep tremor', 'Cum. uplift', 'Macroscopic variation on the deformation pattern', and 'Presence of acid gas'. The highest weighted parameters in Node 2 of the Taupō volcano BET_EF model as 'Non-double-couple earthquakes', 'Presence of gas', and 'Deformation directly related to volume increase or pressure'. Between the two models, only one common parameter has a higher weight - the geochemical parameter. Parameters 'VLP/ULP', 'Presence of tremor', and 'Presence of deep tremor' for the BETEF_CF model have a higher weight; However, the 'LP/VLP/Tremor' parameter of the Taupō volcano BET_EF model does not, as decided upon in the expert elicitation workshop.

For Node 3 of the final BETEF_CF model, parameters 'Acceleration in seismic events', 'Presence of tremor', 'Macroscopic variation on the deformation pattern', 'Presence of acid gases', and 'Phreatic activity' have a high weighting. In Node 3 of the Taupō volcano BET_EF model, parameters 'Duration of current sustained tremor', 'Presence of gas', and 'Acceleration in the deformation rate' have an increased weighting of 2. Compared to Node 2, Node 3 parameters of both BET_EF models share similarly higher-weighted parameters. The geochemical parameters, earthquake tremor, and variation/acceleration in deformation parameters in each BET_EF models have higher weightings. The parameters 'Acceleration in seismic events', and 'Phreatic activity' in the BETEF_CF model have a higher weighting, but the equivalent parameters in the Taupō volcano BET_EF model do not. This was a consequence of considerable discussion, about such events and the neighbouring geothermal systems, at the elicitation workshop.

BETEF_CF, especially compared to the Taupō volcano BET_EF, has a

large number of repeated (nested) monitored parameters (repeating backward from Node 3 to 1) that have the same threshold and 'look back' period. This is needed in order to be able to link the importance of a parameter as an indicator of a transition in the volcanic system, particularly for parameters that are present at all stages and are continuous throughout. For example, the 'presence of acid gases $(HF,\ HCl,\ SO_2)$' parameters is in Nodes 1, 2, and 3, with a lookback period of 'last week', with a binary threshold of YES/NO. Therefore, if acid gases are found at Campi Flegrei, the BET_EF model will abruptly jump from 'no unrest' to an elevated level of eruption probability. However, if acid gases are detected at the surface of Campi Flegrei, it indicates that all the large aquifers have been saturated, which suggests that an eruption is imminent (L. Sandri, personal communication). This also occurs with the parameter 'presence of tremor'. The Taupō volcano BET_EF model also has nested parameters, but a number of these have telescoping thresholds (i.e., the thresholds evolve throughout the nodes), such as LP/VLP/Tremor, and fumarole temperature. The idea is that these would lead to a less abrupt rise in the estimate probability of eruption, assuming frequent enough sampling.

Overall, the Taupō volcano BET_EF model is very similar to the first iteration of the BETEF_CF model, in terms of the presence of the parameters. This is on par with what is expected, considering the Taupō volcano BET_EF model was developed in less than one year compared to the final BETEF_CF model which was developed over five years. This leaves room for further development and refinement of the Taupō volcano BET_EF model as more data comes to light, whether that be another unrest episode, new information about past-eruptive behaviour, new monitoring installations, or even a new conceptual model.

### 5.6. 2019 unrest episode

For the 2019 unrest episode, unlike every other recorded unrest episode at Taupō volcano, has two different seismic datasets. One is classed as the 'processed' dataset, in which Illsley-Kemp et al. (2021) processed publicly available GeoNet data. The other dataset is the publicly available GeoNet data, classed as the 'GeoNet' dataset. The two datasets provide the opportunity to see how improved monitoring may change estimated eruption probabilities.

When run through the Taupō volcano BET_EF model, the GeoNet based 2019 unrest data had an estimated peak probability of eruption of 16%, while the processed 2019 unrest data had an estimated peak probability of eruption of 42% - two and a half times as high, which has potentially disturbing implications. The processed 2019 unrest data produced higher probabilities, due to the identification of non-double couple earthquakes, but the probabilities are more volatile compared to the GeoNet 2019 unrest data (Fig. 5), due to the reprocessing of the seismic data into swarms, as seismic data in monthly/daily rates aligns better with the Taupō volcano BET_EF model.

The potential downside is that there are inherent delays in post-processing monitoring data. Processing information takes time, time that may not be readily available under a period of intense unrest or crisis, where every hour of processing is an hour less for decision and action.

Real-time data is likely to produce less detailed eruption probabilities from the Taupō volcano BET_EF model, as multiple parameters in the model need additional levels of data processing (e.g., non-double-couple earthquakes). However, real-time data is immensely valuable in crisis situations, where there may be little to no time to process data before decisions must be made. A recommendation is to explore automated reprocessing to better exploit monitoring data in real-time (e.g., the automatic detection of non-double-couple earthquakes).

### 5.7. Double weight parameters

In Nodes 2 and 3 of the Taupō volcano BET_EF model, there are several parameters that have a weight of 2, compared to the unit weight of the others. As a sensitivity analysis every unrest episode was run through their retrospective BET_EF calibrations, but with unit weights for each parameter (i.e., every parameter has an equal weight of 1). It was found that differences in the probabilities of magmatic unrest and eruption were minimal, except for the 2019 unrest episode. For the 2019 unrest episode, there was a maximum factor of two difference in the

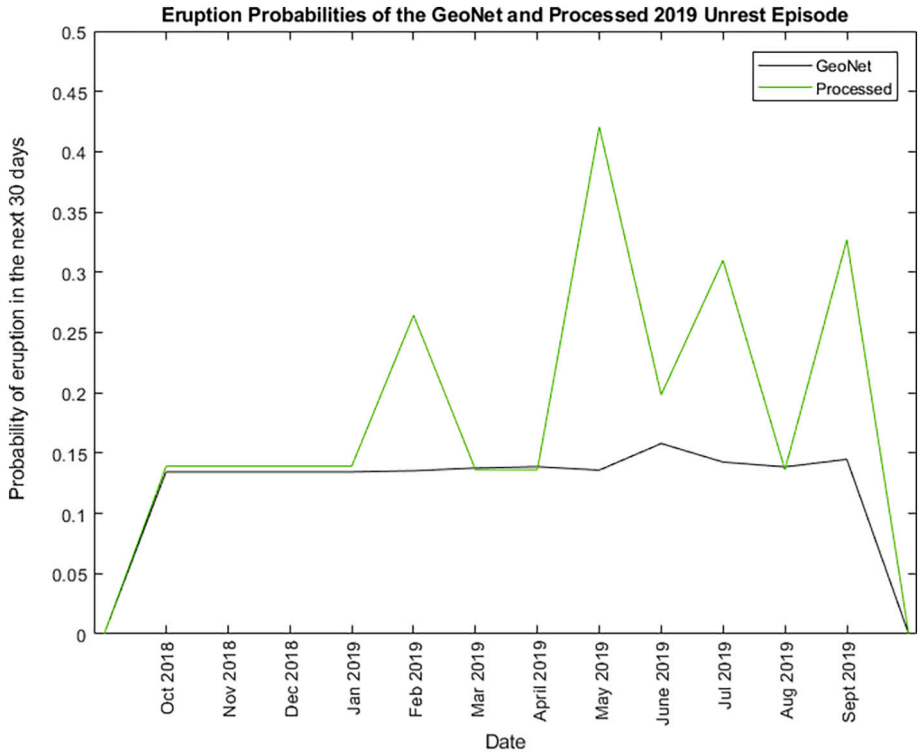



**Fig. 5.** Probabilities of eruption in the next 30 days from the Taupō volcano BET_EF: GeoNet and Processed 2019 Unrest episode.

estimated probability of eruption between the elicited parameter weights (15.8%) and the unit parameter weights (all parameters have an assigned weight of 1) (5.5%). The double weight parameter 'Deformation directly related to volume increase or pressure' in Node 2 is the main cause of the difference.

As the 2019 unrest episode has two different datasets (see Section 5.4), a parameter weight sensitivity analysis was also conducted between the GeoNet and processed 2019 datasets. For the processed 2019 unrest, there was a maximum difference in the estimated probability of eruption of 28% between the elicited parameter weights (42%) and the unit parameter weights (13.6%). Differing from the GeoNet 2019 unrest episode, the parameter 'Non-double-couple earthquakes' was anomalous. This parameter is also in Node 2, with a higher weight of 2. As the processed 2019 unrest episode activated two double weight parameters in Node 2 (compared to only one double weight parameter for the GeoNet 2019 unrest episode), there was greater difference in the estimated eruption probabilities between unit weight and elicited weight parameters.

## 6. Conclusions

Given Taupō volcano's potential for prolonged unrest and/or high impact eruptions, the development of the BET_EF model is an essential first step for expanding the available probabilistic forecasting tools in Aotearoa-New Zealand. Having a usable probabilistic forecasting tool means that when Taupō volcano undergoes another period of unrest or eruption, this model will be available to provide support to scientists and decision-makers. It can also be a stepping stone for which future model improvements or alternatives can be compared. It was found that the time variable monitoring network around Taupō volcano and the parameter weights had a significant impact on estimated eruption and magmatic unrest probabilities. Further work is needed to reduce the uncertainty in the available datasets and estimates. Magmatic unrest and eruption probabilities were found to be relatively high, with an estimated model likelihood of the observed last 140 years without eruption of only 4.7%. We experimented with additional hypothetical past data to obtain a realistic lower bound on this likelihood, but the use of fractional weights in the BET_EF should also be considered, with the caveat that more intensive elicitation is needed to assign them, and more comprehensive monitoring needed to support their use. In general, updated monitoring capabilities and more sophisticated processing of monitoring data, especially seismicity, should prompt updates and therefore refinements of the model.

## CRediT authorship contribution statement

**Emmy Scott:** Methodology, Investigation, Data-curation, Writing-original-draft. **Mark Bebbington:** Conceptualization, Methodology, Writing-review-editing, Supervision. **Thomas Wilson:** Conceptualization, Supervision, Writing-review-editing. **Ben Kennedy:** Supervision, Methodology. **Graham Leonard:** Investigation, Methodology, Supervision.

## 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

Data will be made available on request.

## Acknowledgements

This work was supported by the New Zealand Endeavour Fund programme ECLIPSE, grant RTVU1704. The authors would also like to acknowledge Laura Sandri and one anonymous reviewer for their insightful remarks and careful critiques of this research which lead to major improvements.

## Appendix A

**Prior Distribution and Past Data** The BET_EF nodes require the provision of a prior distribution for each node and, if available, can accommodate past data from observed unrest episode. Here we limit our past data to the observed unrest sequences since April 1878 (1730 months to June 2021).

Our prior distributions are all beta distributions, and because of the restless nature and relatively short observation period, will all have maximum ignorance (number of equivalent data = 1). It remains to set the mean of the distributions.

In historical terms, one in six episodes of unrest at silicic volcanoes worldwide have produced eruptions (Newhall and Dzurisin, 1988). We reject this as a prior mean, however, as it is incompatible with the 17 observed unrest episodes (likelihood of no eruption $= (5/6)^{17} = 0.045$) and even more so with the time since the last eruption, assuming that the observed level of unrest predated the beginning of observation. Instead we will work with:

$$P(Eruption|Unrest) = 1/60 = P(Eruption|Magma)P(Magma|Unrest). \quad (A.1)$$

The latter two quantities are the prior probabilities for Nodes 3 and 2, respectively. This is a subjective prior mean representing our best estimate of the probability of eruption.

**Node 1: Unrest/No Unrest**

Node 1 of BET_EF considers whether there is unrest or no unrest in the time interval $(t_0, t_0 + \tau)$, where $t_0$ is the present time, and $\tau$ is the forward time window considered. For the Taupō volcano BET_EF, a 30-day default window will be used. Taupō volcano is restless and according to some definitions of 'unrest', could be considered as always in a state of unrest. There is no consensus on the "normal" or "background" state of Taupō volcano due to the uncertainties and fluctuations in the baseline behaviour of the volcano. Because of this, Node 1 will be classed as any notable changes to the volcano from the "background". This differentiation is needed to make sense of the eventual probability of eruption, as otherwise "normal" behaviours of the volcano will elevate the eruption probability.

**Prior Distribution:** If we consider the post-Oruanui record of 28 eruptions in 25,500 years, we are led via Eq. (A.1) to approximately $28 \times 60 = 1680$ unrest episodes in 306,000 months, for a prior mean of P(Unrest) = 0.005 (Table 3).

**Past Data:** We have observed 17 unrest episodes covering 4345 days in 1730 months. However, in 128 of these months the system was already in unrest, and hence our past data are 4345/30 out of 1602 months (Table 3).

**Parameter 1:** Moderate Magnitude EQ(s) (>magnitude 4) are uncommon at Taupō volcano. There have only been seven earthquakes of magnitude 4.5 and above in the Taupō area in the last 40 years. Any earthquake magnitude 4 or above would indicate unrest, whether that be tectonic or magmatic in origin. Earthquake data will only be considered if they originate <30 km deep (above the subduction interface), within the crustal mush

and crustal region.

**Parameter 2:** If there were an increased seismicity rate, to between 50–100 earthquakes per month, that would be a notable change. For the current seismic network at Taupō volcano, any magnitude earthquake is included. However, with an enhanced network, more earthquakes will be able to be located and recorded. Therefore, with an enhanced network, only earthquakes of magnitude 2 should be considered.

**Parameter 3:** LP/VLP, and Tremor have never been recorded at Taupō volcano, even during periods of unrest. It would undoubtedly be anomalous if any of these seismic events were to occur.

**Parameter 4:** Currently, geochemical parameters are not monitored, and if they are (i.e., at Wairakei/Tauhara geothermal fields) they are not monitored through a volcanic lens. The sheer size of Lake Taupō inhibits any geochemical monitoring within the lake, as any gases will dissipate in the water. However, it is important to include geochemical parameters for any future periods of unrest in case gas is monitored. All geochemical parameters are clustered into one parameter for this model, as they are unlikely to be independent of each other. This is also the case for the geochemical parameters in Node 2 and 3.

**Parameter 5:** Currently, temperature recordings are done periodically in the geothermal fields outside of Lake Taupō. There are also hot water channels in and around Lake Taupō that members of the public can interact with. If any of these surface features were to have any notable changes in temperature, it is likely to be picked up.

**Parameter 6:** The qualifier notable is used to indicate any deformation (in the long term i.e., more than three months) that is different from fluctuations observed in the past. There are a limited number of GPS/GNSS stations around the lake, which affects the ability to define numbers regarding ground deformation.

**Parameter 7:** Similarly to parameter 6, this parameter is about surface deformation in the long term, but captured by lake levelling surveys. The thresholds of >5–25 mm/yr are derived from five uplift episodes over the last 41 years. For any future periods of unrest, any recordings should be subtracted from the multi-year trend (i.e., background) (xx mm/yr; set date; over X timeframe).

**Parameter 8:** Any notable ground deformation over a couple of days would definitively indicate some for of unrest.

**Parameter 9:** Lake Taupō is a popular tourist attraction. People will likely notice any visual changes in and around the lake. Changes in the lake's appearance, surrounding springs, and geothermal waters may indicate some increased activity at the volcano, e.g., changes in the hydrothermal system or increased gas emissions from fumaroles.

**Node 2: Magma/No Magma**

**Prior Distribution:** Given Eq. (A.1), our prior estimate of P(Magma|Unrest) we subjectively put at 1/10, to reflect the very active hydrothermal system, and the consequence that we believe there are more unrests per magma episode than magma episodes per eruption (Table 3).

**Past Data:** There has been only one unrest episode definitely identified as magmatic in origin - the 2019 unrest episode (Illsley-Kemp et al., 2021). It is likely that some of the other 17 unrest episodes were also magmatic. So, a conservative estimate of two magmatic unrest episodes is considered (Table 3).

**Parameter 1:** Non-double-couple earthquakes have been shown experimentally to result from the involvement of aqueous fluids and are inferred to occur from the movement of magma (Clarke et al., 2019). Because of this, this parameter has a higher weight of 2, as it is more likely to indicate magma on the move than other parameters. Any presence of non-double-couple earthquakes could suggest magma on the move at Taupō volcano.

**Parameter 2:** Migrating hypocentres can indicate magma on the move, whether that be vertically and/or laterally. This level of detail in earthquake location is not feasible with the current seismic network at Taupō volcano. It is unlikely to be operational, yet it is included in the BET_EF model if it is detected. This is one of many examples in this Appendix underlying the recommendation for improved monitoring.

**Parameter 3:** As mentioned previously LP, VLP, and tremor have not been recorded at Taupō volcano. Therefore, any more than 10 would be a great concern.

**Parameter 4:** As there has been no LP/VLP/Tremor recorded at Taupō volcano, there is no precedent for potential amplitudes of these events. Amplitude is an important part of seismic monitoring, as it is helpful to distinguish LP/VLP/Tremor events from normal volcanic-tectonic events. Amplitude is dependent on the location of the LP/VLP/Tremor event, the location of the seismometer, and the attenuation structure of the crust. As these things are highly variable at Taupō volcano, it is hard to define any specific amplitudes. Because of this, to be anomalous, any amplitude of LP/VLP/Tremor have to be larger than nearby earthquakes (>magnitude 2).

**Parameter 5:** In this BET_EF model, a swarm is defined as when multiple earthquakes are within 1.5 magnitude units of the largest earthquake. Swarms can also be differentiated from mainshock-aftershock sequences from b-values and/or Omori decay (Shcherbakov et al., 2004; Burris et al., 2008). At Taupō volcano, if there are more than 50 earthquakes in a swarm, it is anomalous.

**Parameter 6:** If there are more than five earthquakes (<30 km deep) a day at Taupō volcano, it would be considered anomalous. If there were more than 100 earthquakes, that would be considered extremely anomalous. With an enhanced network around Taupō volcano, only earthquakes more than magnitude 2 should be considered for this parameter.

**Parameter 7:** As mentioned in Node 1, geochemistry is not actively monitored at Taupō volcano. Thresholds of each sub-parameter have been proposed for Taupō volcano by geochemical and gas experts at the expert elicitation workshop, so there is considerable uncertainty surrounding specific thresholds. This parameter has a weighting of 2.

**Parameter 8:** If the temperature of any fumarole(s) or surface feature(s) reaches boiling point, that is a cause for concern, as that indicates that magma is close to the surface (i.e., shallow crust). If in close proximity, both magma and magmatic fluids can interact with and cause hydrothermal fluids to heat up.

**Parameter 9:** If any recorded ground deformation is directly related to volume or pressure increase, it is a cause for concern. Pressure and volume increase can be due to the movement of magma into the shallow crust, but can also be caused by pore water pressurisation due to the heating of hydrothermal fluids. Even in hydrothermal heating causing ground deformation, magma is involved in the heating of these fluids. Because of this, a weighting of 2 has been applied to this parameter.

**Parameter 10:** Any notable changes in lake levels at this node are an important visual cue for ground deformation, which may be due to magma. Due to the presence of the lake, ground deformation may go unnoticed. So, if ground deformation was large to cause visible lake level changes, that is a cause for concern. Another important visual cue is increased bubbling in the lake around Horomatangi Reefs. Bubbling is currently occurring in the lake, though increased bubbling would imply an increased discharge rate of the Horomatangi vents.

**Node 3: Eruption/No Eruption**

**Prior Distribution:** To be consistent with Eq. (A.1) and our prior mean at Node 2, the prior mean is set at 1/6 (Table 3).

**Past Data:** There have been no observed eruptions at Taupō volcano during the monitoring period (Table 3). As mentioned in Node 2, a

conservative number of two magmatic unrest episodes are assumed to have occurred at Taupō volcano.

**Parameter 1:** Large local (and potentially regional) earthquakes have the potential to trigger an eruption of there is eruptible magma in the system (e.g., Seropian et al., 2021). Any earthquake larger than 5.5 would be considered anomalous and could potentially trigger an eruption.

**Parameter 2:** Any vertically migrating earthquakes, particularly <5 km depth, would indicate shallow emplacement, if not an impending eruption. With the current seismic network around Taupō volcano, depth is unlikely to be resolvable to this degree.

**Parameter 3:** If there are more than 100 earthquakes (<30 km deep) a day at Taupō volcano, that would be anomalous. Any more than 1000 would be extremely anomalous. With an enhanced network, only earthquakes more than magnitude 2 should be considered.

**Parameter 4:** If there is a continuous saturation of earthquakes or tremors, that could indicate an impending eruption. If the continuous saturation of earthquakes, or tremor, lasted more than 10 min, that would be anomalous. If it lasted almost 24 h, that would almost certainly be anomalous and may indicate an eruption. Because of this, the weight of this parameter is 2.

**Parameter 5:** Before an eruption, it is not uncommon for a sudden reverse of seismic activity. This may be due to magma intrusions pausing in the crust, or that the magma pathways in the crust have already formed (Newhall and Dzurisin, 1988). Newhall and Dzurisin (1988) and Gottsmann et al. (2019) both found that sudden (hours/days) declines in seismicity, after previously increasing earthquake rates, can be a sign of an impending eruption.

**Parameter 6:** This parameter is directly analogous to the other geochemical parameters in Nodes 1 and 2. The higher sub-parameter thresholds would indicate that magma is very close to the surface (<5 km).

**Parameter 7:** If fumarole or surface feature temperate(s) increased over 400°C, it would suggest that a magma body was only hundreds of meters from the surface (Menyailov et al., 1986; Taran et al., 1995; Shinohara et al., 2002). From geochemical analysis of Oruanui eruption products from Taupō volcano, erupted magma pooled in the shallow crust for less than 600 years, at only 3.5 to 6 km depth (Allan et al., 2017). Therefore, any magma that is only a few hundred meters away from the surface is considered very likely to erupt.

**Parameter 8:** Accelerating deformation rates can indicate that something (magma or hydrothermal fluids) is moving rapidly within the crust. Observing ground deformation at Taupō volcano is equipment-dependent. Locals would certainly notice meters scale uplift, but due to the lack of GPS/GNSS sensors in the lake, high-rate analysis may not occur. Therefore, threshold values are currently not, and will likely never be, attached to this parameter. https://www.overleaf.com/project/6125bc377d7c7c3f88f52fd1.

**Parameter 9:** Similarly to the parameter above, migration of the surface deformation pattern may well indicate that magma is moving.

**Parameter 10:** Phreatic activity has never been witnessed at Taupō volcano. If phreatic activity/eruptions were to occur at Taupō volcano, that would indicate that magma is very close to the surface. For ground cracking and minor visual faulting to occur, the ground must be deforming rapidly. In conjunction with other unrest indicators, ground cracking and minor visual faulting are good indicators for magma moving in the shallow crust. They could also indicate tectonic unrest; however, tectonic unrest can also trigger eruptions if the magma is primed (i.e., a molten body of magma in the crust).

## References


Acocella, V., Di Lorenzo, R., Newhall, C., Scandone, R., 2015. An overview of recent (1988 to 2014) caldera unrest: Knowledge and perspectives. Rev. Geophys. 53 (3), 896–955.

Adams, N.K., Houghton, B.F., Fagents, S.A., Hildreth, W., 2006. The transition from explosive to effusive eruptive regime: the example of the 1912 Novarupta eruption, Alaska. Geol. Soc. Am. Bull. 118 (5–6), 620–634.

Allan, A.S.R., Barker, S.J., Millet, M.-A., Morgan, D.J., Rooyakkers, S.M., Schipper, C.I., Wilson, C.J.N., 2017. A cascade of magmatic events during the assembly and eruption of a super-sized magma body. Contrib. Miner. Petrol. 172 (7), 49.

Aspinall, W., Cooke, R., 1998. Expert Judgement and the Montserrat Volcano Eruption, pp. 2113–2118.

Barberi, F., Bertagnini, A., Landi, P., Principe, C., 1992. A review on phreatic eruptions and their precursors. J. Volcanol. Geoth. Res. 52 (4), 231–246.

Barker, S.J., Wilson, C.J., Illsley-Kemp, F., Leonard, G.S., Mestel, E.R., Mauriohooho, K., Charlier, B.L., 2020. Taupō: an overview of New Zealand's youngest supervolcano. NZ J. Geol. Geophys. 1–27.

Barker, S.J., Wilson, C.J.N., Allan, A.S.R., Schipper, C.I., 2015. Fine-scale temporal recovery, reconstruction and evolution of a post-supereruption magmatic system. Contrib. Miner. Petrol. 170 (1), 5.

Basualto, D., Pena, P., Delgado, C., Gallegos, C., Roa, H., Munoz, J., 2008. Seismic Activity Related to the Evolution of the Explosive Eruption of Chaiten Volcano in the Southern Andes Volcanic Zone. Eos Trans. Am. Geophy. Union 89.

Bebbington, M., 2020. Temporal-volume probabilistic hazard model for a supervolcano: Taupo, New Zealand. Earth Planet. Sci. Lett. 536, 116141.

Bebbington, M., Marzocchi, W., 2011. Stochastic models for earthquake triggering of volcanic eruptions. J. Geophys. Res.: Solid Earth 116 (B5).

Bebbington, M., Stirling, M.W., Cronin, S., Wang, T., Jolly, G., 2018. National-level long-term eruption forecasts by expert elicitation. Bull. Volcanol. 80 (6), 56.

Bebbington, M., Zitikis, R., 2016. Dynamic Uncertainty in Cost-Benefit Analysis of Evacuation Prior to a Volcanic Eruption. Math. Geosci. 48 (2), 123–148.

Bibby, H.M., Caldwell, T.G., Davey, F.J., Webb, T.H., 1995. Geophysical evidence on the structure of the Taupo Volcanic Zone and its hydrothermal circulation. J. Volcanol. Geoth. Res. 68 (1–3), 29–58.

Biggs, J., Ebmeier, S.K., Aspinall, W.P., Lu, Z., Pritchard, M.E., Sparks, R.S.J., Mather, T. A., 2014. Global link between deformation and volcanic eruption quantified by satellite imagery. Nat. Commun. 5 (1), 3471.

Brancato, A., Gresta, S., Alparone, S., Andronico, D., Bonforte, A., Caltabiano, T., Cocina, O., Corsaro, R.A., Cristofolini, R., di Grazia, G., Distefano, G., Ferlito, C., Gambino, S., Giammanco, S., Greco, F., Napoli, R., Sandri, L., Selva, J., Tusa, G., Viccaro, M., 2011. Application of BET_EF to Mount Etna: A retrospective analysis (Years 2001–2005). Ann. Geophys. 54 (5), 642–661.

Brancato, A., Gresta, S., Sandri, L., Selva, J., Marzocchi, W., Alparone, S., Andronico, D., Bonforte, A., Caltabiano, T., Cocina, O., Corsaro, R.A., Cristofolini, R., Di Grazia, G. D., Distefano, G., Ferlito, C., Gambino, S., Giammanco, S., Greco, F., Napoli, R., Tusa, G., Viccaro, M., 2012. Quantifying probabilities of eruption at a well-monitored active volcano: An application to Mt. Etna (Sicily, Italy). Boll. Geofis. Teor. Appl. 53 (1), 55–74.

Browne, P.R.L., Lawless, J.V., 2001. Characteristics of hydrothermal eruptions, with examples from New Zealand and elsewhere. Earth Sci. Rev. 52 (4), 299–331.

Burris, L., Christensen, B., Mcnutt, S., 2008. Swarms Versus Mainshock-Aftershock Sequences: A Systematic Study in Terms of Energy. AGU Fall Meeting Abstracts.

Carn, S.A., Zogorski, J.S., Lara, L., Ewert, J.W., Watt, S., Prata, A.J., Thomas, R.J., Villarosa, G., 2009. The unexpected awakening of Chaitén Volcano, Chile. Eos Trans. Am. Geophys. Union 90 (24), 205–206.

Castro, J.M., Dingwell, D.B., 2009. Rapid ascent of rhyolitic magma at Chaitén volcano, Chile. Nature 461 (7265), 780–783.

Charlier, B.L.A., Wilson, C.J.N., Lownstern, J.B., Blake, S., Van Calstern, P.W., Davidson, J.P., 2005. Magma Generation at a Large, Hyperactive Silicic Volcano (Taupo, New Zealand) Revealed by U-Th and U–Pb Systematics in Zircons. J. Petrol. 46 (1), 3–32.

Clarke, J., Adam, L., Sarout, J., van Wijk, K., Kennedy, B., Dautriat, J., 2019. The relation between viscosity and acoustic emissions as a laboratory analogue for volcano seismicity. Geology 47 (6), 499–503.

Colson, A.R., Cooke, R.M., 2018. Expert Elicitation: Using the Classical Model to Validate Experts' Judgments. Rev. Environ. Econ. Policy 12 (1), 113–132.

Connor, C.B., Hill, B.E., Brandi, W., Franklin, N.M., Femina, P.C.L., 2001. Estimation of Volcanic Hazards from Tephra Fallout. Nat. Hazards Rev. 2 (1), 33–42.

Constantinescu, R., Robertson, R., Lindsay, J.M., Tonini, R., Sandri, L., Rouwet, D., Smith, P., Stewart, R., 2016. Application of the probabilistic model BET_UNREST during a volcanic unrest simulation exercise in Dominica, Lesser Antilles. Geochem. Geophys. Geosyst. 17 (11), 4438–4456.

Constantinescu, R., Rouwet, D., Gottsmann, J., Sandri, L., Tonini, R., 2015. Tracking volcanic unrest at cotopaxi, ecuador: The use of the bet_ef tool during unrest simulation exercise. Geophys. Res. Abstr. EGU General Assembly 17, 2015–2251.

Cooke, R.M., 1991. Experts in Uncertainty: Opinion and Subjective Probability in Science. Oxford University Press, Cary, UNITED STATES.

Davy, B.W., Caldwell, T.G., 1998. Gravity, magnetic and seismic surveys of the caldera complex, Lake Taupo, North Island, New Zealand. J. Volcanol. Geoth. Res. 81 (1), 69–89.

De Ronde, C.E.J., Stoffers, P., Garbe-Schönberg, D., Christenson, B.W., Jones, B., Manconi, R., Browne, P.R.L., Hissmann, K., Botz, R., Davy, B.W., Schmitt, M., Battershill, C.N., 2002. Discovery of active hydrothermal venting in Lake Taupo, New Zealand. J. Volcanol. Geoth. Res. 115 (3–4), 257–275.

Endo, E.T., Murray, T.L., Power, J.A., 1996. A comparison of preeruption real-time seismic amplitude measurements for eruptions at mount st. helens, redoubt volcano,

mount spurr, and mount pinatubo. PHIVOLCS, University of Washington Press, pp. 233–247.
Fournier, N., Williams, C., Wallace, L., Sherburn, S., Jolly, A., Chardot, L., Ristau, J., Bourguignon, S., Hurst, T., Scott, B., Gibbs, M., Unglert, K., Beavan, R., 2013. From subduction processes to volcanic unrest: unraveling domino effects at Lake Taupo caldera. New Zealand (Invited) 2013, V44C–01.
Gottsmann, J., Komorowski, J., Barclay, J., 2019. Volcanic Unrest and Pre-eruptive Processes: A Hazard and Risk Perspective. In: Gottsmann, J., Neuberg, J., Scheu, B. (Eds.), Volcanic Unrest: From Science to Society. Springer International Publishing, Cham, pp. 1–21.
Illsley-Kemp, F., Barker, S.J., Wilson, C.J.N., Chamberlain, C.J., Hreinsdóttir, S., Ellis, S., Hamling, I.J., Savage, M.K., Mestel, E.R.H., Wadsworth, F.B., 2021. Volcanic unrest at Taupō volcano in 2019: causes, mechanisms and implications. Geochem. Geophys. Geosyst. 22 (6), e2021GC009803.
Johnston, D., Scott, B., Houghton, B., Paton, D., Dowrick, D., Villamor, P., Savage, J., 2002. Social and economic consequences of historic caldera unrest at the Taupo volcano, New Zealand and the management of future episodes of unrest. Bull. N. Z. Soc. Earthq. Eng. 35(4. SE - Articles).
Jolly, G., Beavan, J., Christenson, B., Ellis, S., Jolly, A., Miller, C., Peltier, A., Scott, B., Sherburn, S., Wallace, L., McCaffrey, R., 2008. What constitutes unrest at Taupo caldera. New Zealand? 2008, V44A–07.
Kennedy, B.M., Holohan, E.P., Stix, J., Gravley, D.M., Davidson, J.R.J., Cole, J.W., 2018. Magma plumbing beneath collapse caldera volcanic systems. Earth Sci. Rev. 177, 404–424.
Lavigne, F., Thouret, J.C., Voight, B., Young, K., LaHusen, R., Marso, J., Suwa, H., Sumaryono, A., Sayudi, D.S., Dejean, M., 2000. Instrumental lahar monitoring at Merapi Volcano, Central Java, Indonesia. J. Volcanol. Geoth. Res. 100 (1), 457–478.
Lee, S., Kang, N., Park, M., Hwang, J.Y., Yun, S.H., Jeong, H.Y., 2018. A review on volcanic gas compositions related to volcanic activities and non-volcanological effects. Geosci. J. (Seoul, Korea) 22 (1), 183–197.
Lindsay, J., Marzocchi, W., Jolly, G., Constantinescu, R., Selva, J., Sandri, L., 2010. Towards real-time eruption forecasting in the Auckland Volcanic Field: application of BET_EF during the New Zealand National Disaster Exercise 'Ruaumoko'. Bull. Volcanol. 72 (2), 185–204.
Manga, M., Brodsky, E., 2006. Seismic Triggering of Eruptions in the Far Field: Volcanoes and Geysers. Annu. Rev. Earth Planet. Sci. 34 (1), 263–291.
Manville, V., Segschneider, B., Newton, E., White, J.D.L., Houghton, B.F., Wilson, C.J.N., 2009. Environmental impact of the 1.8 ka Taupo eruption, New Zealand: Landscape responses to a large-scale explosive rhyolite eruption. Sed. Geol. 220 (3), 318–336.
Marzocchi, W., Bebbington, M., 2012. Probabilistic eruption forecasting at short and long time scales. Bull. Volcanol. 74 (8), 1777–1805.
Marzocchi, W., Newhall, C., Woo, G., 2012. The scientific management of volcanic crises. J. Volcanol. Geoth. Res. 247–248, 181–189.
Marzocchi, W., Sandri, L., Gasparini, P., Newhall, C., Boschi, E., 2004. Quantifying probabilities of volcanic events: The example of volcanic hazard at Mount Vesuvius. J. Geophys. Res.: Solid Earth 109 (B11).
Marzocchi, W., Sandri, L., Selva, J., 2008. BET_EF: a probabilistic tool for long- and short-term eruption forecasting. Bull. Volcanol. 70 (5), 623–632.
Marzocchi, W., Woo, G., 2007. Probabilistic eruption forecasting and the call for an evacuation. Geophys. Res. Lett. 34 (22).
Marzocchi, W., Woo, G., 2009. Principles of volcanic risk metrics: Theory and the case study of Mount Vesuvius and Campi Flegrei, Italy. J. Geophys. Res.: Solid Earth 114 (B3).
Massey, C.I., Manville, V., Hancox, G.H., Keys, H.J., Lawrence, C., McSaveney, M., 2010. Out-burst flood (lahar) triggered by retrogressive landsliding, 18 March 2007 at Mt Ruapehu, New Zealand—a successful early warning. Landslides 7 (3), 303–315.
Menyailov, I.A., Nikitina, L.P., Shapar, V.N., Pilipenko, V.P., 1986. Temperature increase and chemical change of fumarolic gases at Momotombo Volcano, Nicaragua, in 1982–1985: Are these indicators of a possible eruption? J. Geophys. Res.: Solid Earth 91 (B12), 12199–12214.
Newhall, C., Hoblitt, R., 2002. Constructing event trees for volcanic crises. Bull. Volcanol. 64 (1), 3–20.
Newhall, C.G., Albano, S.E., Matsumoto, N., Sandoval, T., 2001. Roles of groundwater in volcanic unrest. J. Geol. Soc. Philippines 56, 69–84.
Newhall, C.G., Dzurisin, D., 1988. Historical unrest at large calderas of the world. U.S. Geological Survey Bulletin, p. 1.
Newhall, C.G., Pallister, J.S., 2015. Using Multiple Data Sets to Populate Probabilistic Volcanic Event Trees. In: Shroder, J.F., Papale, P.B.T.V.H. (Eds.), Volcanic Hazards, Risks and Disasters. Elsevier, Boston, pp. 203–232.
Oskarsson, N., 1984. Monitoring of fumarole discharge during the 1975–1982 rifting in Krafla volcanic center, north Iceland. J. Volcanol. Geoth. Res. 22 (1), 97–121.
Otway, P.M., 1989. Vertical Deformation Monitoring by Periodic Water Level Observations, Lake Taupo, New Zealand. In: Latter, J.H. (Ed.), Volcanic Hazards. Springer, Berlin Heidelberg, Berlin, Heidelberg, pp. 561–574.
Otway, P.M., Blick, G.H., Scott, B.J., 2002. Vertical deformation at Lake Taupo, New Zealand, from lake levelling surveys, 1979–99. NZ J. Geol. Geophys. 45 (1), 121–132.
Otway, P.M., Sherburn, S., 1994. Vertical deformation and shallow seismicity around Lake Taupo, New Zealand, 1985–90. NZ J. Geol. Geophys. 37 (2), 195–200.
Pallister, J.S., McNutt, S.R., 2015. Synthesis of Volcano Monitoring. In: Sigurdsson, H. (Ed.), The Encyclopedia of Volcanoes, second ed. Academic Press, Amsterdam, pp. 1151–1171.
Peltier, A., Hurst, T., Scott, B., Cayol, V., 2009. Structures involved in the vertical deformation at Lake Taupo (New Zealand) between 1979 and 2007: New insights from numerical modelling. J. Volcanol. Geoth. Res. 181 (3), 173–184.
Phillipson, G., Sobradelo, R., Gottsmann, J., 2013. Global volcanic unrest in the 21st century: an analysis of the first decade. J. Volcanol. Geoth. Res. 264, 183–196.
Pierson, T.C., Wood, N.J., Driedger, C.L., 2014. Reducing risk from lahar hazards: concepts, case studies, and roles for scientists. J. Appl. Volcanol. 3 (1), 16.
Poland, M.P., Anderson, K.R., 2020. Partly Cloudy With a Chance of Lava Flows: Forecasting Volcanic Eruptions in the Twenty-First Century. J. Geophys. Res.: Solid Earth 125 (1), e2018JB016974.
Potter, S.H., Scott, B.J., GE, J., 2012. Caldera Unrest Management Sourcebook. GNS Science Report, 2012/12.
Potter, S.H., Scott, B.J., Jolly, G.E., Johnston, D.M., Neall, V.E., 2015. A catalogue of caldera unrest at Taupo Volcanic Centre, New Zealand, using the Volcanic Unrest Index (VUI). Bull. Volcanol. 77 (9), 78.
Pritchard, M.E., Mather, T.A., McNutt, S.R., Delgado, F.J., Reath, K., 2019. Thoughts on the criteria to determine the origin of volcanic unrest as magmatic or non-magmatic. Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci. 377 (2139), 20180008.
Rinawati, D.I., Sari, D.P., Handayani, N.U., Siwi, B.R., 2018. Predicting the probability of Mount Merapi eruption using Bayesian Event Tree_Eruption Forecasting. MATEC Web of Conferences, 154.
Rosenberg, M.D., Wilson, C.J.N., Bignall, G., Ireland, T.R., Sepulveda, F., Charlier, B.L.A., 2020. Structure and evolution of the Wairakei-Tauhara geothermal system (Taupo Volcanic Zone, New Zealand) revisited with a new zircon geochronology. J. Volcanol. Geoth. Res. 390, 106705.
Rouwet, D., Sandri, L., Marzocchi, W., Gottsmann, J., Selva, J., Tonini, R., Papale, P., 2014. Recognizing and tracking volcanic hazards related to non-magmatic unrest: a review. J. Appl. Volcanol. 3 (1), 17.
Rowe, G., Wright, G., 2001. Expert Opinions in Forecasting: The Role of the Delphi Technique. In: Armstrong, J.S. (Ed.), Principles of Forecasting, pp. 125–144.
Sahetapy-Engel, S., Self, S., Carey, R.J., Nairn, I.A., 2014. Deposition and generation of multiple widespread fall units from the c. AD 1314 Kaharoa rhyolitic eruption, Tarawera, New Zealand. Bull. Volcanol. 76(8):836.
Sandri, L., Acocella, V., Newhall, C., 2017. Searching for patterns in caldera unrest. Geochem. Geophys. Geosyst. 18 (7), 2748–2768.
Sandri, L., Guidoboni, E., Marzocchi, W., Selva, J., 2009. Bayesian event tree for eruption forecasting (BET_EF) at Vesuvius, Italy: a retrospective forward application to the 1631 eruption. Bull. Volcanol. 71 (7), 729–745.
Sandri, L., Jolly, G., Lindsay, J., Howe, T., Marzocchi, W., 2012. Combining long- and short-term probabilistic volcanic hazard assessment with cost-benefit analysis to support decision making in a volcanic crisis from the Auckland Volcanic Field, New Zealand. Bull. Volcanol. 74 (3), 705–723.
Sandri, L., Marzocchi, W., Zaccarelli, L., 2004. A new perspective in identifying the precursory patterns of eruptions. Bull. Volcanol. 66 (3).
Sandri, L., Tonini, R., Rouwet, D., Constantinescu, R., Mendoza-Rosas, A., Andrade, D., Bernard, B., 2017. In: Gottsmann, J., Neuberg, J., Scheu, B. (Eds.), The Need to Quantify Hazard Related to Non-magmatic Unrest: From BET_EF to BET_UNREST. Volcanic Unrest.
Selva, J., Marzocchi, W., Papale, P., Sandri, L., 2012. Operational eruption forecasting at high-risk volcanoes: the case of Campi Flegrei, Naples. J. Appl. Volcanol. 1 (1), 5.
Seropian, G., Kennedy, B.M., Walter, T.R., Ichihara, M., Jolly, A.D., 2021. A review framework of how earthquakes trigger volcanic eruptions. Nat. Commun. 12 (1), 1004.
Shcherbakov, R., Turcotte, D.L., Rundle, J.B., 2004. A generalized Omori's law for earthquake aftershock decay. Geophys. Res. Lett. 31 (11).
Sherburn, S., 1992. Seismicity of the Lake Taupo region, New Zealand, 1985–90. NZ J. Geol. Geophys. 35 (3), 331–335.
Shinohara, H., Kazahaya, K., Saito, G., Matsushima, N., Kawanabe, Y., 2002. Degassing activity from Iwodake rhyolitic cone, Satsuma-Iwojima volcano, Japan: Formation of a new degassing vent, 1990–1999. Earth Planet. Space 54 (3), 175–185.
Soosalu, H., Einarsson, P., Þorbjarnardóttir, B.S., 2005. Seismic activity related to the 2000 eruption of the hekla volcano, Iceland. Bull. Volcanol. 68, 21–36.
Sparks, R.S.J., 2003. Forecasting volcanic eruptions. Earth Planet. Sci. Lett. 210 (1–2), 1–15.
Sparks, R.S.J., Cashman, K.V., 2017. Dynamic Magma Systems: Implications for Forecasting Volcanic Activity. Elements 13 (1), 35–40.
Stagpoole, V., Miller, C., Caratori Tontini, F., Brakenrig, T., Macdonald, N., 2021. A two million-year history of rifting and caldera volcanism imprinted in new gravity anomaly compilation of the Taupō Volcanic Zone, New Zealand. NZ J. Geol. Geophys. 1–14.
Stix, J., de Moor, J.M., 2018. Understanding and forecasting phreatic eruptions driven by magmatic degassing. Earth Planet. Space 70 (1), 83.
Sutton, A.N., Blake, S., Wilson, C.J.N., 1995. An outline geochemistry of rhyolite eruptives from Taupo volcanic centre, New Zealand. J. Volcanol. Geoth. Res. 68 (1), 153–175.
Tabayag, S.G., 2010. Disaster risk reduction through land use planning in lahar-devastated footslopes of Mayon Volcano in Albay [Philippines]. Int. Soc. Southeast Asian Agric. Sci. (Philippines) 16.
Taran, Y., Hedenquist, J.W., Korzhinsky, M.A., Tkachenko, S.I., Shmulovich, K.I., 1995. Geochemistry of magmatic gases from Kudryavy volcano, Iturup. Geochim. Cosmochim. Acta 59 (9), 1749–1761.
Thompson, G., 1960. Increases in temperature measured at 35cm along established traverse lines at Wairakei. Department of Scientific and Industrial Research, New Zealand, 7(G.E.K.T. Geothermal circular).
Tierz, P., Clarke, B., Calder, E.S., Dessalegn, F., Lewi, E., Yirgu, G., Fontijn, K., Crummy, J.M., Bekele, Y., Loughlin, S.C., 2020. Event Trees and Epistemic Uncertainty in Long-Term Volcanic Hazard Assessment of Rift Volcanoes: The Example of Aluto (Central Ethiopia). Geochem. Geophys. Geosyst. 21 (10).

Tierz, P., Woodhouse, M.J., Phillips, J.C., Sandri, L., Selva, J., Marzocchi, W., Odbert, H. M., 2017. A Framework for Probabilistic Multi-Hazard Assessment of Rain-Triggered Lahars Using Bayesian Belief Networks. Front. Earth Sci. 5.
Tonini, R., Sandri, L., Rouwet, D., Caudron, C., Marzocchi, W., Suparjan, 2016. A new Bayesian Event Tree tool to track and quantify volcanic unrest and its application to Kawah Ijen volcano. Geochem. Geophys. Geosyst. 17 (7), 2539–2555.
White, R., McCausland, W., 2016. Volcano-tectonic earthquakes: a new tool for estimating intrusive volumes and forecasting eruptions. J. Volcanol. Geoth. Res. 309, 139–155.
Whitehead, M.G., Bebbington, M., 2021. Method selection in short-term eruption forecasting. J. Volcanol. Geoth. Res. 419, 107386.
Wilson, C.J., 1993. Stratigraphy, Chronology, Styles and Dynamics of Late Quaternary Eruptions from Taupo Volcano, New Zealand. Philos. Trans.: Phys. Sci. Eng. 343 (1668), 205–306.
Wilson, C.J., 2001. The 26.5 ka Oruanui eruption, New Zealand: An introduction and overview. J. Volcanol. Geoth. Res. 112 (1–4), 133–174.
Wilson, C.J., Blake, S., Charlier, B.L., Sutton, A.N., 2006. The 26.5 ka Oruanui eruption, Taupo Volcano, New Zealand: Development, characteristics and evacuation of a large rhyolitic magma body. J. Petrol. 47 (1), 35–69.
Wilson, C.J., Charlier, B.L.A., 2009. Rapid Rates of Magma Generation at Contemporaneous Magma Systems, Taupo Volcano, New Zealand: Insights from U-Th Model-age Spectra in Zircons. J. Petrol. 50 (5), 875–907.
Wilson, C.J., Rowland, J.V., 2016. The volcanic, magmatic and tectonic setting of the Taupo Volcanic Zone, New Zealand, reviewed from a geothermal perspective. Geothermics 59, 168–187.