

METHODOLOGY

Open Access



MatHaz: a Matlab code to assist with probabilistic spatio-temporal volcanic hazard assessment in distributed volcanic fields

Daniel Bertin^{1*} , Jan M. Lindsay¹, Laura Becerril², Shane J. Cronin¹ and Lizette J. Bertin³

Abstract

This paper introduces an open source computer code to perform an integrated probabilistic spatio-temporal volcanic hazard assessment in distributed volcanic fields. The program, named MatHaz, is a set of Matlab scripts that follows a sequential methodology. After the user has provided a set of input files, this tool first estimates the spatial probability of future volcanic vents, then the temporal probability of future volcanic events, and finally models up to five volcanic phenomena (pyroclastic density currents, ballistic projectiles, lava flows, lahars, and tephra fallout) following a probabilistic approach. These results can be combined and depicted as an integrated quantitative (and/or qualitative) volcanic hazard map, with weightings of hazard factors chosen by the user. We illustrate the use of this tool by applying it to the Carrán-Los Venados Volcanic Field in southern Chile. The open-source, replicable, and user-friendly nature of the code allows its application to any volcanic region of the world, regardless of its extent, type, and amount of volcano-structural data.

Keywords: Matlab, Spatial probability analysis, Temporal probability analysis, Probabilistic volcanic hazard assessment, Integrated volcanic hazard map

Introduction

At present, apart from analyses of vent location and/or onset time of eruption, few volcanic hazard assessments in volcanic fields have dealt with the distribution of specific eruptive hazards. This strong bias towards spatial and temporal assessments is mainly due to the difficulty in addressing, in a single assessment, the multiple volcanic phenomena that might be produced by any subsequent eruption. Some attempts concentrate on a single volcanic phenomenon, mainly pyroclastic density currents (PDCs; Sandri et al. 2012; Neri et al. 2015) and lava flows (Connor et al. 2012; Gallant et al. 2018); there are very few studies that have included several volcanic phenomena (Alcorn et al. 2013; Bartolini et al. 2014, 2015; Becerril et al. 2014, 2017; Sandri et al. 2014). Integration of multiple hazards into a single map has only been carried out qualitatively (or in one-off scenario

events), because a quantitative procedure for hazard integration has been lacking.

There is much past work on estimating the distribution of volcanic phenomena via computer modeling. Indeed, the volcanological literature includes numerous analytical, empirical and numerical models that simulate, with strongly varying degrees of detail and sophistication, specific volcanic phenomena. In general, the more sophisticated the model is, the longer the execution times. This is a challenge for hazard assessment in volcanic fields where there are many potential vents, adding an extra complexity, so that a strategy to adequately (and timely) assess multiple volcanic phenomena is desirable. This strategy places a premium on simplified models that are easy to manage and replicate and quick to execute, acknowledging that by meeting these conditions model resolution is lost.

In order to address the challenge of assessing hazard in volcanic fields, this paper introduces an open source computer code, named MatHaz, developed to carry out

* Correspondence: daniel.bertin.u@gmail.com

¹School of Environment, University of Auckland, Auckland, New Zealand
Full list of author information is available at the end of the article



an integrated (quantitative and/or qualitative) spatio-temporal volcanic hazard assessment for multiple volcanic phenomena. The way in which this model works can be summarized in three sequential steps. The first step combines all the volcano-structural data provided by the user and produces a map of spatial density of vent-location. The second step performs a temporal probability analysis with the eruptive record provided by the user, combining it with the spatial analysis to generate maps showing the probability of vent opening within specific time intervals. The third step separately models up to five volcanic phenomena (PDCs, ballistic projectiles, lava flows, lahars, and tephra fallout) based on the spatio-temporal analysis. All these results can be merged into a volcanic hazard map that displays the probability that a specific area is affected by any of the modeled volcanic processes in a specific time interval.

Our sequential approach is implemented in Matlab, enabling ease of use (compared to compiled codes) and replication. Being an open source code, the user can test and modify the files at any step, either to include other models or to migrate its routines to other programming languages.

The intended users of MatHaz are researchers who need to make hazard maps for volcanic fields, in order to either provide a regional context for hazards, or to rapidly develop a broad perspective on the hazards. This means that unlike other hazard simulation codes, MatHaz has not been developed to model the physics of volcanic eruptions and their products, or to model the details of specific eruptive scenarios.

Here, we introduce the model and illustrate its use by applying it to the Carrán–Los Venados Volcanic Field, a well-studied volcanic field in southern Chile.

Background

Distributed volcanism

Volcanic regions characterized by distributed volcanism are common on Earth. Volcanic activity in these regions frequently consists of individual (i.e., monogenetic), basaltic to basaltic andesitic, small-volume ($< 1 \text{ km}^3$) eruptions, characterized by episodic periods of activity that can last up to several years (Németh 2010; Kereszturi et al. 2013; Le Corvec et al. 2013; Valentine and Connor 2015). Over time these eruptions tend to form volcanic fields, built of scoria cones, maars, tuff cones, tuff rings, small shields and/or lava flows (Valentine and Gregg 2008; Valentine and Connor 2015). The life span of a distributed volcanic field can cover millions of years, suggesting recurrence rates on the order of 10^{-4} – 10^{-5} events/year, very low compared to the frequencies of eruptions at individual composite volcanoes (Németh 2010; Kereszturi et al. 2013; Valentine and Connor 2015). Hence, every new eruption is a rare phenomenon (Connor et al. 2018), so is typically treated in temporal hazard

models using a renewal process based on the history of past eruptions (Bebbington and Cronin 2011). A greater challenge is forecasting the location of future events in volcanic fields, because they frequently occur in completely new locations, often unrelated to time-sequence (Bebbington and Cronin 2011; Connor et al. 2018). For this reason, over the last 50 years several probabilistic and geological approaches have been applied to forecast the location, timing, type, and impacts/effects of future volcanic events in distributed volcanic fields.

Spatial, temporal, and volcanic hazard assessments

Spatial analyses for modeling possible locations of future volcanic events in regions of distributed volcanism have usually been carried out via kernel density estimation (Connor et al. 2018, and references therein). This is a non-parametric statistical method for estimating the probability density function (PDF) of a distributed sample (e.g., volcanic vents) in which the statistical parameters that rule it are unspecified. In volcanology, the most frequently used kernel functions are the radially symmetric (i.e., isotropic) and the elliptical (i.e., anisotropic) Gaussian kernels (Connor et al. 2012, and references therein). Both assume that the volcanic region has no boundaries, which is hard to reconcile with geological and/or geophysical evidence in some volcanic fields supporting a tectonic, magmatic or lithological control on the distribution of the volcanic vents (Connor et al. 2000; Martin et al. 2004; Germa et al. 2013; Deng et al. 2017).

Many models share an approach that recognises that inter-event variability lies generally within the constraints of long-term patterns in time, space, and style exhibited over the history of a volcanic field (Connor et al. 2015, and references therein). However, this is only reliable if long records are known, if dating is good, and if there are stable tectonic/volcanic conditions. In many cases, strongly time-variant behaviour occurs, which may affect frequency, location, and size of eruptions (Bebbington and Cronin 2011). In consequence, forecasting the rate of activity in volcanic fields can be a difficult task to accomplish (Bebbington and Cronin 2011; Runge et al. 2014; Richardson et al. 2017a). A comprehensive review is provided by Bebbington (2012), who grouped several classes of stochastic models into three categories: stationary, stationary with covariates, and non-stationary models, depending on how the intensity of the volcanic process varies with time. For volcanic fields with low recurrence rates (e.g., Auckland Volcanic Field, $\sim 10^{-4}$ events/year; Bebbington and Cronin 2011), or scarce temporal data (e.g., Harrat Al-Madinah; El Difrawy et al. 2013), it is usually assumed that the temporal component of the hazard assessment is independent of the spatial component (Bebbington 2013, 2015). Following this assumption, a long-term

average recurrence rate (a stationary model) can be calculated for the whole field based on the number of events over a specific time interval, which has been a common approach for several volcanic fields (e.g., El Difrawy et al. 2013; Runge et al. 2014; Bartolini et al. 2015; Gallant et al. 2018; Nieto-Torres and Martin Del Pozzo 2019).

Several techniques to integrate multiple background data have been developed during the last few years (e.g., Cappello et al., 2012, 2013; Alcorn et al. 2013; Becerril et al. 2013, 2014, 2017; El Difrawy et al. 2013; Bartolini et al. 2014, 2015; Gallant et al. 2018; Jiménez et al. 2018). Most of these approaches apply a similar statistical treatment of the available volcano-structural information resulting in spatial density maps showing the probability of vent location. In general, only the results obtained after the spatial analysis have been used as inputs for any subsequent volcanic hazard assessment, mainly because a temporal analysis was not conducted, and if it was, its results were not combined with those of the spatial analysis. In these studies the simulated volcanic processes have been restricted to: PDCs, lava flows, lahars and tephra fallout, i.e., ballistic projectiles are absent. Finally, in cases where the modeling results are combined, this has been done either via superposition or following a qualitative procedure for which the methodology has not been clearly stated.

Existing models and research gaps

A single package allowing the modeling of several volcanic processes for distributed volcanism was previously attempted in a Geographical Information System (GIS) framework (Felpeto et al. 2007). This tool, named VORIS, is able to determine the spatial probability of vent opening and to simulate lava flows, PDCs, and tephra fallout. The program has to be loaded as a toolbar in ArcMap (up to version 9.3) and a complete user's guide is provided. VORIS was a pioneer initiative in including different simulation tools in a single framework and it is still frequently used in volcanic hazard assessments (e.g., Alcorn et al. 2013; Bartolini et al. 2015; Becerril et al. 2017). However, this program assumes a unique scalar bandwidth for the spatial probability analysis, does not perform a temporal analysis, does not model either ballistic projectiles or lahars, and it is not possible to combine hazards into an integrated map. The code is not open access and therefore editing and testing of the program are not possible.

The QVAST tool (Bartolini et al. 2013) goes a step further in terms of estimating the spatial probability of vent opening. This is a Quantum GIS (QGIS) plugin useful for estimating the spatial density of future volcanic vents based on volcano-structural data. The program is able to use different methods to calculate the bandwidth of

every volcano-structural dataset, after which each probability map is weighted and combined into a single spatial density map. Its main advantages are its user-friendly interface and its development for the free and open-source geographic information system QGIS. However, QVAST mainly works with scalar bandwidths and the user cannot access the code. It also requires other tools to perform temporal and individual hazard analyses, so it has been commonly coupled with the VORIS tool for volcanic hazard assessments (e.g., Bartolini et al. 2015; Becerril et al. 2017; Jiménez et al. 2018).

More recently, some computer codes have been uploaded to the GitHub hosting service (<https://github.com/geoscience-community-codes>). This platform allows the user to conduct a spatial probability analysis of vent opening as well as to simulate lava flows (MOLASSES tool; Richardson et al. 2017b), PDCs (Energy cone model; Malin and Sheridan 1982), and tephra fallout (Tephra2 code; Bonadonna et al. 2005a). Unlike VORIS and QVAST, the user can freely access these codes. Some advantages are that, in the spatial probability analysis, the code automatically runs the scripts for obtaining the bandwidths, and that simulations for each volcanic process can be run online. Some drawbacks are their different programming languages for different volcanic processes (i.e., integration is not possible), the inability to model either ballistic projectiles or lahars, and that the user cannot upload anything, including topography.

MatHaz code

MatHaz consists of six files written in Matlab. To start the program, the file *MatHaz.m* has to be run in the Matlab command window. A flowchart describing MatHaz is shown in Fig. 1.

The first operational step of MatHaz is to decide on the analysis type: probabilistic spatial, spatio-temporal or volcanic hazard assessment. A full probabilistic spatio-temporal volcanic hazard assessment includes a base step and three sequential steps, each of which has to be run separately. The base step (called here 'Step 0'; Fig. 1a), loads the file *MatHaz0.m* and generates several text files (.txt extension) based on the volcano-structural data provided by the user, which then have to be loaded individually into the statistical program R and their results noted in the file *MatHazR.m*. The first step ('Step 1'; Fig. 1b) loads the file *MatHaz1.m*, reads the file *MatHazR.m*, and produces a spatial density map by estimating the probability of vent opening. The second step ('Step 2'; Fig. 1c) loads the file *MatHaz2.m* and performs a temporal probability analysis, generating a spatio-temporal map showing the spatial probability of vent opening within a specific time interval. The third step ('Step 3'; Fig. 1d) loads the file *MatHaz3.m* and models separately PDCs, ballistic projectiles, lava flows, lahars, and tephra fallout based on the

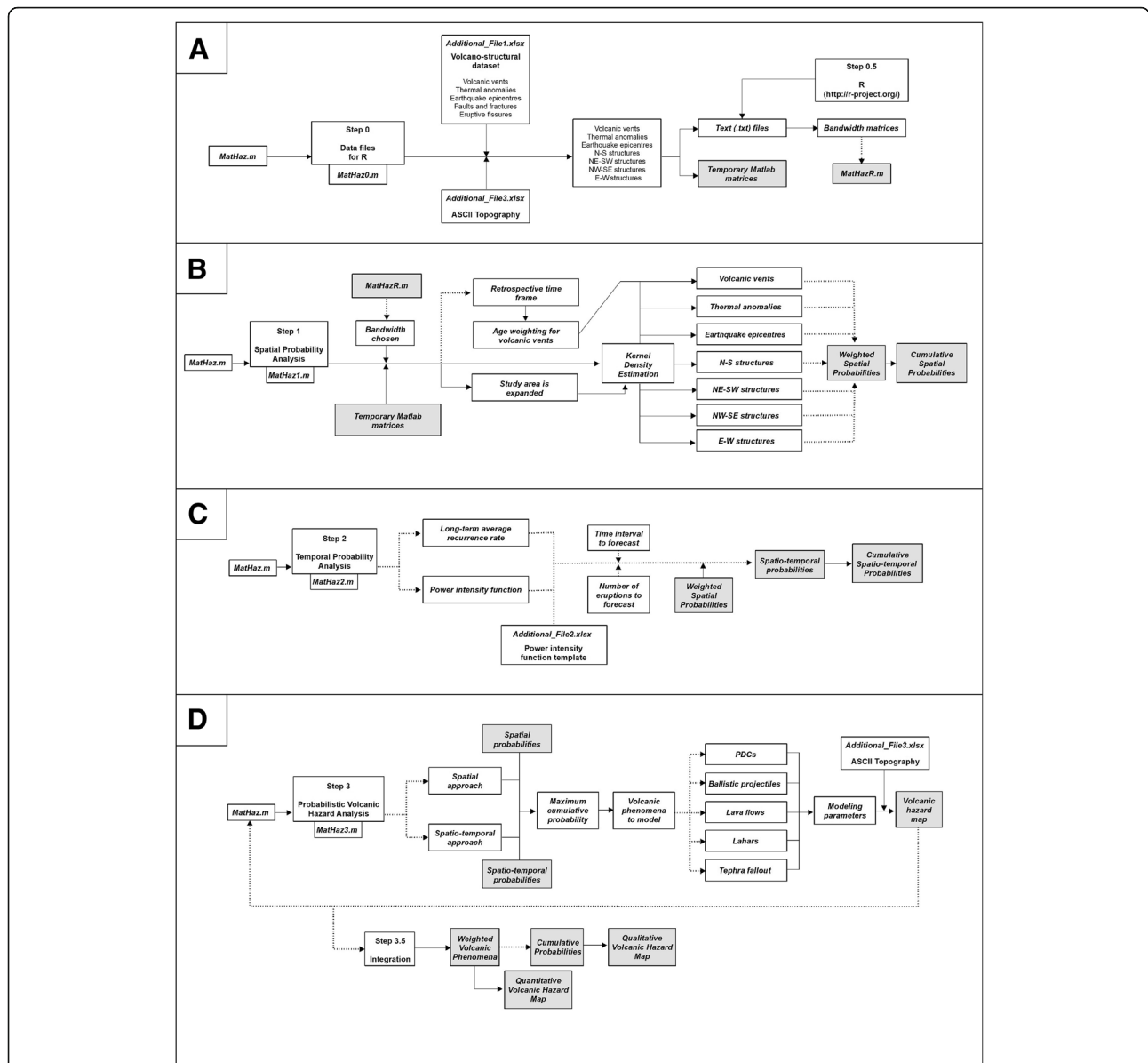


Fig. 1 Flowchart depicting the procedure for a full probabilistic spatio-temporal volcanic hazard assessment. **a** Step '0' (Data files for R) and Step '0.5' (R). **b** Step '1' (Spatial probability analysis). **c** Step '2' (Temporal probability analysis). **d** Step '3' (Probabilistic volcanic hazard analysis) and Step '3.5' (Integration). White rectangles show all the commands, files and products/subproducts created for each step. Grey rectangles depict the outputs of each step that will be used as inputs for any subsequent step. Solid lines with arrows are shown where the program performs the routine automatically. Dashed lines with arrows are shown where the user has to define the input parameters

spatio-temporal analysis. Finally, these results can be merged, quantitatively and/or qualitatively, into a probabilistic volcanic hazard map showing the probability that a specific area is affected by any volcanic process over a specific time interval.

In MatHaz, the spatial density analysis is based on the kernel density estimation method, whose theoretical framework is found in [Appendix 1](#).

This computer code is tested using the Carrán-Los Venados Volcanic Field in southern Chile. For this

example, five volcano-structural datasets were acquired from the geological knowledge of the area (Additional file 1). Local topography was based on an Advanced Spaceborne Thermal Emission and Reflection Global Digital Elevations Model v2, with a spatial resolution of ~30 m for the study area and a vertical accuracy of 17 m within the 95% confidence interval (Gesch et al. 2011), from which a topographic matrix of 50 × 40 km with a pixel size of 100 m was extracted (Additional file 3).

The following section presents a brief geological summary of the Carrán-Los Venados Volcanic Field, focused on its eruptive record as well as the derivation and justification of each volcano-structural dataset used here. The subsequent section describes the application of MatHaz to this volcanic field.

Case-study: Carrán-Los Venados volcanic field

Geological background

The Carrán-Los Venados Volcanic Field (centered at 40.37°S, 72.14°W; 300 m a.s.l.), hereafter CLV, is a volcanic cluster located in Chile, in the central segment of the Southern Volcanic Zone (SVZ) of the Andes (Fig. 2).

CLV consists of at least 64 scoria cones and maars, with ages from the Upper Pleistocene to three recent events in 1907, 1955, and 1979. Vents are aligned in a general ENE orientation, and some are associated with 14 to <1 km-long lava flows and/or up to 6 km-long PDC deposits, spanning an area of about 160 km².

The local basement consists of Paleozoic to Miocene intrusive, metasedimentary, and volcano-sedimentary rocks (Campos et al. 1998; Rodríguez 1999). The northern boundary of the CLV is marked by the Los Guindos stratovolcano, and the southern boundary by products from the Mencheca, Cordillera Nevada, and Cordón Caulle volcanic complexes, all Mid-to-Late Pleistocene in age (Campos et al. 1998; Lara and Moreno 2006). The early history of the CLV is poorly known, since much of the area was covered by glaciers as recently as ~17 ka (Moreno et al. 2015), so the volcanic record is more reliable in post-glacial times (Bertin et al. 2018; Bertin and Moreno *in press*). The earliest product of the CLV is a basaltic lava that fills the Nilahue River valley for about 14 km, reaching Lake Ranco (Bertin et al. 2018). After this effusive episode, at least 70 basaltic to basaltic andesitic eruptive events have been identified, producing tephra fallout, pyroclastic surges, ballistic ejecta, rain-triggered lahars, and/or short (up to 3 km long) lava flows (Bertin et al. 2018; Bertin and Moreno *in press*).

The CLV area is cut by the main trace of the NNE-trending Liquiñe-Ofqui Fault Zone (LOFZ). This is a 1200 km-long, transpressional dextral strike-slip fault system that dominates the intra-arc region between 38° and 47°S (Cembrano et al. 1996; Rosenau et al. 2006). This fault zone has strongly influenced volcanic activity for the last ~6 Ma (Cembrano and Lara 2009). Further local-scale NW-striking fault zones are documented in this area (Campos et al. 1998; Lara et al. 2006), interpreted as crustal-scale weaknesses associated with ancient faults that reactivated as sinistral-reverse strike-slip faults during arc development (Lara et al. 2006; Rosenau et al. 2006; Glodny et al. 2008; Lange et al. 2008). In addition, an ENE-trending system of small faults is inferred based on morphologic lineaments (Bucchi et al.

2015). Although the kinematics of these features are not well constrained, the orientation of the ENE system suggests extension under the current stress field (cf. Cembrano and Lara 2009). The three modern vents are located close to where these ENE lineaments intersect the main trace of the LOFZ (Bucchi et al. 2015).

Volcano-structural datasets

Volcanic vents

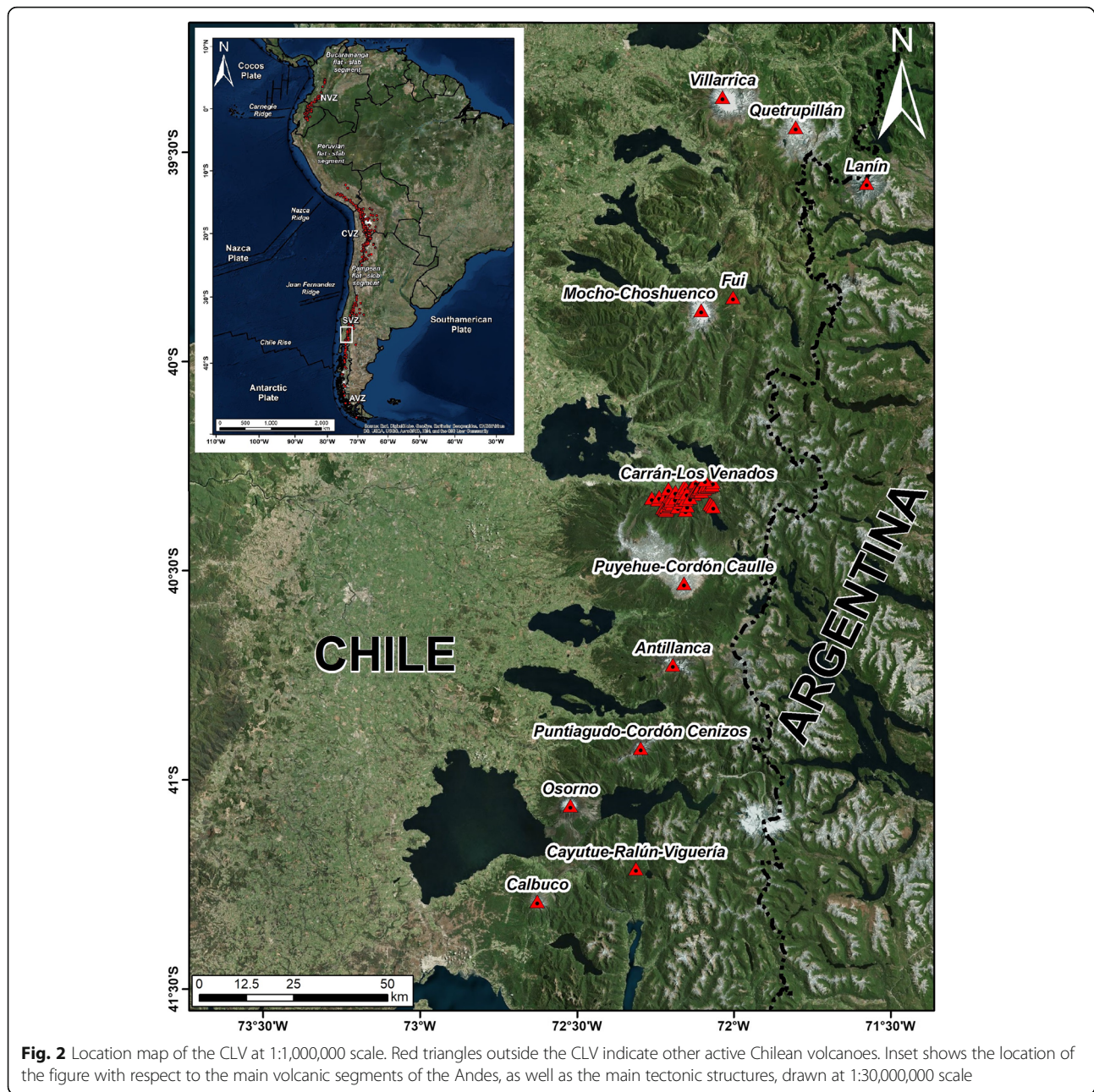
Sixty four volcanic vents are identified in the CLV (Fig. 3a; Bertin et al. 2018; Bertin and Moreno *in press*), mainly consisting of scoria cones and maars. In parallel, up to 70 eruptive events, mainly strombolian and phreatomagmatic, have been described to date and proposed to be related to CLV activity (Bertin et al. 2018). This suggests that some vents may have either erupted more than once and/or been buried by subsequent eruptions (e.g., Wetmore et al. 2009), by sedimentation (e.g., George et al. 2015), or overlooked during mapping. Due to lack of further evidence, it was assumed in our test case that every volcanic vent represents a single event.

Twenty nine eruptive events attributed to the CLV are ¹⁴C dated (Bertin et al. 2018, and references therein). The best-exposed eruptions have been tentatively linked to specific vents based on their distribution and facies (Bertin et al. 2018). We combined these data with a preliminary morphological assessment to make a first-order estimation of the absolute age of every vent (cf. Nieto-Torres and Martin Del Pozzo 2019). Hence, every vent was assigned an absolute age (with no age uncertainties) noted as years before 2019. Although we acknowledge that a more sophisticated treatment of the age data (e.g., including age uncertainties; Connor et al. 2013) should be followed if a probabilistic hazard assessment is envisioned, this approach was deemed adequate for the purposes of our application.

Thermal anomalies

In active volcanic areas, the presence of hot water ponds, heated ground and/or fumarolic activity can reflect heating of shallow aquifers from magma at depth (Goff and Janik 2000; Eppelbaum et al. 2014). These geothermal fluid sources can be further disturbed by shallow magmatic intrusions, leading to phreatomagmatic or phreatic eruptions (Germanovich and Lowell 1995; Valentine et al. 2014). The ample evidence for phreatomagmatic eruptions at CLV (Bertin et al. 2018) suggests that the identification of thermal anomalies can be used to estimate the location of future volcanic activity (Bartolini et al. 2014).

Geothermal field studies conducted in the zone have identified a single area of hot water ponds (up to 60°C), located a few kilometres east of the main trace of the LOFZ (Lemus et al. 2015, and references therein). In



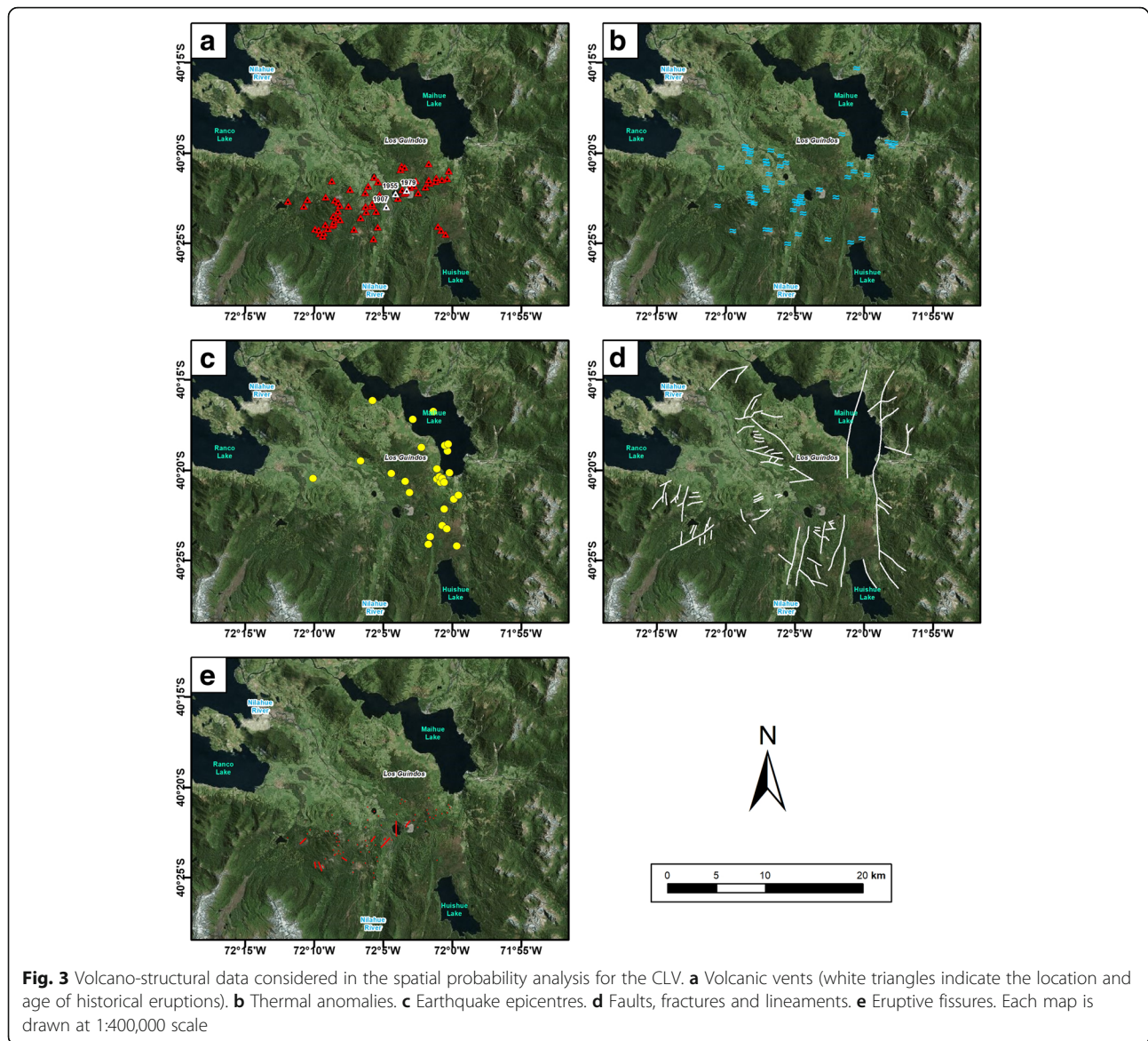
parallel, thermal infrared remote sensing analyses based on satellite imagery (spatial resolution of ~ 90 m) have found that this whole volcanic region has thermal anomalies with temperatures higher than 3°C (Lemus et al. 2015). Taking this into account, only those pixels with anomalies in temperature higher than 10°C were considered in a recent hazard assessment by Bertin et al. (2018) and used in this study, which results in 52 hotspots (Fig. 3b).

Earthquake epicentres

Seismic analyses at CLV started in 2007 with initially four broadband stations (two near-field and two far-

field). In 2011 new stations were added due to the explosive eruption of the neighbouring Cordón Caulle volcano (Bertin et al. 2015). Currently, eight near-field and two far-field seismic stations monitor in real time this volcanic field. According to the Chilean Geological and Mining Survey (SERNAGEOMIN), five volcano-tectonic events per month have been recorded in average, with local magnitudes (M_L) of up to 3.1 and depths between 2 and 10 km (Bertin et al. 2018).

Seismic events are mainly concentrated along the main trace of the LOFZ and some other lineaments, suggesting that they are caused by brittle failure or shear



fracture along tectonic fault planes. As these structures represent zones of weakness, they can be potential magma pathways, especially if they are seismically active (Cembrano and Lara 2009). Consequently, volcano-tectonic events can highlight those fault segments that are active at present (Cappello et al. 2012). In our CLV case, the epicenters of all $>M_L$ 0.1 seismic events with depths < 5 km and recorded after 2010 were selected for our application, which results in 27 events (Fig. 3c).

Faults, fractures and lineaments

Every fault, fracture and lineament on previously published geological maps (Campos et al. 1998; Lara and Moreno 2006; Bertin and Moreno *in press*) was collated into a central database of 82 structures (Fig. 3d). Structure lengths vary between 300 m and 19 km and follow

three main orientations: N-S, NE-SW, and NW-SE. The N-S-trending structures form part of the transpressional dextral strike-slip LOFZ, the NW-SE-trending features are sinistral-reverse strike-slips faults, and the NE-SW-trending structures are normal faults with a strike-slip component, i.e., transtensional faults.

Eruptive fissures

Eruptive fissures were mapped by classifying volcanic vents with the morphometric criteria of Corazzato and Tibaldi (2006). In our test case it was assumed that overlapping and superimposed cones are the surficial expression of eruptive fissures (Corazzato and Tibaldi 2006, and references therein), which may display the same direction as their related feeder dikes (Becerril et al. 2013; Tadini et al. 2014). Field mapping indicates that long-

runout lava flows filling the Nilahue River valley were fed by fissures (Bertin et al. 2018; Bertin and Moreno in press). In addition, the 1979 eruption began as a fissure (Moreno 1980). Collectively, 80 eruptive fissures are identified, most of which trend E-W (Fig. 3e).

All volcano-structural datasets were processed in ArcMap with a Universal Transverse Mercator (UTM) projection. For point data (i.e., volcanic vents, thermal anomalies, and earthquake epicentres), northing and easting coordinates were extracted. For line data (i.e., structures and eruptive fissures), starting and ending coordinates were noted. Finally, all these layers were then exported to an Excel file in which each dataset was assigned to a separate sheet (Additional file 1).

Method and results

MatHaz needs two Excel files (.xlsx extension) as inputs. The initial file has to contain as many sheets as there are volcano-structural datasets; with separate sheets required for volcanic vents, faults, fissures, and any other spatial feature considered important (Additional file 1). The second file has to be a Digital Elevation Model written in ASCII (American Standard Code for Information Interchange) format (Additional file 3).

Step 0: data files for R

The base step of MatHaz loads the volcano-structural file and generates text files based on these data. For the processing of point data, the code reads and rewrites each sheet as a text file. For the processing of line data, the program calculates both lengths and azimuths of every line. Later on, it defines a parameter called ds , which is the pixel size of the topographic matrix. It then calculates a variable called $steps$ by dividing every line's length by ds , and defines dx and dy based on every line's unit projection. All these calculations can be written as:

$$d_l = \sqrt{(x_1 - x_n)^2 + (y_1 - y_n)^2} \tag{1}$$

$$\theta_l = \text{atan}\left(\frac{x_n - x_1}{y_n - y_1}\right) \tag{2}$$

$$\Delta = \left\lceil \frac{d_l}{ds} \right\rceil \tag{3}$$

$$dx = |\sin\theta_l| \tag{4}$$

$$dy = |\cos\theta_l| \tag{5}$$

where (x_1, y_1) and (x_n, y_n) are a line's starting and ending coordinates, respectively, d_l and θ_l are a line's length and azimuth, respectively, ds is the pixel size, Δ is the $steps$ parameter (where $\lceil \cdot \rceil$ is the nearest integer function), and dx and dy are the line's unit projection.

Once these parameters are obtained for every line, the program performs the following iteration:

$$x_{m+1} = \begin{cases} x_m + dsdx \Leftrightarrow x_m < x_n & \Leftrightarrow 1 \leq m \leq \Delta \\ x_m - dsdx \Leftrightarrow x_m > x_n & \end{cases} \tag{6}$$

$$y_{m+1} = \begin{cases} y_m + dsdy \Leftrightarrow y_m < y_n & \Leftrightarrow 1 \leq m \leq \Delta \\ y_m - dsdy \Leftrightarrow y_m > y_n & \end{cases} \tag{7}$$

where (x_m, y_m) are the starting coordinates of the m -th segment of a line.

The above procedure is performed in order to split every line in a consistent way throughout the entire dataset. That is, the longer the line is, the more segments it is split into, and the more it will contribute to the final kernel. Faults, fractures, lineaments and fissures are managed in the same way, and they are grouped according to their azimuths into N-S, NE-SW, NW-SE, and E-W structures. Therefore, the 'Step 0' produces at most seven text files. The user can edit the code to include any other spatial feature considered important and/or to remove those that are absent or not relevant. All these datasets are also saved as matrices for the first step of the program.

Step 0.5: R

The performance of a kernel density estimation strongly depends on the bandwidth selection (Duong 2005). In the bi-dimensional case, the bandwidth matrix controls both the degree and direction of smoothing (i.e., orientation and rate of change of the spatial density with distance from events), so its selection is not straightforward. A widely used way for finding an optimal bandwidth matrix is through the R package 'ks' (<https://www.r-project.org>; Duong 2018, and references therein).

There are many methods to estimate an optimal bandwidth matrix, which can be grouped into normal-scale (\hat{H}_{NS}), normal-mixture (\hat{H}_{NM}), plug-in (\hat{H}_{PI}), and cross-validation (least-squares, \hat{H}_{LSCV} ; biased, \hat{H}_{BCV} ; smoothed \hat{H}_{SCV}) selectors (Duong 2018, and references therein). An important issue is how fast they converge to the AMISE-optimal bandwidth matrix, which depends on both the sample size and the dimension of the matrix (Duong 2005). As the sample size is fixed for every dataset and its dimension is 2, the fastest bandwidths in terms of convergence are: \hat{H}_{PI} with AMSE pilot (also called $\hat{H}_{PI,AMSE}$), \hat{H}_{PI} with SAMSE pilot (also called $\hat{H}_{PI,SAMSE}$), and \hat{H}_{SCV} (Duong 2005). To obtain the $\hat{H}_{PI,AMSE}$ bandwidth, for example, the routine in R is:

```
library(ks)
data <- read.table("text file location")
```


bandwidth < -Hpi(x = data, pilot = "amse")

show (bandwidth)

This routine must be run for every text file. Here the $\hat{H}_{PI,AMSE}$, $\hat{H}_{PI,SAMSE}$, and \hat{H}_{SCV} bandwidths were calculated and their results noted into the *MatHazR.m* file. The user is free to choose any bandwidth estimator, but those above are recommended.

Step 1: spatial probability analysis

Once the *MatHazR.m* file is populated, the user can conduct a spatial probability analysis of vent opening by running MatHaz again. Firstly, the user is asked what bandwidth estimator will be used. Then, for the vent data, a retrospective time frame (T_r) is specified. What the code does with this parameter is to give weights (w_k) to every vent k based on its estimated age (t_k) as follows:

$$w_k = e^{-\frac{t_k}{T_r}} \tag{8}$$

T_r is defined by default as the age of the oldest vent, but the user may choose another constant and another weighting function just by editing the code. Equation (8) means that the weight assigned to historical events is close to 1 (t_k close to 0), and the older the vent, the lower its weight and the less it will contribute to the final kernel.

The spatial analysis is based on a Gaussian kernel, which has infinite support, so the resulting PDF reaches zero only at infinity. However, this function decreases rapidly enough to be negligible at large distances, so it can be cropped to a finite domain (Duong 2005; Germa et al. 2013; Bebbington 2015; Connor et al. 2018). Thus, the program increases the study area by ~900%, which ensures that the resulting PDF will decrease towards zero at its edges. Then, the program creates a $m \times n$ grid for this expanded domain using the pixel size ds , and every coordinate obtained at the end of the ‘Step 0’ is adjusted to the pixel size.

MatHaz works with an expanded expression of Equation (34) (see Appendix 1 for derivation), rewritten to show the value of the kernel function in any point of a $m \times n$ grid for each k component of a bivariate sample of size S_q (q being the dataset; e.g., volcanic vents, E-W structures, etc.). That is:

$$\widehat{f}_H^q(r_k) = \frac{(ds)^2 |H|^{-\frac{1}{2}}}{2\pi} \sum_{j=1}^n \sum_{i=1}^m e^{-\frac{1}{2}(x_i-x_k; y_j-y_k)^T H^{-1} (x_i-x_k; y_j-y_k)} \tag{9}$$

Where x_i and y_j are the northern and western coordinates of every segment of the study area, respectively. The $(ds)^2$ term is a normalizing factor added because every coordinate was adjusted to a grid of pixel size ds ,

so the integral of this function across the whole area is quite close to one for every k .

Next, the program performs the following:

$$\widehat{f}_H^q(r) = \begin{cases} \frac{1}{S_q} \sum_{k=1}^{S_q} \widehat{f}_H^q(r_k) w_k \Leftrightarrow q = \text{Vents} \\ \frac{1}{S_q} \sum_{k=1}^{S_q} \widehat{f}_H^q(r_k) \Leftrightarrow q = \text{Any other dataset} \end{cases} \tag{10}$$

Due to the assignment of different weights w_k for each k vent, the resulting PDF has to be normalized. The program performs all these steps automatically and the user chooses the graphic output.

Subsequently, the final PDF is defined as a linear combination of the contribution of every dataset as follows (e.g., Martí and Felpeto 2010; Cappello et al. 2012; Becerril et al. 2013; Bartolini et al. 2014; Bevilacqua et al. 2015; Galindo et al. 2016):

$$\widehat{f}_H(r) = \sum_{q=1}^{S_t} \widehat{f}_H^q(r) w_q \tag{11}$$

where

$$\sum_{q=1}^{S_t} w_q = 1 \tag{12}$$

with S_t the total volcano-structural datasets grouped during ‘Step 0’ and w_q the weight assigned to the q -th volcano-structural dataset.

In our CLV application, the weights w_q were chosen by assigning a relatively higher importance to volcanic vents, N-S structures, and NW-SE structures (Table 1). These features have been considered as the most relevant for controlling the distribution of the volcanism in the region (Cembrano and Lara 2009). The results of equations (10) and (11) are shown as probability isocontours in Figs. 4a-g and Fig. 4h, respectively.

Step 2: temporal probability analysis

Equation (11) is an estimate of the spatial density. That is, it assumes that at least one eruptive event has already taken place, and estimates its probability of occurrence throughout the study area. However, if a spatio-temporal assessment is envisioned, then the temporal probability of the occurrence of a new volcanic event has to be included. There are two ways for doing this:

$$\lambda(r, \Delta t) = \widehat{f}_H(r, \Delta t) \tag{13}$$

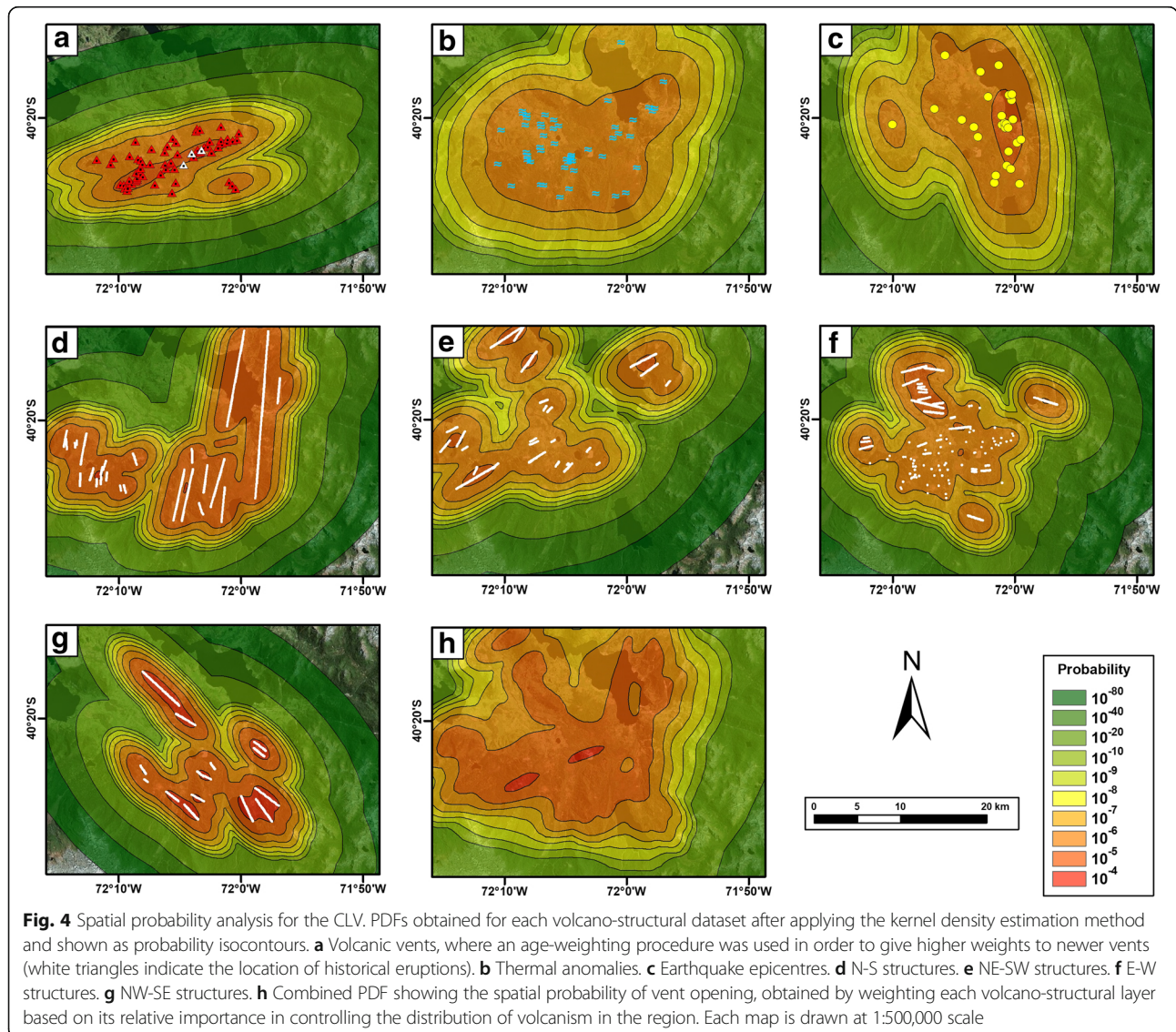
or

Table 1 Weights w_q for each volcano-structural dataset, chosen according to what features have been proposed as most relevant for controlling the distribution of the volcanism in the region

Dataset	w_q
Volcanic vents	0.45
Thermal anomalies	0.04
Earthquake epicentres	0.06
N-S structures	0.20
NE-SW structures	0.14
NW-SE structures	0.03
E-W structures	0.08

$$\lambda(r, \Delta t) = \widehat{f}_H(r)\rho(\Delta t) \tag{14}$$

Where $\lambda(\cdot)$ is the space-time varying intensity function, $\rho(\cdot)$ is a point process intensity function, and Δt is the elapsed time to be forecasted. Equation (13) is used if there is a temporal term explicit at some step during the spatial analysis, and assumes that both spatial and temporal components of the analysis are mutually dependent. Equation (14), on the other hand, assumes that both the spatial and temporal components are independent, so they can be evaluated separately and then multiplied by each other. The latter procedure is based on the common assumption in volcanology that spatial and temporal components can be considered separately (e.g., Bebbington and Cronin 2011; Bebbington 2013; Connor et al. 2013); this assumption will also be made



here, so the focus will be on the different ways of calculating the $\rho(\cdot)$ function.

Given that $\rho(\cdot)$ is defined as a point process intensity function, then a $\phi(\cdot)$ function has to be defined in advance. Such a function can be expressed as a long-term average recurrence rate, and is based on either the number of eruptions over a given observation period, or the number of repose times over a long time interval (Ho et al. 1991). In the second case, this is written:

$$\phi(\Delta t) = \frac{N-1}{t_1-t_N} \tag{15}$$

Where N is the total number of eruptions contained between t_1 and t_N , inclusive, with t_1 being the age of the oldest known eruption and t_N the age of the youngest known eruption. Note that equation (15) does not depend explicitly on Δt , meaning that the eruption rate is considered to be constant throughout the history of the volcanic field, so the volcano record is represented by a simple Poisson process. This inference is questionable since eruption frequencies can vary strongly over time (e.g., Bebbington and Cronin 2011; Leonard et al. 2017; Damaschke et al. 2018), but a Poisson model is a sensible first-order approach in volcanic areas with long eruptive histories and/or patchy eruptive records (e.g., Connor et al. 2013; El Difrawy et al. 2013; Runge et al. 2014; Bartolini et al. 2015; Gallant et al. 2018; Nieto-Torres and Martin Del Pozzo 2019).

On the other hand, if reliable geochronological and/or historical data are available, then more sophisticated estimators can be used, both stationary and non-stationary (Bebbington 2012). In the second case, a useful non-stationary model is a nonhomogeneous Poisson process either known as a Weibull process, power intensity function or power law process (Smethurst et al. 2009; Cappello et al. 2013; Connor et al. 2015), which is written:

$$\phi(t_e) = \frac{\delta}{\theta} \left(\frac{t_e}{\theta}\right)^{\delta-1} \tag{16}$$

Where δ and θ are to-be-determined positive parameters, and t_e the year of interest counted from the age of the oldest known eruption (that is $t_e = 0$ at t_1 and $t_e = t_1$ at present). Both δ and θ parameters can be optimally obtained by minimizing the following residual:

$$\sum_{i=1}^N [\phi(t_i) - \chi(t_i)]^2 \tag{17}$$

where

$$\phi(t_i) = \int_0^{t_i} \phi(t) dt \tag{18}$$

with N being the total number of eruptions, and $\phi(\cdot)$ and

$\chi(\cdot)$ the virtual and actual cumulative number of eruptions observed up to t_i , respectively. The steps listed here have been compiled in an Excel template (Additional file 2).

Having obtained the function $\phi(\cdot)$, the function $\rho(\cdot)$ is modeled as a Poisson point process, which can be either homogeneous (if equation (15) is used) or nonhomogeneous (if equation (16) is used), and can be generally described as:

$$\rho(n_e, \Delta t) = \frac{[\Lambda(\Delta t)]^{n_e}}{n_e!} e^{-\Lambda(\Delta t)} \tag{19}$$

where $\Lambda(\Delta t)$ is the estimated number of eruptions for a forecasting time interval Δt , that is:

$$\Lambda(\Delta t) = \int_{t_1-t_N}^{t_1-t_N+\Delta t} \phi(t) dt \tag{20}$$

Equation (19) calculates the probability of n_e eruptions at some location for a forecasting time interval Δt , attaining its maximum at:

$$\Delta t_{max} = \Lambda^{-1}(n_e) \tag{21}$$

Where $\Lambda^{-1}(\cdot)$ is the function that reverses $\Lambda(\cdot)$. Integration limits in Equation (20) imply that the forecasting time interval is measured from the age of the youngest known eruption (t_N).

If n_e is set at zero (i.e., there are no eruptions for the forecasting time interval), equation (19) becomes:

$$\rho(0, \Delta t) = e^{-\Lambda(\Delta t)} \tag{22}$$

In contrast, if this is not true, then:

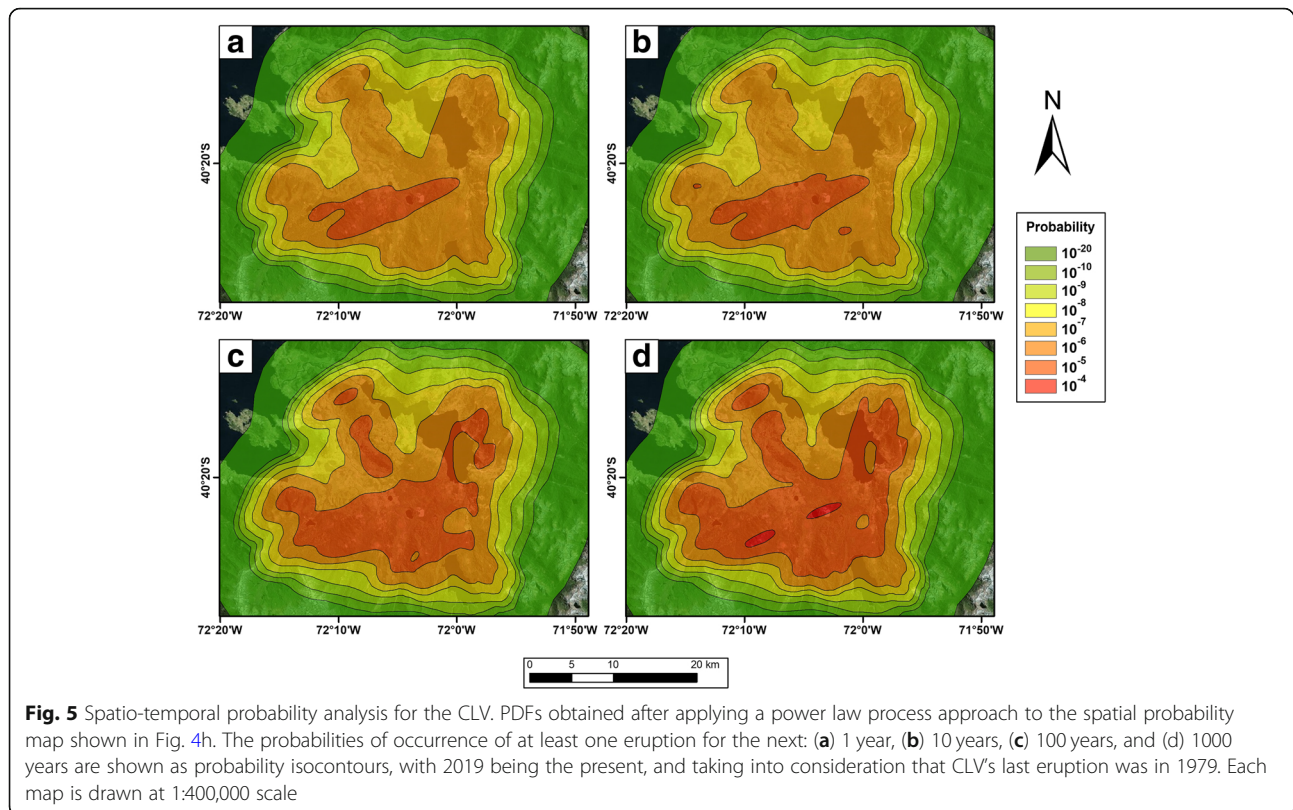
$$\rho(n_e \geq 1, \Delta t) = 1 - e^{-\Lambda(\Delta t)} \tag{23}$$

Equation (23) calculates the cumulative probability of at least one eruption at some location for the forecasting time interval Δt .

Regardless of which option is considered, MatHaz calculates the spatio-temporal probability of n eruptions (or at least one eruption) at r for a forecasting time interval Δt (being r defined as any (x, y) pair of pixel size ds).

Likewise, as was done with its temporal counterpart, an integration of $\widehat{f}_H(r)$ in space is also an option, although a common procedure in volcanology is to keep the area as unitary (defined by the pixel size ds) and to vary the forecasting time interval instead (e.g., Cappello et al. 2013; Connor et al. 2013).

MatHaz allows the user to decide which methodology, forecasting time interval, and number of eruptions will be modeled. In our case study application, equations (16) and (23) are used as an example, modeling the occurrence of at least one eruption in the CLV for the next 1, 10, 100,



and 1000 years (Figs. 5a to d, respectively). As it is noted, the greater the time interval to be forecasted, the more similar the resulting figure (Fig. 5d) is to that obtained at the end of the spatial probability analysis (Fig. 4h). These results can also be integrated for a specific area. In our test case, the whole study area is considered, resulting in probabilities of 24.7%, 29.8%, 62.5%, and 99.9% for the four time periods forecasted. This means that the probability of at least one eruption happening somewhere in the study area during the next year is 24.7%, which increases to 29.8% for the next decade, and so on.

Step 3: probabilistic volcanic hazard analysis

Once the spatio-temporal assessment is completed, the user can perform a probabilistic volcanic hazard analysis by running the code once again. MatHaz works with analytical, semi-empirical and empirical models published in the literature, and each model is run as many times as there are pixels in the study area (i.e., every pixel is a potential vent). Despite this very convenient assumption, the total number of pixels (i.e., $m \cdot n$) can still be too large for manageable simulation times. Consequently, at the beginning of this third step the program asks the user what maximum cumulative probability will be used, picking out only those pixels with the highest spatio-temporal probabilities, so the closer this number is to 1, the more pixels will be

selected. In the example of the CLV followed throughout this work, cumulative probabilities of 0.9, 0.99, and 0.999 are reached at 5.5%, 9%, and 11.3% of the total number of pixels, respectively (Fig. 6).

After this substep, the program reloads the topographic matrix of the area and asks the user what volcanic phenomena will be modeled. In this first version of MatHaz, the following phenomena are considered: PDCs, ballistic projectiles, lava flows, lahars, and tephra fallout. Debris avalanches are absent due to the very low probability of occurrence; nevertheless, they can be included in an updated version of this tool. This is discussed further in the next section.

PDCs (which include pyroclastic flows, pyroclastic surges, blasts, and block-and-ash flows) are simulated via a simple energy cone model (e.g., Malin and Sheridan 1982). Input parameters for this model are the collapse height h_c and the ratio between $h_v + h_c$ and L , with h_v being the elevation of the chosen vent and L the run-out length of the phenomena. The model generates a cone with a slope m_v (defined by the aforementioned ratio) with its apex located h_c meters over the vent. The intersection of this cone with the topography defines the area that is prone to be affected by the phenomena (Malin and Sheridan 1982). Both h_c and L can range widely, mainly depending on the type and volume of the current (Newhall and Hoblitt 2002, and references therein).

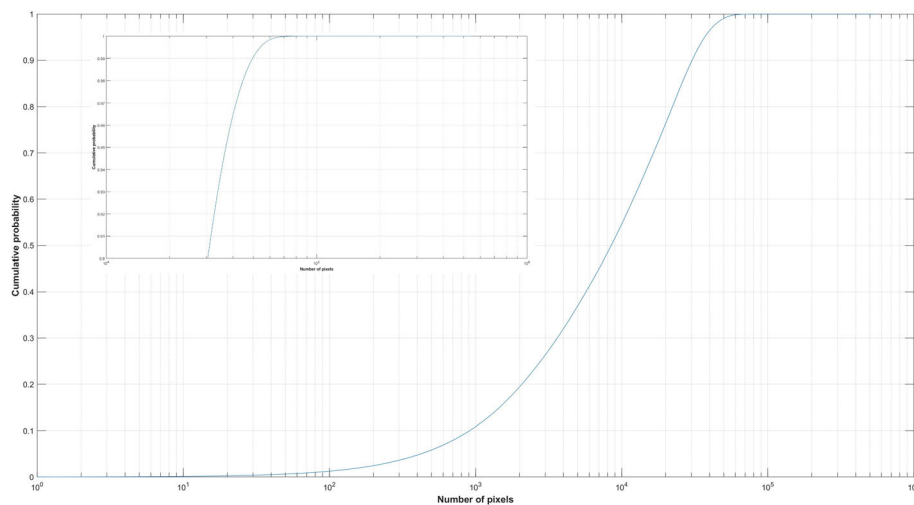


Fig. 6 Semi-log plot showing the cumulative probability curve versus the number of pixels considered. Inset shows that cumulative probabilities of 0.9, 0.99, and 0.999 are reached at 5.5%, 9%, and 11.3% of the number of pixels with the highest probabilities

There are many compilations of these parameters worldwide (e.g., Sheridan and Malin 1983; Hayashi and Self 1992; Freundt et al. 2000; Saucedo et al. 2005), so the user can choose them based on analogues of what is expected in the study area. At CLV there is evidence of at least 25 base surges, with run-outs of up to 6 km and deposits evenly distributed around the vents. These are mainly associated with phreatomagmatic eruptions, although a few column-collapse scoria flow deposits generated during strombolian eruptions are also present (Bertin et al. 2018). Taking into account all these data, collapse heights (h_c) between 100 and 300 m, and slopes (m_v) between 0.1 and 0.3 were defined in our test case.

Ballistic projectiles are modeled based on an analytical solution of the equations of movement of a dense projectile in the atmosphere (Bertin 2017). This model simulates a two-dimensional movement of an ellipsoidal particle under no wind conditions assuming a constant drag coefficient. Its input parameters are the projectile's density (ρ_S), semi-axes (d , e , f), launch velocity (V_0), ejection angle (θ_0), and drag coefficient (C_D), whose values can range widely (Bertin 2017, and references therein). The density of the atmosphere is assumed constant and calculated at the vent. It is also assumed that there is no preferred launch azimuth, so all plausible angles are considered. Where these results intersect with the topography contours are drawn, with the area enclosed by this contour defined as the zone prone to be affected by ballistic projectiles with decreasing launch velocities. At CLV, ballistic projectiles are recognized both in lapilli fallout and pyroclastic surge deposits, at distances of up to 2 km from the vents (Bertin et al. 2018). Inverse modeling parameters found for these projectiles include: bulk densities ρ_S 1500–2500 kgm⁻³, semi-axes d , e , and f

between 0.2 and 1 m, launch velocities V_0 between 100 and 200 ms⁻¹, ejection angles between 30 and 60°, and drag coefficients between 0.2 and 1.

Lava flows are modeled on the basis of their areal extent, slightly modified from the approach in Connor et al. (2012). Such a model is ruled by simple algorithms, whose input parameters are the total lava volume (V_T) and the lava unit thickness (h_b) emitted for a pixel. The unit blocks of lava are sequentially emitted from the same pixel until their cumulative volume reaches V_T (Connor et al. 2012). After a block is created, the elevation of the source pixel and of its eight neighbors are compared. If the lowest elevation pixel + h_b is lower than the elevation of the source pixel, then the block is transferred to that pixel, otherwise the elevation of the source pixel is increased by h_b . This procedure is repeated while the block is transferred from one cell to another, then, the former and latter topographies are compared and the affected pixels depicted as the lava inundation zone. The model output strongly depends on h_b , so the smaller this parameter, the greater the run-out lengths and the more intricate the flow, including meanders around small topographic features. At CLV, up to 32 basaltic to basaltic andesite lava flows are identified (Rodríguez 1999), most with a ã morphologies and emitted during Hawaiian style eruptions (Bertin et al. 2018, and references therein). Their lengths vary from 14 to < 1 km and their thicknesses from 5 to 55 m, with the longest and thickest flows related to the oldest activity of the field. During the Holocene, almost every lava flow was < 3 km (Bertin et al. 2018). These data helped to constrain the input parameters used for modeling, resulting in volumes V_T between 0.05 and 0.1 km³, and block thicknesses h_b between 1 and 10 m.

Lahars (which include debris flows and hyperconcentrated streamflows) can be modeled following the same inundation criteria for lava flows. In this case $h_b < 1$ m should be used, and volumes $V_T < 0.2 \text{ km}^3$ are recommended based on the literature (e.g., Vallance and Scott 1997; Gudmundsson 2015). At CLV, lahars are generated by rainfall remobilisation of tephra falls, with the largest known generated in 1907. The products of this eruption dammed the Nilahue River, generating a lake upstream that collapsed several months later, flooding areas up to 30 km away (Bertin et al. 2018, and references therein). However, the lahar deposit record during the history of the CLV is poorly known and generally not linked to vents, thus lahar events were not modeled in our test application.

Tephra fallout was empirically modeled assuming that the deposit thins exponentially away from source and that the isopachs are either circular or elliptical (Pyle 1989). This decay law becomes clear if the natural logarithm of the thickness is plotted against the square root of each isopach area, so a straight line can be fitted using a least squares approach (Pyle 1989; Fierstein and Nathenson 1992). Such a line has two parameters: its slope $-k$ and its y-intercept $\ln(T_0)$, where T_0 is the extrapolated maximum deposit thickness, so the deposit thickness can be estimated as (Pyle 1989):

$$T = T_0 e^{-k\sqrt{A}} \tag{24}$$

If equation (24) is integrated with respect to the area A , the total tephra volume can be obtained (Fierstein and Nathenson 1992):

$$V = \frac{2T_0}{k^2} \tag{25}$$

Equation (25) has been commonly used as a first-order approach to estimate the volume of several explosive eruptions, but it requires a minimum number of isopachs covering a widespread area to confidently define $-k$ and T_0 . At CLV, up to 56 fallout deposits are identified, but volumes are only well-constrained for historical events, which vary between 0.009 and 0.37 km^3 (Bertin et al. 2018, and references therein). Based on these data, several thousands of $(-k, T_0)$ pairs were randomly generated by a Monte Carlo simulation and selected only if the calculated volume was between 0.0001 and 0.5 km^3 . Every chosen pair was then used in the following equation, which relates elliptical isopachs to distance, with the source at one of the foci of the ellipse (Nathenson 2017; see Appendix 2 for derivation):

$$T_{ij} = T \left(r_{ij}, \phi_{ij} \right) = T_0 e^{-k r_{ij} [1 - \varepsilon \cos(\phi_{ij} + \phi_0)]} \sqrt{\frac{\pi}{(1 - \varepsilon^2)^{3/2}}} \tag{26}$$

Where ε is the isopach eccentricity, r_{ij} is the distance from the source (x_0, y_0) to any (x_i, y_j) , and ϕ_{ij} is the angle between these two points measured from the ellipse's major axis. The term ϕ_0 allows a rotated ellipse, i.e., its major axis does not coincide with the E-W axis of the study area, and can be considered as a proxy for wind direction (e.g., Bonadonna et al. 2005b; Connor and Connor 2006). The eccentricity ε elongates the deposit downwind and compresses it upwind, and is partly related to how the wind affected the dispersal pattern (e.g., Burden et al. 2013; Mastin et al. 2014). The parameters ε and ϕ_0 do not have any influence on the calculated volume, but they do affect both the shape and the orientation of the isopachs. In our test case, eccentricities ε between 0 and 1, and rotation angles ϕ_0 between 200 and 340° (measured clockwise from north) were defined. The latter parameter reflects a predominantly westerly wind in the region (Bertin et al. 2018).

In our CLV case study, each model was run only for those source pixels whose probability contribute the most to the 99.9% of the spatio-temporal cumulative probability (that is, 11.3% of the total number of pixels; Fig. 6). The parameters for each model were randomly defined following a Latin-hypercube sampling approach (e.g., McKay 1992; Bertin 2017), with up to n' intervals, with n' being the number of pixels considered in this analysis (i.e., $n' \leq m \cdot n$). For each model and for each pixel, the resulting area was assigned the spatio-temporal probability of its source pixel.

Step 3.5: integration

The areas affected by each volcanic phenomenon are summed in this way:

$$p_h = \sum_{p=1}^{n'} p_p^h \tag{27}$$

where p_p^h is the spatio-temporal probability grid (of size $m \cdot n$) related to the p-th source pixel for the h-th volcanic phenomenon. The final sum p_h is normalized, so its integral throughout the area is one.

These probability grids can be quantitatively combined into a single grid by assigning weighting factors w_h to every volcanic phenomenon as follows:

$$p = \sum_{h=1}^{H} p_h w_h \tag{28}$$

where

$$\sum_{h=1}^{h'} w_h = 1 \quad (29)$$

with h' being the total number of phenomena modeled during 'Step 3' (in this case four) and w_h the weight assigned to the h -th volcanic hazard. The user may specify a value of every weight w_h . In our CLV case study, the weighting values reflected the relative frequency of each volcanic phenomenon according to the geological knowledge in the zone (Table 2).

Results of Equation (27) are shown as probability isocontours for the four volcanic phenomena modeled: PDCs, ballistic projectiles, lava flows, and tephra fallout (Figs. 7a to d, respectively), while results of Equation (28) are shown as probability isocontours of an integrated quantitative volcanic hazard map (Fig. 7e).

Finally, if the results are sorted from largest to smallest, they can be binned as percentiles. For our case study, three boundaries were chosen to represent high (20%), moderate (40%), and low hazard (80%). This zoning is shown as an integrated qualitative volcanic hazard map following a red-yellow sequential colour scheme (Fig. 8), after having been passed through the Color Brewer tool (Brewer et al. 2013).

Discussion and conclusions

Limitations of the code

A significant contribution of MatHaz is its ability to combine several volcanic hazards independently modeled for every source pixel. This procedure required some assumptions in order to achieve manageable simulation times (always a consideration for a single computer simulation). Namely, empirical, semi-empirical, and analytical models are prioritized over physics-based numerical models. That is, unlike other hazard codes, MatHaz is not designed to model the physics of volcanic eruptions and their products, nor to model the details of specific eruptive scenarios. The code is not concerned with the rates of these processes other than the rate of vent formation. However, if process rates need to be tackled, a user could run MatHaz to obtain a spatial or a spatio-temporal probability map and then run more sophisticated numerical models on other platforms based on these results.

Table 2 Weights w_h for each volcanic hazard, chosen according to the relative frequency of each volcanic phenomenon based on its relative abundance in the CLV eruptive record

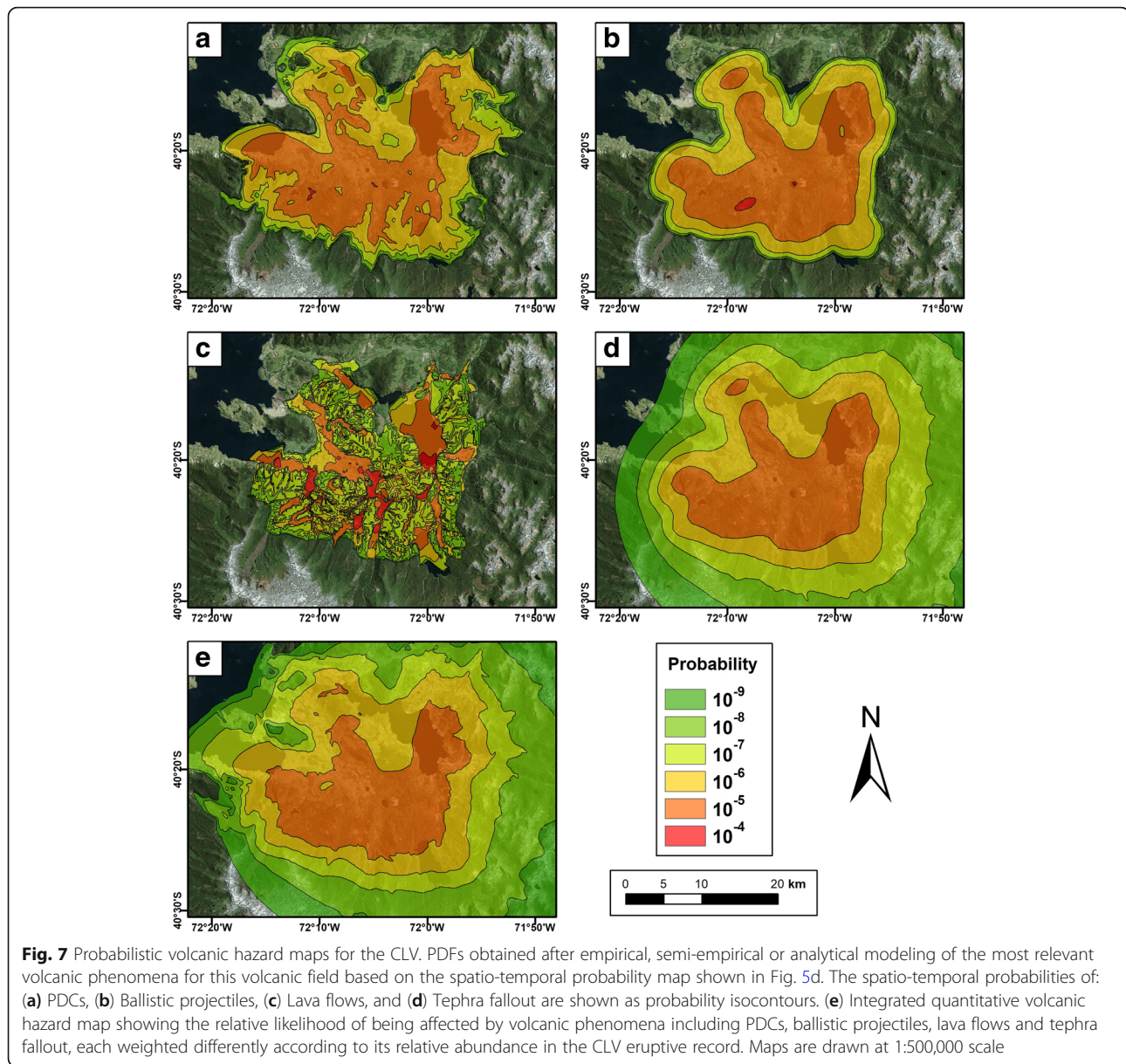
Volcanic hazard	w_h
PDCs	0.20
Ballistic projectiles	0.30
Lava flows	0.10
Tephra fallout	0.40

Even the simplest models considered in MatHaz have their own assumptions. Tephra fallout modeling, for example, was performed by drawing oriented elliptical isopachs assuming a deposit that thins exponentially away from source. The parameters that control this exponential decay (i.e., $-k$ and T_0) were obtained assuming that the isopach data only define a single slope on a $\ln(T)$ versus \sqrt{A} plot. This is a very simple approach and has been proven to not necessarily be true. Indeed, research on several well-documented eruptions has shown different data behavior, e.g., multi-segment exponential curves, power-law curves, and Weibull functions, which in turn have their own parameters (Bonadonna and Costa 2012, and references therein). A potential user can, however, adapt the code to include any of these functions and routines to obtain their optimized parameters.

Alternative semi-empirical models, such as the tephra attenuation model (González-Mellado and De La Cruz-Reyna 2010; Kawabata et al. 2013), can also be incorporated into the code. Numerical modeling of tephra fallout, on the other hand, has already been accomplished in some hazard assessments of distributed volcanic fields, but results have typically been either depicted by isopachs superimposed on the integrated volcanic hazard map (e.g., Alcorn et al. 2013; Bartolini et al. 2014; Becerril et al. 2017), or integrated in some fashion with the other hazards (e.g., Becerril et al. 2014; Bartolini et al. 2015). However, such integration is challenging if the other hazards represent cumulative probabilities of inundation yet the tephra hazard is based on a single eruptive scenario modeled with specific atmospheric conditions. Sophisticated numerical tephra modeling tools are available (e.g., Hurst and Turner 1999; Bonadonna et al. 2005a; Folch et al. 2009; Schwaiger et al. 2012) and are continuously being adapted to perform probabilistic analyses, varying either the atmospheric parameters (Amigo et al. 2012; Sandri et al. 2014) or the eruptive scenarios (Biass et al. 2016), and tested to obtain better-constrained eruptive parameters via inverse modeling (White et al. 2017). Running any of these models on their respective platforms for those vents with the highest spatio-temporal probabilities might be a sensible option.

Other volcanic processes included in MatHaz, such as lahars, are modeled based on a simple 'flow-routing' code. That is, the lahar will always follow the thalweg of the valley and will never overbank. There are, however, more sophisticated tools available to model the dynamics of the flow taking into account mass and momentum, for example, the Titan2D model (Patra et al. 2005), which has proven to be useful to simulate block-and-ash flows as well (Charbonnier and Gertisser 2009).

Debris avalanches (e.g., sector collapses) have not been considered in this first version of MatHaz due to their very low probability of occurrence in volcanic fields,

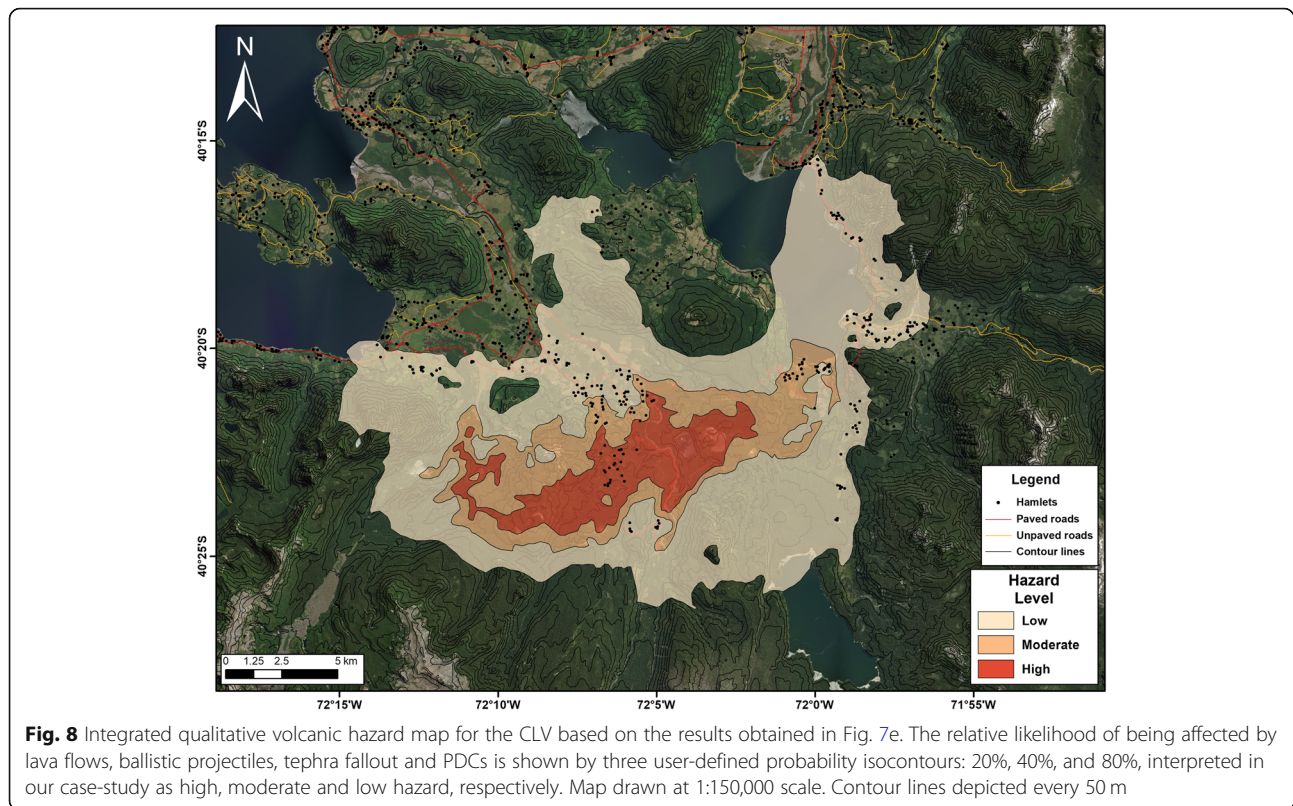


although a model can be incorporated in any subsequent version of the program. A routine for debris avalanche modeling could be developed by simply editing the energy cone tool. Despite its simplicity, the energy cone tool is a good estimator of the extent of debris avalanches worldwide, where an inverse relationship between the slope of the cone (m_v) and the volume of the deposit has been proposed (Ui 1983; Schuster and Crandell 1984; Siebert 1984; Ui et al. 2000). If the user wishes to model this phenomenon, m_v values between 0.04 and 0.18 are recommended (Hayashi and Self 1992; Siebert 1996).

For the modeling methodology, it was assumed that every volcanic phenomenon had its source in a single squared pixel of size ds . This is an oversimplification,

especially for lava flows, since they quite often start erupting from a fissure (e.g., several pixels), which may or may not evolve into a single vent (Valentine and Gregg 2008). In the CLV, there are examples of these fissure eruptions in both the historical and prehistorical record. In MatHaz, simulation of fissure eruptions adds a complexity to the model similar to that which arises when simulating rain-triggered lahars, since many assumptions about the additional source pixels have to be made and its implementation is not straightforward, although for rain-triggered lahars some attempts have been made (e.g., Jones et al. 2017; Mead and Magill 2017).

MatHaz is programmed to run regardless of the input files or parameters given by a potential user. In order to optimize performance of future applications, we



recommend that the user maintain the format and structure of each of the input files used in our test case application (see Additional files 1 and 3). The user can, however, edit the code to enable it to read other file formats and structures if required. Likewise, MatHaz is not programmed to give a warning if the input parameters are unrealistic. However, some guidelines and sensible ranges of parameters are given as recommendations in the volcanic hazards' modeling section as the user goes through this task.

Limitations of the application

In our CLV application it was assumed that every volcanic vent represents a single event. This is not necessarily true for distributed volcanism (e.g., Runge et al. 2014; Bevilacqua et al. 2017; Connor et al. 2018; Gallant et al. 2018), so statistical analyses testing the independence of every volcanic vent in the study area would be desirable. Following this idea, vents classified either as overlapping or superimposed (e.g., Corazzato and Tibaldi 2006) might be assumed to be the surficial expression of magma-feeding fractures (e.g., Becerril et al. 2013; Tadini et al. 2014), and then be related to a single magmatic event (e.g., Nieto-Torres and Martin Del Pozzo 2019). However, considering events rather than vents would mean that the spatial probability analysis will estimate the spatial density of future volcanic events, each of which might produce more

than one volcanic vent (Connor et al. 2018). If an analysis of this type is envisioned, then studies focused on detailed stratigraphic correlations including radiometric dating should be carried out in the zone of interest to attempt to attribute vents to events, although this may be difficult to accomplish (Connor et al. 2018; Gallant et al. 2018). On the other hand if, as was done in our case study, a spatial probability of vent opening is visualized, then each vent has to be considered as a single magmatic event. A restriction of this approach is that, because the model only forecasts the location of the next volcanic vent, the location of any forthcoming vent will be influenced by the location of the previous vents (Connor et al. 2018). This means that the use of equations (16) and (23) for more than one eruption during the spatio-temporal assessment is only partly valid. Nevertheless, that would be true for distributed fields with just a few vents, so the larger the number of vents, the less any subsequent vent will affect the spatial density estimation. In order to evaluate how any future vent affects the spatial probability map, a sensitivity analysis simulating different vent locations should be conducted.

Additionally, for the temporal assessment a power law process was used for the CLV. As a detailed temporal eruption record does not exist for this volcanic field, this approach was made possible by estimating the absolute age (with no age uncertainties) of every vent based on its

morphology and on the ^{14}C ages available. This procedure was performed as an ideal case for comparing the results with those obtained assuming a long-term average recurrence rate (see next subsection), which only needs the total number of eruptions and the ages of the oldest and youngest eruptions. However, more sophisticated estimations can be conducted by including age uncertainties, and running some Monte Carlo simulations to obtain better recurrence rates and confidence intervals (e.g., Bebbington 2013; Connor et al. 2013), or by building a very complete eruption age dataset (e.g., Damaschke et al. 2018).

The volcanic record at CLV shows the occurrence of PDCs, ballistic projectiles, lava flows and tephra fallout. Lahar deposits have also been identified in the CVL, but they are likely to have been generated by remobilization of unconsolidated pyroclastic material during intense rainfall, so lahars triggered by eruptive activity were not modeled in our test application. However, the simulation of lahars generated by a volcanic eruption is an option in MatHaz.

MatHaz has been developed to follow a sequential methodology, meaning that each step is based on the results obtained in previous steps. In our case study, the vent location probability (Fig. 4h) is by far the main factor governing all hazards (Figs. 7 and 8), so if the vent location model is wrong, then the entire analysis will be wrong. To address this issue, some alternative models, bandwidths, and weights should be tested for comparison and to manage uncertainties. Some examples of this are given below.

Model achievements and sensitivity analyses

The model presented here performs an integrated (quantitative and/or qualitative) spatio-temporal volcanic hazard assessment for distributed volcanic fields, which presents some improvements over existing methods. One breakthrough is related to the consideration of different options for calculating bandwidths during the spatial hazard assessment. These bandwidths have commonly been considered as scalars in the literature, with studies that have worked with matrices being quite scarce (e.g., Kiyosugi et al. 2010; Bebbington and Cronin 2011; Connor et al. 2012, 2013; El Difrawy et al. 2013; Bebbington 2015; Galindo et al. 2016; Bevilacqua et al. 2017; Connor et al. 2018). Throughout the CLV example followed in this manuscript, the $\hat{H}_{PI,AMSE}$ selector was considered. Also considered were the selectors $\hat{H}_{PI,SAMSE}$ and \hat{H}_{SCV} , and the area enclosed by the resulting isocontours 10^{-20} , 10^{-10} , and 10^{-5} of their respective final spatial probability maps (i.e., after applying Equation (11)) noted for comparison (Table 3). The highest discrepancies (up to 38.6%) were found when comparing

the area of the isocontour 10^{-20} of the $\hat{H}_{PI,AMSE}$ and \hat{H}_{SCV} selectors, with this being lower (up to 5.1%) at higher probability isocontours. These results suggests that, although the selection of bandwidths follows an automatic and data-driven methodology, different bandwidths may have a significant influence on spatial probabilities, so a sensitivity analysis should always be conducted since the spatial probability map generated in this step forms the basis of both the spatio-temporal and the probabilistic hazard analyses. This sensitivity analysis can also be extended to evaluate how kernels with adaptive bandwidths perform, such as the m^{th} nearest neighbour estimate (Connor and Hill 1995; Bebbington 2013), those based on the Botev's algorithm (Botev et al. 2010; Galindo et al. 2016), or kernels with other bandwidths, such as the Kullback-Leibler score (Vere-Jones 1992; Bebbington 2015).

Another highlight compared to previous approaches is that the weight assigned to each vent is left up to the user. To test the effect of this, our weighted approach (i.e., younger vents having a greater influence on forecasts) was compared to an unweighted case. Likewise, the area enclosed by the isocontours 10^{-20} , 10^{-10} , and 10^{-5} of their respective spatial probability maps (i.e., after applying Equation (10)) was noted for comparison (Table 4). The greatest differences (up to 30%) were found at the higher probability isocontour, suggesting that weighting strongly influences spatial probability estimates. In our test case, an exponentially decreasing function was assumed following Cappello et al. (2013), however the user can choose any other function that best represents the age data and then make comparisons, e.g., if there is a distinct spatial trend in the data (e.g., Connor and Hill 1995; Cronin et al. 2001; Ho 2010; Gallant et al. 2018), the weighting could be set up to reflect this.

For the temporal assessment, in order to compare the power law approach with long-term recurrence rates, four different time intervals (1, 10, 100, and 1000 years) were considered and their temporal probabilities (after applying equation (23)) calculated for each method (Table 5). These results suggest differences of up to 43% for the shortest forecasting time interval, which are

Table 3 Sensitivity analysis. Comparison of the area enclosed by three probability isocontours of the combined PDF (shown in Fig. 4h) for three different bandwidth estimators

Probability isocontour	Area (km ²)		
	$\hat{H}_{PI,AMSE}$	$\hat{H}_{PI,SAMSE}$	\hat{H}_{SCV}
10^{-20}	1866	1875	2587
10^{-10}	976	991	1184
10^{-5}	303	311	318

Table 4 Sensitivity analysis. Comparison of the area enclosed by three probability isocontours of the spatial density for volcanic vents (shown in Fig. 4a) for a weighted ($w_k = f(t)$) and a non-weighted ($w_k = 1$) case

Probability isocontour	Area (km ²)	
	$w_k = f(t)$	$w_k = 1$
10^{-20}	682	688
10^{-10}	373	381
10^{-5}	110	143

smaller for larger forecasting horizons. In consequence, special care should be taken when conducting the temporal probability analysis. Ideally, several approaches and age datasets, including uncertainties, should be tested and compared to each other.

For the hazard modeling assessment, a Latin-hypercube sampling approach was used to assign different modeling parameters to every source pixel for each volcanic phenomenon. In order to validate the code against outputs of the same models for the same inputs (after applying equation (27)), up to 10 simulations were run and their results compared. Interestingly, very small differences (between 0.3 and 1%) were found when comparing the areas enclosed by the probability isocontours 10^{-8} , 10^{-6} , and 10^{-4} , suggesting that very few simulations are representative for each hazard. This finding may be due to the sampling technique, which aims at spreading the modeling parameters more evenly than a pure random (e.g., Monte Carlo) approach.

The pixel size influence on the simulation times was also tested. In our test case, a pixel dimension of 100 m was used, which is $\sim 0.1\%$ of the longest side of the expanded study area, providing $\sim 5 \cdot 10^5$ pixels. A pixel size of 1000 m ($\sim 5 \cdot 10^3$ pixels) reduced the simulation times by 2.5 orders of magnitude, but produced coarse probability isocontours. A pixel size of 10 m ($\sim 5 \cdot 10^7$ pixels), on the other hand, was far more computationally intensive, but with little reward, in that the shapes of each probability isocontour were quite similar to the 100 m pixel case. The pixel size/computational efficiency trade-

Table 5 Sensitivity analysis. Comparison of the spatio-temporal cumulative probabilities for the whole study area (shown in Fig. 5) for four different time periods, considering both power law and long-term average recurrence rate approaches

Time period	Spatio-temporal cumulative probabilities	
	Power law process	Long-term average recurrence rate
1	0.2421	0.1381
10	0.2929	0.1695
100	0.6221	0.4055
1000	0.9994	0.9790

off would need to be evaluated for each volcanic field studied, depending on its overall dimensions.

Finally, another innovative step is the aggregate probability for each volcanic phenomenon. In our CLV example, phenomena were weighted accordingly to their frequency of occurrence in the eruptive record. However, for volcanic fields with an incomplete eruptive record, the user could alternatively weight phenomena based on a simplified view of their relative hazard (PDCs might be given a higher weighting than ashfall, for example). If this approach is taken, the restriction shown in Equation (29) would have to be discarded, and the sum of weights would be any integer, and interpreted as a hazard index.

Final considerations

MatHaz has been developed as a tool to help deal with the real and practical complications of assessing hazard for distributed volcanism, so its intended users are researchers who need to make hazard maps for volcanic fields. This tool can also be used to provide a general approach for hazards in large volcanic areas and/or to rapidly generate several scenarios for volcanic fields showing signs of unrest.

Its systematic, sequential, and automated data-driven methodology can be tuned to perform optimally for any volcanic area, regardless of its extent, type, and amount of volcano-structural data. However, in order to achieve this task, and unlike other hazard simulation tools, MatHaz has not been developed to model the physics of volcanic eruptions and their products, or to model the details of specific eruptive scenarios.

MatHaz was created in Matlab R2017b (version 9.3) and has been successfully tested in Matlab R2015a (version 8.5). It does not require any additional toolboxes, however the Statistics and machine learning toolbox is required for 'Step 3' of the program if a Latin-hypercube sampling approach is envisioned.

MatHaz can be found in Additional file 4. The code is open source, so it can be edited and replicated. This code and further updates are planned to be hosted on the Github hosting service (<https://github.com/geoscience-community-codes>).

Disclaimer

It should be emphasized that a volcanic hazard map of the Carrán–Los Venados Volcanic Field already exists, produced by Bertin et al. (2018) and issued by SERNA-GEOMIN, which is the only agency in Chile that has the mandate to develop official hazard maps. The official map developed by Bertin et al. (2018) is therefore the map that should be used for any volcanic planning and mitigation activities. The probabilistic spatio-temporal volcanic hazard assessment conducted here, as well as

every hazard map produced as a result, although based on sound data merely constitute an intellectual exercise performed to illustrate the implementation of the MatHaz tool. Therefore, the qualitative volcanic hazard map obtained (Fig. 8) is not comparable in any regular way with the official volcanic hazard map issued by SERNAGEOMIN.

Appendix 1

Kernel density estimation – Theoretical framework

Kernel density estimation is a non-parametric method for estimating the PDF of a distributed sample in which the parameters that rule it are unspecified. In the bivariate case, that is:

$$\hat{f}(r) = \frac{1}{S} \sum_{k=1}^S K(r-r_k) \tag{30}$$

Where $K(\cdot)$ is the kernel, $r = (x, y)^T$, and $r_k = (x_k, y_k)^T$, being r the coordinates of every point in the two-dimensional space and r_k the coordinates of the k -th component of a bivariate sample of size S .

This kernel has to satisfy the following properties:

$$\int_{\mathbb{R}^2} K(u) du = 1 \tag{31}$$

which ensures that a PDF will be obtained, and

$$K(u) = K(-u) \quad \forall u \tag{32}$$

which ensures that the mean of the distribution will be equal to that of the sample.

If $K(u)$ is a kernel, that means that $\lambda K(\lambda u)$ for some $\lambda > 0$ is a kernel as well. This property allows an appropriate scale to be chosen. To do this, it has commonly been assumed that $\lambda = \frac{1}{h}$, where h is a smoothing parameter, also called bandwidth, which determines how smooth the estimated function is (i.e., it controls the size of the neighborhood around every r_k). In the one-dimensional case, the bandwidth is a scalar (h), but in the two-dimensional case it can be either a scalar (h) or a matrix (H). If the bandwidth is a matrix, the scaled bivariate kernel is:

$$K_H(r-r_k) = |H|^{-\frac{1}{2}} K\left(H^{-\frac{1}{2}}(r-r_k)\right) \tag{33}$$

Although there are various choices among kernels, the most often used in volcanology is the Gaussian kernel (e.g., Connor and Hill 1995; Conway et al. 1998; Connor et al. 2000), mainly due to its small loss of efficiency and its mathematical properties, although its support extends to infinity. If a Gaussian kernel is considered, equation (30) becomes:

$$\hat{f}_H(r) = \frac{|H|^{-\frac{1}{2}}}{2\pi S} \sum_{k=1}^S e^{-\frac{1}{2}b^T b} \tag{34}$$

where $b = H^{-\frac{1}{2}}(r-r_k)$.

The performance of a kernel density estimator depends more strongly on the bandwidth selection than on the kernel function chosen (Duong 2005). In the Gaussian case, a small bandwidth leads to very spiky estimates, while a large one an overly smoothed estimate (Connor et al. 2009). If the bandwidth is a matrix, its notation is:

$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{12} & h_{22} \end{bmatrix} \tag{35}$$

This matrix has to be positive-definite, which amongst many other properties means that is invertible, has a square root, and $h_{11} > 0$, $h_{22} > 0$, $|h_{12}| < h_{11}h_{22}$. Notation in equation (35) implies two smoothing patterns of the kernel function, not necessarily aligned along the coordinate axes of the grid. That is, if a scalar bandwidth (h) is used, its shape is circular, whereas if it is a matrix (H), its shape is elliptical and smoothed by the major and minor axes lengths. To obtain those parameters, both bandwidth's eigenvalues and eigenvectors have to be calculated.

There are many methods to estimate an optimal bandwidth matrix (Duong 2018, and references therein). No one method has emerged as the best, mainly due to limitations on measuring their relative rate of convergence to the AMISE (asymptotic mean integrated squared error) optimal bandwidth matrix (Duong 2005). However, selectors based on either the AMSE (asymptotic mean squared error) or the SAMSE (sum of AMSE) procedures are recommended because of their simplicity (i.e., statistical parsimony) and because they produce finite pilot bandwidths, displaying faster rates of convergence (Duong 2005). To obtain an optimal bandwidth, the freely available statistical program R can be used (<https://www.r-project.org/>).

Appendix 2

Derivation of Equation (26)

According to Pyle (1989), tephra thickness can be empirically modeled as an exponential function of either the square root of the isopach area A or the distance to the source r , that is:

$$T = T_0 e^{-k\sqrt{A}} = T_0 e^{-qr} \tag{35}$$

Where $-k$ and $-q$ are the slopes on plots of $\ln(T)$ versus \sqrt{A} and r , respectively, while T_0 is the extrapolated maximum deposit thickness. Assuming that all the isopachs are elliptical in shape, both r and A can be expressed as follows:

$$r(\phi) = \frac{r_0(1-\varepsilon)}{1-\varepsilon \cos\phi} \tag{36}$$

$$A = (1-\varepsilon)^2 \int_0^\pi \frac{r_0^2}{(1-\varepsilon \cos\phi)^2} d\phi = \frac{\pi r_0^2(1-\varepsilon)^2}{(1-\varepsilon^2)^{3/2}} \tag{37}$$

Where r_0 and ε are the major axis and the eccentricity of the ellipse, with the source at one of the foci of the ellipse, whereas ϕ is the angular coordinate measured from the horizontal axis of the study area.

Equating exponents in equation (35), then:

$$k \sqrt{\frac{\pi r_0^2(1-\varepsilon)^2}{(1-\varepsilon^2)^{3/2}}} = q \frac{r_0(1-\varepsilon)}{1-\varepsilon \cos\phi} \tag{38}$$

Simplifying and solving for q :

$$q = k(1-\varepsilon \cos\phi) \sqrt{\frac{\pi}{(1-\varepsilon^2)^{3/2}}} \tag{39}$$

ϕ and r can be written as:

$$\phi = \phi_{ij} = \tan^{-1} \left(\frac{y_j - y_0}{x_i - x_0} \right) \tag{40}$$

$$r = r_{ij} = \sqrt{(x_i - x_0)^2 + (y_j - y_0)^2} \tag{41}$$

Where x_i and y_j are the respective i -th western and j -th northern coordinates of the study area, and x_0 and y_0 are the western and northern coordinates of the source.

If the ellipse is rotated (i.e., its major axis does not coincide with the horizontal axis of the study area), an angular term ϕ_0 can be added, so equation (40) becomes:

$$\phi_{ij} = \phi_0 + \tan^{-1} \left(\frac{y_j - y_0}{x_i - x_0} \right) \tag{42}$$

Equation (26) can be obtained by inserting equation (42) in equation (39), and by replacing equations (39) and (41) in equation (35).

Notation

A area enclosed within a given isopach.

C_D drag coefficient.

d major radius of the particle.

d_l length of a line.

ds squared pixel size of the

topographic matrix.

dx E-W unit projection of a line.

dy N-S unit projection of a line.

e intermediate radius of the particle.

f minor radius of the particle.

$\hat{f}(\cdot), \hat{f}_H(\cdot)$ PDFs of a distributed sample.

$\hat{f}_H^q(\cdot)$ PDF of the q -th bivariate sample.

h scalar bandwidth.

h' total number of hazards modeled.

h_b unit thickness of every block.

h_c collapse height.

h_{ij} components of a bandwidth matrix H .

h_v elevation of the vent

H matrix bandwidth

\hat{H}_{NS} normal-scale selector

\hat{H}_{NM} normal-mixture selector

\hat{H}_{PI} plug-in selector

$\hat{H}_{PI,AMSE}$ plug-in selector with AMSE pilot

$\hat{H}_{PI,SAMSE}$ plug-in selector with SAMSE pilot

\hat{H}_{LSCV} least-squares cross-validation selector

\hat{H}_{BCV} biased cross-validation selector

\hat{H}_{SCV} smoothed cross-validation selector

$-k$ slope on a $\ln(T)$ versus \sqrt{A} plot

$K(\cdot)$ (kernel)

$K_H(\cdot)$ scaled bivariate kernel

L run-out length of the phenomena

m number of pixels in the E-W direction of the expanded study area

m_v slope of the energy cone

n number of pixels in the N-S direction of the expanded study area

n' number of pixels considered in the analysis

n_e number of eruptions to be forecasted

N total number of eruptions contained between t_1 and t_N

p integrated spatio-temporal probability matrix

p_h integrated spatio-temporal probability matrix for the h -th hazard

p_p^h spatio-temporal probability matrix related to the p -th pixel for the h -th hazard

q number of bivariate samples

$-q$ slope on a $\ln(T)$ versus r plot

$r = r_{ij}$ distance from the source (x_0, y_0) to any (x_i, y_j)

$r = (x, y)^T$ two-dimensional Cartesian coordinate system

r_0 ellipse's major axis with the source at one of its foci

$r_k = (x_k, y_k)^T$ coordinates of the k -th component of a bivariate sample

S size of a bivariate sample

S_q size of the q -th bivariate sample

S_t total volcano-structural datasets

Δt forecasting time interval

t_1 age of the oldest known eruption

t_e year of interest counted from the age of the oldest known eruption

t_i age of the i -th eruption

t_k age of the k -th volcanic vent

Δt_{max} time when the temporal probability attains its maximum

t_N age of the youngest known eruption

$T = T_{ij}$ tephra thickness at (x_i, y_j)

T_0 extrapolated maximum tephra thickness
 T_r retrospective time frame
 V tephra fall volume
 V_0 launch velocity of the particle
 V_T total volume of the phenomena
 x_i western coordinate of the i -th E-W segment of the study area
 y_j northern coordinate of the j -th N-S segment of the study area
 w_h weight assigned to the h -th volcanic hazard
 w_k weight assigned to the k -th volcanic vent
 w_q weight assigned to the q -th volcano-structural dataset
 (x_0, y_0) coordinates of the tephra source
 (x_1, y_1) starting coordinates of a line
 (x_m, y_m) starting coordinates of the m -th segment of a line
 (x_n, y_n) ending coordinates of a line
 δ positive parameter that controls the power law process
 ε isopach eccentricity
 θ positive parameter that controls the power law process
 θ_0 ejection angle of the particle with the horizontal plane
 θ_i azimuth of a line
 $\lambda(\cdot)$ space-time varying intensity function
 $\rho(\cdot)$ point process intensity function
 ρ_s block density
 $\phi(\cdot)$ recurrence rate
 ϕ_0 angle between the ellipse's major axis and the E-W axis of the study area
 $\chi(\cdot)$ actual cumulative number of eruptions
 Δ steps parameter
 $\Lambda(\cdot)$ estimated number of eruptions
 $\phi = \phi_{ij}$ angle between (x_0, y_0) and (x_i, y_j) measured from the ellipse's major axis
 $\Phi(\cdot)$ virtual cumulative number of eruptions

Additional files

Additional file 1: Excel file containing sheets for different volcano-structural datasets for the CLV. (XLSX 36 kb)

Additional file 2: Excel template for obtaining the parameters δ and θ of a Weibull function. (XLSX 23 kb)

Additional file 3: Excel file containing the topography of the CLV area written in ASCII format. (XLSX 1035 kb)

Additional file 4: MatHaz code. (ZIP 11 kb)

Abbreviations

AMISE: Asymptotic mean integrated squared error; AMSE: Asymptotic mean squared error; ASCII: American Standard Code for Information Interchange; CLV: Carrán-Los Venados Volcanic Field; ENE: East-northeast; E-W: East-west; GIS: Geographic Information System; LOFZ: Liquiñe-Ofqui Fault Zone; M_L : Local Magnitude; NE-SW: Northeast-southwest; NNE: North-northeast; N-S: North-south; NW-SE: Northwest-southeast; PDC: Pyroclastic density current; PDF: Probability density function; QGIS: Quantum GIS; SAMSE: Sum of AMSE; SERNAGEOMIN: Chilean Geological and Mining Survey; SVZ: Southern Volcanic Zone of the Andes; UTM: Universal Transverse Mercator

Acknowledgments

DB acknowledges CONICYT-Becas Chile for a PhD scholarship and the University of Auckland for funding support. JML acknowledges support from the New Zealand Earthquake Commission. LB acknowledges Becas Iberoamérica-Santander Investigación 2016 for giving her the opportunity to conduct an internship in SERNAGEOMIN-Chile. SJC acknowledges the support of the Natural Hazards Research Platform project "Quantifying volcanic hazards from multiple, cascading events". Further data for this paper, if needed, are available by contacting the corresponding author at daniel.bertin.u@gmail.com. Associate Editor Charles B. Connor and an anonymous reviewer are deeply thanked for their thorough comments and suggestions that substantially improved the quality and structure of this paper.

Author's contributions

DB conceptualized and wrote the program. DB wrote the manuscript with several inputs from JML, LB and SJC. LJB provided all the geological data, published and unpublished. All authors read and approved the final manuscript.

Funding

This work was partially funded through a CONICYT-Becas Chile PhD scholarship held by DB (2018–2021: Folio 72170365) and supported by the University of Auckland, New Zealand. JML is partially supported by the New Zealand Earthquake Commission. SJC is partially supported by the Natural Hazards Research Platform.

Availability of data and materials

The datasets analysed during the current study are available in Additional files 1, 2 and 3. MatHaz is available in Additional file 4.

Competing interests

The authors declare that they have no competing interests.

Author details

¹School of Environment, University of Auckland, Auckland, New Zealand.

²Servicio Nacional de Geología y Minería, Red Nacional de Vigilancia Volcánica, Santiago, Chile. ³Servicio Nacional de Geología y Minería, Dirección Regional de Atacama, Copiapó, Chile.

Received: 11 April 2019 Accepted: 20 June 2019

Published online: 18 July 2019

References

- Alcorn R, Panter KS, Gorsevski PC. A GIS-based volcanic hazard and risk assessment of eruptions sourced within Valles caldera, New Mexico. *J Volcanol Geotherm Res.* 2013;267:1–14.
- Amigo A, Bertin D, Orozco G. Peligros volcánicos de la zona norte de Chile, Regiones de Arica y Parinacota, Tarapacá, Antofagasta y Atacama, Servicio Nacional de Geología y Minería, Carta Geológica de Chile, Serie Geología Ambiental, vol. 17; 2012. 45.
- Bartolini S, Bolós X, Martí J, Riera Pedra E, Planagumà L. Hazard assessment at the quaternary La Garrotxa volcanic field (NE Iberia). *Nat Hazards.* 2015;78(2): 1349–67.
- Bartolini S, Cappello A, Martí J, Del Negro C. QVAST: a new quantum GIS plugin for estimating volcanic susceptibility. *Nat Hazards Earth Syst Sci.* 2013;13: 3031–42.
- Bartolini S, Geyer A, Martí J, Pedrazzi D, Aguirre-Díaz G. Volcanic hazard on Deception Island (South Shetland Islands, Antarctica). *J Volcanol Geotherm Res.* 2014;285:150–68.
- Bebbington M, Cronin SJ. Spatio-temporal hazard estimation in the Auckland volcanic field, New Zealand, with a new event-order model. *Bull Volcanol.* 2011;73:55–72.
- Bebbington MS. Models for temporal volcanic hazard. *Stat Volcanol.* 2012;1:1–24.
- Bebbington MS. Assessing spatio-temporal eruption forecasts in a monogenetic volcanic field. *J Volcanol Geotherm Res.* 2013;252:14–28.
- Bebbington MS. Spatio-volumetric hazard estimation in the Auckland volcanic field. *Bull Volcanol.* 2015;77:39.
- Becerril L, Bartolini S, Sobradelo R, Martí J, Morales JM, Galindo I. Long-term volcanic hazard assessment on El Hierro (Canary Islands). *Nat Hazards Earth Syst Sci.* 2014;14:1853–70.

- Becerril L, Cappello A, Galindo I, Neri M, Del Negro C. Spatial probability distribution of future volcanic eruptions at El Hierro Island (Canary Islands, Spain). *J Volcanol Geotherm Res.* 2013;257:21–30.
- Becerril L, Martí J, Bartolini S, Geyer A. Assessing qualitative long-term volcanic hazards at Lanzarote Island (Canary Islands). *Nat Hazards Earth Syst Sci.* 2017; 17:1145–57.
- Bertin D. 3-D ballistic transport of ellipsoidal volcanic projectiles considering horizontal wind field and variable shape-dependent drag coefficients. *J Geophys Res: Solid Earth.* 2017;122:1126–51.
- Bertin D, Lara LE, Basualto D, Amigo A, Cardona C, Franco L, Gil F, Lazo J. High effusion rates of the Cordón Caulle 2011–2012 eruption (southern Andes) and their relation with the quasi-harmonic tremor. *Geophys Res Lett.* 2015; 42:7054–63.
- Bertin L, Moreno H. *Geología del Campo Volcánico Carrán - Los Venados, Región de los Ríos, Servicio Nacional de Geología y Minería, Carta Geológica de Chile, Serie Geología Básica*; 2019. in press.
- Bertin L, Moreno H, Becerril L. Peligros del Campo Volcánico Carrán - Los Venados, Región de los Ríos, Servicio Nacional de Geología y Minería, Carta Geológica de Chile, Serie Geología Ambiental, vol. 33; 2018. 52.
- Bevilacqua A, Bursik M, Patra A, Pitman EB, Till R. Bayesian construction of a long-term vent opening map in the Long Valley volcanic region (CA, USA). *Stat Volcanol.* 2017;3(1):1–36.
- Bevilacqua A, Isaia R, Neri A, Vitale S, Aspinall WP, Bisson M, Flandoli F, Baxter PJ, Bertagnini A, Esposti Ongaro T, Iannuzzi E, Pistolesi M, Rosi M. Quantifying volcanic hazard at Campi Flegrei caldera (Italy) with uncertainty assessment: 1. Vent opening maps. *J Geophys Res: Solid Earth.* 2015;120(4):2309–29.
- Biass S, Falcone J-L, Bonadonna C, Di Traglia F, Pistolesi M, Rosi M, Lestuzzi P. Great balls of fire: a probabilistic approach to quantify the hazards related to ballistic – a case study at La Fossa volcano, Vulcano Island, Italy. *J Volcanol Geotherm Res.* 2016;325:1–14.
- Bonadonna C, Connor CB, Houghton BF, Connor L, Byrne M, Laing A, Hincks TK. Probabilistic modeling of tephra dispersal: hazard assessment of a multiphase rhyolitic eruption at Tarawera, New Zealand. *J Geophys Res: Solid Earth.* 2005a;110:B03203.
- Bonadonna C, Costa A. Estimating the volume of tephra deposits: a new simple strategy. *Geology.* 2012;40:415–8.
- Bonadonna C, Phillips JC, Houghton BF. Modeling tephra sedimentation from a Ruapehu weak plume eruption. *J Geophys Res: Solid Earth.* 2005b;110:B08209.
- Botev ZI, Grotowski JF, Kroese DP. Kernel density estimation via diffusion. *Ann Stat.* 2010;38(5):2916–57.
- Brewer, C.A., Harrower, M., Sheetsley, B., Woodruff, A., Heyman, D. (2013). *ColorBrewer 2.0*, the Pennsylvania State University, state college, PA, www.colorbrewer2.org
- Bucchi F, Lara LE, Gutiérrez F. The Carrán - Los Venados volcanic field and its relationship with coeval and nearby polygenetic volcanism in an intra-arc setting. *J Volcanol Geotherm Res.* 2015;308:70–81.
- Burden RW, Chen L, Phillips CJ. A statistical method for determining the volume of volcanic fall deposits. *Bull Volcanol.* 2013;75:707.
- Campos, A., Moreno, H., Muñoz, J., Antinao, J.L., Clayton, J., Martin, M. (1998). *Área de Futrono - Lago Ranco, Región de los Lagos, Servicio Nacional de Geología y Minería, Mapas Geológicos No.8.*
- Cappello A, Bilotta G, Neri M, Del Negro C. Probabilistic modeling of future volcanic eruptions at Mt Etna. *J Geophys Res: Solid Earth.* 2013;118:1925–35.
- Cappello A, Neri M, Acocella V, Gallo G, Vicari A, Del Negro C. Spatial vent opening probability map of Mt Etna volcano (Sicily, Italy). *Bull Volcanol.* 2012; 74:2083–94.
- Cembrano J, Hervé F, Lavenu A. The Liquiñe-Ofqui fault zone: a long-lived intraarc fault zone in southern Chile. *Tectonophysics.* 1996;259:55–66.
- Cembrano J, Lara L. The link between volcanism and tectonics in the southern volcanic zone of the Chilean Andes: a review. *Tectonophysics.* 2009;471:96–113.
- Charbonnier SJ, Gertisser R. Numerical simulations of block-and-ash flows using the Titan2D flow model: examples from the 2006 eruption of Merapi volcano, Java, Indonesia. *Bull Volcanol.* 2009;71:953–9.
- Connor CB, Bebbington MS, Marzocchi W. Probabilistic volcanic hazard assessment. In: Sigurdsson H, Houghton BF, McNutt SR, Rymer H, Stix J, editors. *Encyclopedia of volcanoes* (2nd edition). Amsterdam: Elsevier; 2015. p. 897–910.
- Connor CB, Connor LJ, Germa A, Richardson JA, Bebbington M, Saballos JA. How to use kernel density estimation as a diagnostic and forecasting tool for distributed volcanic vents. *Stat Volcanol.* 2018;4:1–21.
- Connor CB, Connor LJ, Jaquet O, Wallace L, Kiyosugi K, Chapman N, Sparks RSJ, Goto J. Spatial and temporal distribution of future volcanism in the Chugoku region – a partial application of NUMO's ITM and topaz probabilistic tectonic assessment methodology, technical report NUMO-TR-13-03. Tokyo: Nuclear Waste Organization of Japan; 2013. 83.
- Connor CB, Hill BE. Three nonhomogeneous Poisson models for the probability of basaltic volcanism: application to the Yucca Mountain region. *J Geophys Res.* 1995;100:10107–25.
- Connor CB, Sparks RSJ, Díez M, Volentik ACM, Pearson SCP. The nature of volcanism. In: Connor CB, Chapman NA, Connor LJ, editors. *Volcanic and tectonic hazard assessment for nuclear facilities*. Cambridge: Cambridge University Press; 2009. p. 74–115.
- Connor CB, Stamatakis JA, Ferrill DA, Hill BE, Ofoegbu GI, Conway FM, Sagar B, Trapp J. Geologic factors controlling patterns of small volume basaltic volcanism: application to a volcanic hazards assessment at Yucca Mountain, Nevada. *J Geophys Res.* 2000;105(1):417–32.
- Connor LJ, Connor CB. Inversion is the key to dispersion: understanding eruption dynamics by inverting tephra fallout. In: Coles SG, Connor CB, Connor LJ, Mader HM, editors. *London: Statistics in Volcanology, Geological Society of London*; 2006. p. 231–42.
- Connor LJ, Connor CB, Meliksetian K, Savov I. Probabilistic approach to modeling lava flow inundation: a lava flow hazard assessment for a nuclear facility in Armenia. *J Applied Volcanol.* 2012;1:3.
- Conway FM, Connor CB, Hill BE, Condit CD, Mullaney K, Hall CM. Recurrence rates of basaltic volcanism in SP cluster, San Francisco volcanic field, Arizona. *Geology.* 1998;26:655–8.
- Corazzato C, Tibaldi A. Fracture control on type, morphology and distribution of parasitic volcanic cones: an example from Mt, Etna, Italy. *J Volcanol Geotherm Res.* 2006;158:177–94.
- Cronin SJ, Bebbington M, Lai CD. A probabilistic assessment of eruption recurrence on Taveuni volcano, Fiji. *Bull Volcanol.* 2001;63:274–88.
- Damaschke M, Cronin SJ, Bebbington MS. A volcanic event forecasting model for multiple tephra records, demonstrated on Mt, Taranaki, New Zealand. *Bull Volcanol.* 2018;80:9.
- Deng F, Connor CB, Malservisi R, Connor LJ, White JT, Germa A, Wetmore PH. A geophysical model for the origin of volcano vent clusters in a Colorado plateau volcanic field. *J Geophys Res: Solid Earth.* 2017;122:8910–24.
- Duong T. Bandwidth selectors for multivariate kernel density estimation, Ph.D. thesis, University of Western Australia, School of Mathematics and Statistics; 2005. p. 173.
- Duong T. (2018), 'R package 'ks' – Reference manual', <https://cran.r-project.org/web/packages/ks/ks.pdf>. Accessed 2018-07-26.
- El Difrawy MA, Runge MG, Moufti MR, Cronin SJ, Bebbington M. A first hazard analysis of the quaternary Harrat Al-Madinah volcanic field, Saudi Arabia. *J Volcanol Geotherm Res.* 2013;267:39–46.
- Eppelbaum L, Kutasov I, Pilchin A. *Applied Geothermics*: Springer; 2014. 751.
- Felpeo A, Martí J, Ortiz R. Automatic GIS-based system for volcanic hazard assessment. *J Volcanol Geotherm Res.* 2007;166:106–16.
- Fierstein J, Nathenson M. Another look at the calculation of fallout tephra volumes. *Bull Volcanol.* 1992;54:156–67.
- Folch A, Costa A, Macedonio G. A computational model for transport and deposition of volcanic ash. *Comput Geosci.* 2009;35(6):1334–42.
- Freundt A, Wilson CJN, Carey SN. Ignimbrites and block-and-ash flow deposits. In: Sigurdsson H, Houghton BF, McNutt SR, Rymer H, Stix J, editors. *San Diego: Encyclopedia of Volcanoes, Academic Press*; 2000. p. 581–600.
- Galindo I, Romero MC, Sánchez N, Morales JM. Quantitative volcanic susceptibility analysis of Lanzarote and Chinijo Islands based on kernel density estimation via a linear diffusion process. *Sci Rep.* 2016;6:28381.
- Gallant E, Richardson J, Connor C, Wetmore P, Connor L. A new approach to probabilistic lava flow hazard assessments, applied to the Idaho National Laboratory, eastern Snake River plain, Idaho, USA. *Geology.* 2018;46(10):895–8.
- George OA, McClrath J, Farrell A, Gallant E, Tavarez S, Marshall A, McNiff C, Njoroge M, Wilson J, Connor CB, Connor LJ, Kruse S. High-resolution ground-based magnetic survey of a buried volcano: anomaly B, Amargosa Desert, NV. *Stat Volcanol.* 2015;1(3):1–23.
- Germa A, Connor LJ, Cañon-Tapia E, Le Corvec N. Tectonic and magmatic control on the location of post-subduction monogenetic volcanoes in Baja California, México, revealed through spatial analysis of eruptive vents. *Bull Volcanol.* 2013;75:782.
- Germanovich LN, Lowell RP. The mechanism of phreatic eruptions. *J Geophys Res.* 1995;100(39):8417–34.

- Gesch D, Oimoen M, Zhang Z, Danielson J, Meyer D. Validation of the ASTER global digital elevation model (GDEM) version 2 over the conterminous United States, report to the ASTER GDEM version 2 validation team; 2011.
- Glodny J, Gräfe K, Echter H, Rosenau M. Mesozoic to quaternary continental margin dynamics in south-Central Chile (36°–42°S): the apatite and zircon fission track perspective. *Int J Earth Sci.* 2008;97:1271–91.
- Goff F, Janik CJ. Geothermal systems. In: Sigurdsson H, Houghton B, McNutt SR, Rymer H, Stix J, editors. *Encyclopedia of Volcanoes*. San Diego: Academic Press; 2000. p. 817–34.
- González-Mellado AO, De La Cruz-Reyna S. A simple semi-empirical approach to model thickness of ash-deposits for different eruption scenarios. *Nat Hazards Earth Syst Sci.* 2010;10:2241–57.
- Gudmundsson MT. Hazards from lahars and jökulhlaups. In: Sigurdsson H, Houghton BF, McNutt SR, Rymer H, Stix J, editors. *Encyclopedia of volcanoes* (2nd edition). Amsterdam: Elsevier; 2015. p. 971–84.
- Hayashi JM, Self S. A comparison of pyroclastic flow and debris avalanche mobility. *J Geophys Res.* 1992;97:9063–71.
- Ho CH. Hazard area and recurrence rate time series for determining the probability of volcanic disruption of the proposed high-level radioactive waste repository at Yucca Mountain, Nevada, USA. *Bull Volcanol.* 2010;72:205–19.
- Ho CH, Smith EI, Feuerbach DL, Naumann TR. Eruptive calculation for the Yucca Mountain site, USA: statistical estimation of recurrence rates. *Bull Volcanol.* 1991;54:50–6.
- Hurst AW, Turner R. Performance of the program ASHFALL for forecasting ashfall during the 1995 and 1996 eruptions of Ruapehu volcano. *N Z J Geol Geophys.* 1999;42(4):615–22.
- Jiménez D, Becerril L, Bartolini S, Martí J. Spatio-temporal hazard estimation in San Miguel volcano, El Salvador. *J Volcanol Geotherm Res.* 2018;358:171–83.
- Jones R, Manville V, Peakall J, Froude MJ, Odbert HM. Real-time prediction of rain-triggered lahars: incorporating seasonality and catchment recovery. *Nat Hazards Earth Syst Sci.* 2017;17:2301–12.
- Kawabata E, Bebbington MS, Cronin SJ, Wang T. Modeling thickness variability in tephra deposition. *Bull Volcanol.* 2013;75:738.
- Kereszturi G, Németh K, Cronin SJ, Agustín-Flores J, Smith IEM, Lindsay J. A model for calculating eruptive volumes for monogenetic volcanoes – implication for the quaternary Auckland volcanic field, New Zealand. *J Volcanol Geotherm Res.* 2013;266:16–33.
- Kiyosugi K, Connor CB, Zhao D, Connor LJ, Tanaka K. Relationships between volcano distribution, crustal structure, and P-wave tomography: an example from the Abu monogenetic volcano group, SW Japan. *Bull Volcanol.* 2010; 72(3):331–40.
- Lange D, Cembrano J, Rietbrock A, Haberland C, Dahm T, Bataille K. First seismic record for intra-arc strike-slip tectonics along the Liquiñe-Ofqui fault zone at the obliquely convergent plate margin of the southern Andes. *Tectonophysics.* 2008;455:14–24.
- Lara LE, Lavenu A, Cembrano J, Rodríguez C. Structural control on volcanism in transversal chains: resheared faults and neotectonics in the Cordón Caulle-Puyehue area (40.5°S), southern Andes. *J Volcanol Geotherm Res.* 2006;158:70–86.
- Lara LE, Moreno H. Geología del Complejo Volcánico Puyehue-Cordón Caulle, Región de los Lagos, Servicio Nacional de Geología y Minería, Carta Geológica de Chile, Serie Geología Básica; 2006. 26.
- Le Corvec N, Bernhard Spörli K, Rowland J, Lindsay J. Spatial distribution and alignments of volcanic centers: clues to the formation of monogenetic volcanic fields. *Earth Sci Rev.* 2013;124:96–114.
- Lemus M, Honores C, Aguilera F, Pérez Y, Morales D, Cáceres D, Neira H. Evaluación de los recursos geotérmicos de la Región de Los Ríos, Servicio Nacional de Geología y Minería, Informe Registrado IR-15-59; 2015. 289.
- Leonard GS, Calvert AT, Hopkins JL, Wilson CJN, Smid ER, Lindsay JM, Champion DE. High-precision ⁴⁰Ar/³⁹Ar dating of quaternary basalts from Auckland volcanic field, New Zealand, with implications for eruption rates and paleomagnetic correlations. *J Volcanol Geotherm Res.* 2017;343:60–74.
- Malin MC, Sheridan MF. Computer-assisted mapping of pyroclastic surges. *Science.* 1982;217:637–40.
- Martí J, Felpeto A. Methodology for the computation of volcanic susceptibility: an example for mafic and felsic eruptions on Tenerife (Canary Islands). *J Volcanol Geotherm Res.* 2010;195:69–77.
- Martin AJ, Umeda K, Connor CB, Weller JN, Zhao D, Takahashi M. Modeling long-term volcanic hazards through Bayesian inference: an example from the Tohoku volcanic arc, Japan. *J Geophys Res.* 2004;109:B10208.
- Mastin LG, Van Eaton AR, Lowenstern JB. Modeling ash fall distribution from a Yellowstone supereruption. *Geochem Geophys Geosyst.* 2014;15:3459–75.
- McKay MD. Latin hypercube sampling as a tool in uncertainty analysis of computer models. In proceedings of the 24th winter simulation conference, Arlington; 1992.
- Mead SR, Magill CR. Probabilistic hazard modeling of rain-triggered lahars. *J Applied Volcanol.* 2017;6:8.
- Moreno, H. (1980), La erupción del volcán Mirador en Abril-Mayo de 1979, Lago Ranco-Riñinahue, Andes del Sur, Universidad de Chile, Comunicaciones, 28, 1–23.
- Moreno PI, Denton GH, Moreno H, Lowell TV, Putnam AE, Kaplan MR. Radiocarbon chronology of the last glacial maximum and its termination in northwestern Patagonia. *Quat Sci Rev.* 2015;122:233–49.
- Nathenson M. Revised tephra volumes for Cascade Range volcanoes. *J Volcanol Geotherm Res.* 2017;341:42–52.
- Németh K. Monogenetic volcanic fields: origin, sedimentary record, and relationship with polygenetic volcanism. *Geol Soc Am Spec Pap.* 2010;470: 43–66.
- Neri A, Bevilacqua A, Esposti Ongaro T, Isaia R, Aspinall WP, Bisson M, Flandoli F, Baxter PJ, Bertagnini A, Iannuzzi E, Orsucci S, Pistolesi M, Rosi M, Vitale S. Quantifying volcanic hazard at Campi Flegrei caldera (Italy) with uncertainty assessment: 2Pyroclastic density current invasion maps. *J Geophys Res: Solid Earth.* 2015;120(4):2330–49.
- Newhall CG, Hoblitt RP. Constructing event trees for volcanic crises. *Bull Volcanol.* 2002;64:3–20.
- Nieto-Torres A, Martin Del Pozzo AL. Spatio-temporal hazard assessment of a monogenetic volcanic field, near México City. *J Volcanol Geotherm Res.* 2019;371(1):46–58.
- Patra AK, Bauer AC, Nichita CC, Pitman EB, Sheridan MF, Bursik MI, Rupp B, Webber A, Stinton AH, Namikawa LM, Renschler CS. Parallel adaptive simulation of dry avalanches over natural terrain. *J Volcanol Geotherm Res.* 2005;139:1–22.
- Pyle DM. The thickness, volume and grain size of tephra fall deposits. *Bull Volcanol.* 1989;51(1):1–15.
- Richardson JA, Connor LJ, Connor CB, Gallant E. Probabilistically modeling lava flows with MOLASSES, American Geophysical Union, fall meeting 2017, abstract #V41B-02; 2017b.
- Richardson JA, Wilson JA, Connor CB, Bleacher JE, Kiyosugi K. Recurrence rate and magma effusion rate for the latest volcanism on Arsia Mons, Mars. *Earth Planetary Sci Lett.* 2017a;458:170–8.
- Rodríguez C. Geología del Grupo Carrán – Los Venados, Andes del Sur (40.3°S), Memoria de título (Inédito), Universidad de Chile, Departamento de Geología; 1999. p. 133.
- Rosenau M, Melnick D, Echter H. Kinematic constraints on intra-arc shear and strain partitioning in the southern Andes between 38°S and 42°S latitude. *Tectonics.* 2006;25:TC4013.
- Runge MG, Bebbington MS, Cronin SJ, Lindsay JM, Kenedi CL, Moufti MR. Vents to events: determining an eruption event record from volcanic vent structures for the Harrat Rahat, Saudi Arabia. *Bull Volcanol.* 2014;76:804.
- Sandri L, Jolly G, Lindsay J, Howe T, Marzocchi W. 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.* 2012;74:705–23.
- Sandri L, Thouret J-C, Constantinescu R, Biass S, Tonini R. Long-term multi-hazard assessment for El Misti volcano (Peru). *Bull Volcanol.* 2014;76:771.
- Saucedo R, Macías JL, Sheridan MF, Bursik MI, Komorowski JC. Modeling of pyroclastic flows of Colima Volcano, México: implications for hazard assessment. *J Volcanol Geotherm Res.* 2005;139:103–15.
- Schuster RL, Crandell DR. Catastrophic debris avalanches from volcanoes. In *The IV International Symposium on Landslides*, vol. 1984;1:567–72.
- Schwaiger HF, Denlinger RP, Mastin LG. Ash3d: a finite-volume, conservative numerical model for ash transport and tephra deposition. *J Geophys Res: Solid Earth.* 2012;117:B04204.
- Sheridan MF, Malin MC. Application of computer-assisted mapping to volcanic hazard evaluation of surge eruptions: Vulcano, Lipari. *J Volcanol Geotherm Res.* 1983;17:187–202.
- Siebert L. Large volcanic debris avalanches: characteristics of source areas, deposits, and associated eruptions. *J Volcanol Geotherm Res.* 1984;22:163–97.
- Siebert L. Hazards of large volcanic debris avalanches and associated eruptive phenomena. In: Scarpa R, Tilling RI, editors. *Monitoring and Mitigation of Volcano Hazards*. Berlin: Springer-Verlag; 1996. p. 541–72.
- Smethurst L, James MR, Pinkerton H, Tawn JA. A statistical analysis of eruptive activity on Mount Etna, Sicily. *Geophys J Int.* 2009;179:655–66.

- Tadini A, Bonali FL, Corazzato C, Cortés JA, Tibaldi A, Valentine GA. Spatial distribution and structural analysis of vents in the lunar crater volcanic field (Nevada, USA). *Bull Volcanol.* 2014;76(11):877.
- Ui T. Volcanic dry avalanche deposits – identification and comparison with nonvolcanic debris stream deposits. *J Volcanol Geotherm Res.* 1983;18:135–50.
- Ui T, Takarada S, Yoshimoto M. Debris avalanches. In: Sigurdsson H, Houghton B, McNutt SR, Rymer H, Stix J, editors. *Encyclopedia of Volcanoes*: Academic Press, San Diego; 2000. p. 617–26.
- Valentine GA, Connor CB. Basaltic volcanic fields. In: Sigurdsson H, Houghton BF, McNutt SR, Rymer H, Stix J, editors. *Encyclopedia of volcanoes (2nd edition)*: Amsterdam: Elsevier; 2015. p. 423–39.
- Valentine GA, Graettinger AH, Sonder I. Explosion depths for phreatomagmatic eruptions. *Geophys Res Lett.* 2014;41:3045–51.
- Valentine GA, Gregg TKP. Continental basaltic volcanoes – processes and problems. *J Volcanol Geotherm Res.* 2008;177:857–73.
- Vallance J, Scott K. The Osceola mudflow from Mount Rainier: sedimentology and hazards implications of a huge clay-rich lahar. *Geol Soc Am Bull.* 1997;109: 143–63.
- Vere-Jones D. Statistical methods for the description and display of earthquake catalogues. In: Walden AT, Guttorp P, editors. *Statistics in the environmental and earth sciences*. London: Edward Arnold; 1992. p. 220–46.
- Wetmore PH, Hughes SS, Connor LJ, Caplinger M. Spatial distribution of eruptive centers about the Idaho National Laboratory. In: Connor CB, Chapman NA, Connor LJ, editors. *Volcanic and tectonic hazard assessment for nuclear facilities*. Cambridge: Cambridge University Press; 2009. p. 385–405.
- White JT, Connor CB, Connor L, Hasenaka T. Efficient inversion and uncertainty quantification of a tephra fallout model. *J Geophys Res: Solid Earth.* 2017; 122(1):281–94.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

