How to use kernel density estimation as a diagnostic and forecasting tool for distributed volcanic vents

C. B. CONNOR¹
L. J. CONNOR¹
A. GERMA¹
J. A. RICHARDSON²
M. BEBBINGTON³
E. GALLANT¹
J. A. SABALLOS⁴

(1) *University of South Florida, School of Geosciences, USA*
(2) *NASA Goddard Space Flight Center, USA*
(3) *Massey University, NZ*
(4) *Instituto Nicaragüense de Estudios Territoriales, Nicaragua*

KEYWORDS:

Vent opening probability maps, spatial density, kernel density, volcanic vent, monogenetic volcano, probabilistic volcanic hazard assessment, Lassen volcanic center, Miraflores-Nejapa vent alignment, Arsia Mons, Mars

CORRESPONDENCE:

Chuck Connor: cbconnor@usf.edu

University of South Florida, School of Geosciences, NES 107, 4202 E. Fowler Ave, Tampa, FL 33620-5550 (USA)

CITATION:

Connor, C. B., L. J. Connor, A. Germa, J. A. Richardson, M. Bebbington, E. Gallant, and J. A. Saballos (2019) How to use kernel density estimation as a diagnostic and forecasting tool for distributed volcanic vents, *Statistics in Volcanology* 4.3 : 1 – 25.

DOI: <http://dx.doi.org/10.5038/2163-338X.4.3>

Abstract

Volcanic activity often results in the formation of new volcanic vents. These new vents can create hazards in unexpected areas. Therefore, the probability of new vent formation should be assessed as part of volcanic hazard assessments. This paper describes our use of kernel density estimation (KDE) as a way to estimate the spatial density of future volcanic vents. The bivariate Gaussian kernel function is described step-by-step using pseudocode. Our computer code, written in PERL, is used to calculate the spatial density of existing vents and then create a contour map using GMT (Generic Mapping Tools). Application of this method and code relies on several assumptions about the definition of volcanic events, independence of events, the type of kernel function used, and the selection of kernel bandwidth. Three examples using the code are provided: (1) for volcanic vents located west of the city of Managua (Nicaragua), (2) for volcanic vents distributed within the Arsia Mons caldera (Mars), weighted by volume, and (3) for vents of the Lassen volcanic system (northern California), sub-divided by geochemistry.

Introduction

The opening of a new volcanic vent is a geologically infrequent phenomenon, but on occurrence can produce significant geological hazards. These volcanic hazards might include the formation of lava flows, explosive pyroclastic fallouts and flows, ballistic projectiles, airborne ash and ground accumulations of tephra.

Most volcanic systems are distributed, meaning that new eruptions often occur at unique locations which have not experienced volcanic activity in the past. This distributed nature applies to complex volcanoes (*e.g.*, Mt Etna) as well as to monogenetic volcanic fields, such as the San Francisco Volcanic Field located north of Flagstaff, Arizona ([Polacci & Papale, 1997](#); [Riggs et al., 2019](#)). Geologic mapping shows that over time a volcanic system can form tens to hundreds of vents ([Hasenaka & Carmichael, 1985](#); [Valentine & Connor, 2015](#)). This diffusion of vents suggests that the hazards associated with new vent formation need to be assessed as part of probabilistic volcanic hazard assessments ([Connor et al., 2015](#)).

Kernel density estimation (KDE), a nonparametric statistical method for estimating spatial density, is widely used by researchers as a method for predicting probable locations of future eruptive vents based solely on the observed locations of vents formed in the past (*e.g.*, [Connor & Connor, 2009](#); [Richardson et al., 2012](#); [Germa et al., 2013](#); [El Difrawy et al., 2013](#); [Bevilacqua et al., 2015](#); [Bartolini et al., 2015](#); [Galindo et al., 2016](#); [Mazzarini et al., 2016](#); [Tadini et al., 2017](#); [van den Hove et al., 2017](#)). In application of KDE, it is assumed that the past pattern of vent formation is governed by the same tectonic and magmatic factors that govern future vent formation. The statistical basis for KDE is well described in the statistics literature (*e.g.*, [Silverman, 1981](#); [Wand & Jones, 1994](#); [Jones et al., 1996](#); [Duong et al., 2007](#)). Problems often arise during volcanic hazard assessments when assessment teams do not understand how to execute a KDE. The scientific literature does a poor job of making the KDE methods used by researchers accessible to non-researchers. Our goal is to assist the geoscience community (experts and non-experts) by providing a step-by-step procedure for using KDE for the opening of new volcanic vents. We include a description of the equations used in KDE, pseudocode for implementing these equations, and a computer model, implemented in PERL, that calculates a spatial density grid from a list of known vent locations and contours the spatial density values using Generic Mapping Tools (GMT) ([Wessel et al., 2019](#); [contributors, 2019](#)). The output is a map that shows the resulting spatial density values as color-coded areas of high to low density. Our hope is that this step-by-step system will facilitate a better understanding of spatial density and provide a working solution that can be used for volcanic hazards assessments and also by geoscience students.

Background

Volcanism is a global phenomenon, more widely distributed in some areas than others. Before 1900, geologists Lyell and Desmarest both described the highly dispersed nature of volcanism, in Mexico and in France, respectively. The term *monogenetic volcanic field* appears to have been coined by A. Rittmann to describe these globally distributed groups of

volcanic vents ([Rittmann, 1962](#)). [Nakamura \(1977\)](#) may have been the first to use the term monogenetic volcanic field in a scientifically reviewed publication.

In the broadest sense, monogenetic volcanic fields usually lack a single central vent. An example is the Kirishima volcano complex (Japan). This complex lacks a central vent ([Nagaoka & Okuno, 2011](#)) and eruptions occur within a group of distributed vents, often resulting in the formation of a new volcanic vent that remains active for many years. At Kirishima volcano, the most recent activity has been from the Shinmoe-dake vent, which has been active for at least several hundred years ([Nakada et al., 2013](#)). In comparison, the eruption of Parícutin volcano, the most recently active vent within the monogenetic Michoacán-Guanajuato volcanic field, Mexico, lasted only 9 years; [Williams \(1950\)](#) described the dispersed nature of volcanism while mapping hundreds of older vents near Parícutin.

In contrast, some volcanoes are dominated by a central vent or vent complex, such as Mt. Fuji, Japan and Mt. Etna, Italy, but often experience multiple flank eruptions from newly formed vents ([Takada, 1997](#); [Favalli et al., 2009](#)). The distribution of these flank vents can be assessed using KDE. Volcán de Colima, Mexico, at almost the opposite end of the volcanic activity spectrum, is dominated by eruptive activity from a single central vent ([González et al., 2002](#)). Although some flank vents have formed in the past, almost all hazards at Volcán de Colima are associated with its central vent ([De la Cruz-Reyna et al., 2019](#)).

KDE was first applied to the distribution of volcanic vents by [Lutz & Gutmann \(1995\)](#). [Connor & Hill \(1995\)](#) used KDE for a probabilistic volcanic hazard assessment around the proposed Yucca Mountain nuclear waste repository. With growing use, KDE was expanded to include spatio-temporal models of volcanism ([Bebbington & Cronin, 2011](#)) and coupled to other hazard models, particularly lava flows hazards (e.g., [Favalli et al., 2012](#); [Connor et al., 2012](#); [Cappello et al., 2012](#); [Runge et al., 2014](#)), and to include uncertainty quantification ([Jaquet et al., 2008](#); [Bevilacqua et al., 2017](#)). Spatial density analysis has been incorporated into commonly used hazard applications including QVAST ([Bartolini et al., 2013](#)), VORIS ([Felpeto et al., 2007](#); [Scainia et al., 2014](#)), and MatHaz ([Bertin et al., 2019](#)), although in some of these applications the term spatial density has been replaced by spatial susceptibility or spatial vulnerability. KDE played a major role for a Yucca Mountain probabilistic volcanic hazard analysis update (PVHA-U), to further inform a panel of experts during the formal expert elicitation process ([Coppersmith et al., 2009](#)).

Spatial density is used for probabilistic volcanic hazard assessments to determine possible locations of future events (e.g., volcanic eruptions), or to estimate the probability of an event occurring at a specific location, given that such an event occurs within the region. Spatial density estimates contain no information about the *when* of future events, only the *where*, and implicitly estimates *where* a new event might occur. KDE is also used to characterize patterns of past volcanic activity and to infer mechanisms for the creation of these patterns ([Germa et al., 2013](#); [Deng et al., 2017](#); [Germa et al., 2019](#)).

The output of a spatial density estimate is often a map that shows which areas have a higher probability of future volcanic vent formation, given the distribution of past vent formation (e.g., Figure 1). A map spatially smooths the distribution of past events to estimate the likelihood of new vent formation in a small area around points of interest across the region. The *kernel* is an equation (actually a spatial density) that describes the smoothing of the intensity away from each vent. If a spatial grid of points is overlaid across an entire region, then the spatial density can be calculated for each point and contoured. The goal of this *how to* paper is to use the KDE method to calculate spatial density and provide a detailed step-by-step approach that results in a spatial density contoured map.

A method for calculating spatial density

Example: Nejapa volcano alignment, Nicaragua

We illustrate the step-wise approach to spatial density estimation using an example of 28 distributed Quaternary vents located immediately west of the city of Managua (Nicaragua). Numerous authors have noted the striking alignment of volcanic vents of different ages in this area, which is often referred to as the Nejapa alignment, or the Nejapa-Miraflores alignment ([McBirney, 1955](#); [Walker, 1984](#)). This is a distributed volcanic field with mafic-silicic vents and a wide range of eruption styles, including phreatomagmatic and plinian eruptions ([Pardo et al., 2008](#); [Freundt et al., 2010](#); [Kutterolf](#)

Figure 1: This statistical model of spatial density for volcanic vent locations (white circles) in the Nejapa volcano alignment on the western side of the Managua graben (Nicaragua) is superimposed on a shaded-relief digital elevation model (DEM). Areas most likely to experience future volcano vent formation are colored, estimated using an elliptical kernel. There is a 95% chance that a future vent will form in the contour enclosing the blue-shaded area; the highest probability zones are red. One vent, in downtown Managua, appears to be an outlier, but shows the smoothing pattern of the kernel function. The smoothing pattern is based on the existing vent locations and is objectively determined using the SAMSE method. If one or more undiscovered or buried vents exist within Lago de Managua (within the non-colored gap in the contours), then the elliptical smoothing pattern might associate this seemingly lone vent with the two northern vents that also lie east of the larger vent group. Red and yellow lines indicate faults (the red faults slipped during the 1972 earthquake that destroyed the city of Managua).

et al., 2011), with most recent activity around 1 ka (*Pardo et al.*, 2009; *Rausch & Schmincke*, 2010), and a recurrence rate of vent formation of around one event per 800 years, based on the stratigraphic work of *Avellán et al.* (2012). The N-S volcano alignment coincides with the topographic western margin of the Managua graben and is slightly oblique to Quaternary faults and faults that slipped during the 1972 Managua earthquake (*La Femina et al.*, 2002). Overall, the region is characterized by closely linked volcanism and seismicity (*Diez et al.*, 2005). Consequently, it is important to forecast potential locations of future volcanic vents with a model that is sensitive to the volcano-tectonic setting of this vent alignment.

Step 1: Selecting the vent locations

A major task in preparing a KDE is defining the set of events that will be used to estimate the spatial density. Certainly, a major expense in hazard assessment is data gathering to support the interpretation of mapped volcanic features as events. Hazard assessments sometimes consider alternative event data sets and account for the effect of these varying data sets on spatial density estimates (*Gallant et al.*, 2018).

Spatial density estimates made by KDE use the mapped locations of previously formed volcanic vents (*e.g.*, the white circles in Figure 1). Minimally, these data comprise the Cartesian coordinates of each volcanic vent. Often these coordinates are given as Universal Transverse Mercator (UTM) coordinates, relative to a single UTM zone. UTM coordinates are convenient because KDE uses the calculated distance between each vent location and each grid location within a defined map space. UTM coordinates are already expressed in units of meters, simplifying the distance calculations. When vent locations are known only as latitude and longitude coordinates, degree location coordinates can be converted to UTM coordinates; we use PROJ (*PROJ contributors*, 2019), an open source computer code. Many existing GIS platforms use PROJ for coordinate conversion (*e.g.*, GRASS, QGIS, OpenLayers, PostGIS, MapServer, Mapnik, to name a few).

KDE uses the distribution of past events as a guide to the distribution of potential events in the future. This assumption immediately raises a fundamental question; what are the past events that should be used to develop the spatial density estimate? Volcano vent location datasets used to estimate the spatial density of potential future vents need to be consistent with several features of geological processes.

First, the spatial density of volcanic vents changes with time (*Condit & Connor*, 1996; *Valentine & Perry*, 2006; *Bebbington*, 2013; *Germa et al.*, 2013; *Tadini et al.*, 2014). For processes like volcanism, where a geologic record of past events usually persists for tens of millions of years, consideration needs to be given as to which events best represent the distribution of future volcanism. For example, the distribution of Miocene volcanoes in a given area might be much less relevant than the distribution of Pliocene and Quaternary volcanoes (*Connor et al.*, 2000). Thus, in order to develop an estimate of the spatial density, a conceptual model of volcanism in the region is needed. This conceptual model is used to justify the inclusion of some volcanic vents in the analysis and to justify the exclusion of others. The optimal kernel function may also change with time (*Bebbington*, 2013), which may inform the conceptual model.

Second, how well is the distribution of past volcanic vents actually known? Even in well-studied regions, volcanic vents can be overlooked during mapping. In some areas, bias develops because volcanic vents tend to be buried by subsequent eruptions (*Wetmore et al.*, 2009). In other areas, volcanic vents are buried by sedimentation (*George et al.*, 2015). The sensitivity of the analysis to these types of potential bias in the set of volcanic vent locations is important to consider and assess.

Third, volcanic vents, even when they are all identified and mapped, may be so few as to present an incomplete picture of the underlying geological process, and therefore may not reflect the true distribution of potential volcanic vents to be formed in the future.

Fourth, which volcanic *vents* are actually independent *events*? In monogenetic volcanic fields, alignments of volcanic cones often develop in response to single magmatic events, or episodes of magma rise through the shallow crust. Single igneous dikes ascending through the crust might form segments and rotate within the shallow crust, with each segment feeding a separate vent and each building a volcanic cone (*Reches & Fink*, 1988; *Kiyosugi et al.*, 2012). If the goal of analysis is to forecast the distribution of future magmatic *events*, each of which might produce more than one monogenetic volcano, geological data must be gathered and volcanoes formed by the same magmatic event must be somehow grouped as single events (*Runge et al.*, 2014; *Bevilacqua et al.*, 2017; *Gallant et al.*, 2018). Similarly, the spatial distribution of polygenetic

volcanoes reflects processes of magma generation and rise through the crust. However, the distribution of small vents (sometimes referred to as parasitic or adventive cones) on the flanks of these volcanoes does not necessarily reflect the distribution of polygenetic volcanoes; in this case, a spatial density estimate that includes all *vents* as *events* would not necessarily correctly model spatial density of the polygenetic volcanoes or flank vents.

The independence of events is not easy to determine. This problem of deciding if a set of vents was formed during one episode of volcanic activity is an example of the difficulty in determining independence of events. Rather than simply counting volcanoes on a geologic map, a geological assessment must be made to determine the set of vents that are related to a single eruption or relatively short period of eruptive activity and to distinguish this event from those vents formed during different eruption episodes. Determining independence, or mapping vents into events, is accomplished through detailed analyses of radiometric age determinations, stratigraphic correlations, and the gathering of related geological data. Often, detailed analyses do not resolve whether or not specific vents should be grouped as single events or treated as separate, independent events.

Consequently, the key modeling issue to resolve is *consistency*. As described in the statistical literature, a model is *consistent* if it qualitatively and quantitatively produces output data similar to its input data, when simulated. Context is everything. If the modeling objective is to forecast *during quiescence* the location of the next event, then a model that considers events independently will be accurate, whereas the formation of subsequent events will not occur according to the model, as they are *conditioned* by the location of the first event. This can be a very important distinction in practice, if there is a long-term spatio-temporal dependence. A good example of this distinction is the study performed by [Magill et al. \(2005\)](#), although the input age data therein (in fact the age-order) are now considered inaccurate ([Bebbington & Cronin, 2011](#)).

Step 2: Implementing a spatial density algorithm

After compiling a set of event locations, the map area for the KDE can be defined. For a spatial density map (i.e., a probability map that integrates to one), the map area must contain all of the event data and include a large enough area around each event so that the probability can decrease toward zero at the map edges. The spatial density at any location within the map area can be calculated using KDE. The algorithm transforms each point location into a density, as shown in Figure 2, smoothing it in space, and averaging these densities to represent the underlying geological process. The KDE procedure sums the kernel function at each volcanic event in the data set for each defined map grid location (Figure 3). If each spatial density value is normalized to the grid spacing, the resulting spatial density value is the probability of a future volcanic event for each $\Delta x \times \Delta y$ grid cell, given that such an event occurs within the region.

Alternative kernel functions (e.g., the Epanechnikov kernel or the Cauchy kernel) are possible with KDE and often tailored for different purposes. In general, the shape of the kernel is considered far less important than the smoothing bandwidth, the kernel having little impact on the final density estimate in most cases ([Martin et al., 2004](#)). Kernels should be unimodal, non-negative and integrable. A more substantive question is whether a kernel with infinite support (e.g., the Gaussian kernel), which results in small, but non-zero, probability at great distances from the nearest past event, is appropriate. In hazard assessment, kernel functions with infinite tails (e.g., the Gaussian kernel) are preferred, as the probability is positive and real everywhere, albeit very small, at locations far from past events. Alternatively, a kernel with bounded support (e.g., the Epanechnikov kernel), which is non-zero only within a finite distance of a past event, could be used.

Here we demonstrate KDE using a Gaussian kernel function. The Gaussian kernel is not a bad choice since most natural processes are random processes, and the computational machinery for using the Gaussian is well advanced. Ultimately, the processes of magma generation, magma rise and volcanic eruption are all controlled by chemical and thermal diffusion, and the Gaussian kernel is mathematically associated with diffusion processes. In this case, the spatial density is continuous, differentiable and non-zero everywhere. Although the probability of a volcanic event at a specific location may vary substantially across a region, it is not reasonable to say that it is zero anywhere, just perhaps very nearly zero. The question of where a *region* should *end* remains open, but one possibility is examined by [Bebbington \(2015\)](#).

Figure 2: Shaded-relief map of a 2D radially-symmetric Gaussian kernel calculated on a grid of point locations drawn about a single volcano. Grid point locations near the volcano have the highest values of spatial density within the map grid.

Figure 3: Example set-up for spatial density estimation. Three volcanoes (v_1, v_2, v_3) are located at distances (d_1, d_2, d_3) from grid point g with location coordinates, x_2, y_2 . The grid is arbitrarily placed with respect to the volcanoes, and has dimension X in the west-east direction with grid spacing Δx , and dimension Y in the north-south direction with grid spacing Δy .

The Gaussian kernel function, in one-dimension, can be written as

$$f(x) = a \exp \left[-\frac{1}{2} \left(\frac{x-b}{c} \right)^2 \right], \quad (1)$$

where a controls the amplitude of the kernel function, b the expected location, and c the standard deviation. If all events have equal weight, then $a = (2\pi c^2)^{-1/2}$, which is the well-known Gaussian (or *normal*) probability density function.

In our example, we extend the Gaussian kernel function to the two-dimensional bivariate case and estimate spatial density, λ_g , at a map location (g) by

$$\lambda_g = \frac{1}{2\pi h^2 N} \sum_{e=1}^N \exp \left[-\frac{1}{2} \left(\frac{d_e}{h} \right)^2 \right], \quad (2)$$

where

$$d_e = \sqrt{(x_g - x_e)^2 + (y_g - y_e)^2}.$$

The total number of volcanic events is equal to N , g is the grid location where spatial density is estimated, and x_g and y_g are the map coordinates of that grid location. The local spatial density estimate, λ_g , depends on the distance, d_e , from each volcanic event location (x_e, y_e) to that grid location (x_g, y_g) , and the smoothing bandwidth, h . Spatial density decreases with distance from each event based on the value of the smoothing bandwidth and the kernel function. As a first example, we use a Gaussian kernel that is radially symmetric (isotropic), that is, h is constant in all directions. Note, that if all coordinates and distances are given in meters, then the units of λ_g are m^{-2} . If one multiplies λ_g by an area (m^2) then the result is the probability of a volcanic event forming within that area, given an event occurs within the region. Typically one multiplies λ_g by $\Delta x \times \Delta y$ (Figure 3) to normalize the spatial density by the grid cell area.

Pseudocode is provided to solve equation 2 on a map grid. An implementation of this pseudocode using a selection of volcanic vents near Managua, Nicaragua, demonstrates the symmetric nature of an isotropic smoothing bandwidth (Figure 4a).

The pseudocode accepts as input: (1) an isotropic Gaussian bandwidth (h) of 2 km, (2) a file of 28 volcano locations (x_e, y_e) in the Nejapa region of Nicaragua, and (3) a computational grid with a grid spacing ($\Delta x, \Delta y$) of $100 \text{ m} \times 100 \text{ m}$. Grid boundaries ($west = 565000$, $east = 590000$, $north = 1365000$, $south = 1325000$) and vent locations are defined in Universal Transverse Mercator coordinates (UTM Zone 16 North). The grid spacing is user definable where a larger interval will generate a coarser grid of spatial density values with a faster code execution time, while a smaller grid interval will create a smoother map at the cost of longer code runtime; changing the grid spacing will not affect the shape of contours of spatial density across the map area, only values at individual grid locations, since the values of Δx and Δy are changing.

The pseudocode outputs a table of triplets (x, y, z) where x is the easting, y is the northing, and z is the spatial density at that grid location, normalized to the area of the grid cell. The integral of the spatial density contained within the map boundaries is

$$\Lambda = \sum_{x_g=0}^{X-1} \sum_{y_g=0}^{Y-1} \lambda_g(x_g, y_g) \times \Delta x \times \Delta y, \quad (3)$$

which should be close to 1. If the sum is much less than 1, the map boundaries could be extended, or there might be an error in the implementation of the pseudocode.

If the spatial density values are sorted from largest to smallest, they can be binned as percents (e.g., 5%, 16%, 33%, 50%, 67%, 84%, 95%, 99%) that represent a fraction of the total spatial density across the map area (Figure 4).

Pseudocode to implement an isotropic kernel

Begin: The model inputs

Specify the smoothing bandwidth: h
 Specify the location (easting, northing) of each volcanic event: (x_e, y_e)
 Specify the number of volcanic events: N
 Specify the map boundaries: *west, east, south, north*
 Specify the map grid spacing: $\Delta x, \Delta y$

Internal Variables

Initialize a variable to hold the spatial density at a grid location (x_g, y_g) : $\lambda_g = 0$
 Initialize the sum of normalized spatial densities for the entire grid: $\Lambda = 0$

Loop: Do for each grid location (x_g, y_g)

For ($y_g = \text{south}$ up to $y_g = \text{north}$) for each northing

For ($x_g = \text{west}$ up to $x_g = \text{east}$) for each easting

A. Initialize variables for summation and spatial density

$\Lambda = 0$

$\lambda_g = 0$

B. Loop: Do for each volcanic event (e)

For ($e = 0$ up to $e = N$)

1. Calculate the distance between the current grid location (x_g, y_g)

and each volcanic event (x_e, y_e) : $d = \sqrt{(x_g - x_e)^2 + (y_g - y_e)^2}$

2. Calculate the spatial density: $\lambda_e = \frac{1}{(2\pi h^2 N)} \times \exp\left(-0.5 \times \left(\frac{d}{h}\right)^2\right)$
3. Normalize the spatial density by the grid spacing: $\lambda_e = \lambda_e \times \Delta x \times \Delta y$
4. Sum the spatial density for each volcanic event: $\lambda_g = \lambda_g + \lambda_e$

Next event (e)

C. Integrate across the entire grid space: $\Lambda = \Lambda + \lambda_g$
D. Print out the coordinate and spatial density at each grid location: print x_g, y_g, λ_g

Next grid easting x_g

Next grid northing y_g

End: Integration check for correctness; grid spatial densities should sum to 1: print Λ

The *isotropic* smoothing described by equation 2 is a simplification of the more general case, where the amount of smoothing that is controlled by the bandwidth h , varies in magnitude depending on direction. Often this *anisotropic* bandwidth is useful if structural or tectonic controls (e.g., dikes, faults) might be influencing vent distribution and forming alignments of a particular orientation.

A two-dimensional elliptical kernel with a direction varying bandwidth is given by (Wand & Jones, 1993),

$$\hat{\lambda}(\mathbf{g}) = \frac{1}{2\pi N \sqrt{|\mathbf{H}|}} \sum_{e=1}^N \exp\left[-\frac{1}{2} \mathbf{b}^T \mathbf{b}\right], \quad (4)$$

where,

$$\mathbf{b} = \mathbf{d} \times \mathbf{H}^{-1/2}, \quad \mathbf{d} = [d_x \ d_y], \quad d_x = x_g - x_e, \quad d_y = y_g - y_e.$$

Figure 4: A statistical model of spatial density or a probability of new vent formation per 100sq km (*i.e.*, $\Delta x \times \Delta y$). White circles represent volcanic vent locations. Areas most likely to experience future volcano vent formation are shown by colored areas, estimated using an isotropic, 2 km bandwidth and a Gaussian kernel function at 100 m intervals across the map area. There is 99% probability that a future vent will form within the contour enclosing the white-shaded area and a 5% probability that a future vent will appear in the red zone. The scale bar gives the spatial density range for each shaded contour area.

Figure 5: An example of an elliptical bivariate Gaussian kernel density function shows closed contours that represent 1, 2, and 3 standard deviations.

The bandwidth ($\mathbf{H}$) is now expressed as a 2×2 element matrix,

$$\mathbf{H} = \begin{bmatrix} xx & rr \\ rr & yy \end{bmatrix}$$

that is positive definite (*important because the matrix must have a square root*). The elements of the $\mathbf{H}$ matrix have units of squared distance. The matrix $\mathbf{H}$ is a co-variance matrix that fully describes the shape and orientation of the anisotropic kernel function; xx is the variance of the kernel bandwidth in the x direction, yy is the variance in the y direction, and rr is directly proportional to the covariance of xx and yy ,

$$rr = \rho\sqrt{xx}\sqrt{yy},$$

where ρ is the correlation coefficient between bandwidth in the N–S and E–W directions. Therefore,

$$-\sqrt{xx}\sqrt{yy} \leq rr \leq \sqrt{xx}\sqrt{yy}$$

since $-1 \leq \rho \leq 1$, and rr can be thought of as reflecting the rotation of the bandwidth from north. The determinant of the matrix is $|\mathbf{H}|$, and $\mathbf{H}^{-1/2}$ is the square root of its inverse. The distance between the grid location and the volcanic event is now represented by a 1×2 element distance matrix ($\mathbf{d}$), $\mathbf{b}$ is the cross product of $\mathbf{d} \times \mathbf{H}^{-1/2}$, and $\mathbf{b}^T$ is the transpose of the cross product.

The implementation of equation 4 is nearly the same as the implementation of equation 2, except that it involves some additional linear algebra steps.

An examination of the distribution of vents in Figure 4 shows an elongation of vent locations in the N–S direction. The smoothing bandwidth can accommodate this elongation of vent locations with an anisotropic bandwidth, that specifies greater smoothing in the N–S direction and less smoothing in the E–W direction, as shown in Figure 4b.

Pseudocode is provided to implement equation 4 using a two-dimensional anisotropic Gaussian kernel.

Step 3: Estimating a smoothing bandwidth

A kernel bandwidth has to be specified for the implementation of equations 2 or 4. There is no fixed procedure for arriving at this smoothing bandwidth. One recommended, although very dated, approach is to choose different values of smoothing (h , equation 2 or $\mathbf{H}$, equation 4) and visually inspect the resulting map (*e.g.*, [Silverman, 1981](#)), basically using expert judgment to choose the appropriate smoothing. Different values can be chosen and plotted to show the change in spatial density, at a particular point of interest, as the smoothing changes ([Connor & Hill, 1995](#)).

The subjective nature of simply choosing a value for the smoothing bandwidth is prone to bias. Statistically, an unbiased estimator is preferred over a biased estimator since bias in the estimator will generate a hazard estimate that is very flexible, with a greater tendency to over- or under-estimate the true or optimal. A further difficulty with using elliptical bandwidths over radially symmetric bandwidths is that three values (xx, yy, rr) are needed to specify the smoothing bandwidth, rather than a single value.

A more modern approach chooses the bandwidth to optimize some criteria, in effect estimating it from the data ([Jones et al., 1996](#)). The basic idea is to minimize the error in predicting each point using the remainder, in a leave-one-out cross-validation scheme. If the objective function is the sum of expected squared distances, then the result is least squares cross validation (LSCV) ([Duong et al., 2007](#)). In two dimensions this tends to *overfit* the data in the sense that the kernels become quite elongated and narrow. To counter this tendency [Duong & Hazelton \(2003\)](#) developed the modified asymptotic mean integrated squared error (AMISE) method, resulting in the SAMSE pilot bandwidth ([Duong et al., 2007](#)). This method is designed to be far more robust to single events than the LSCV bandwidth, possibly at the expense of missing useful information ([Bebbington, 2013](#)).

Pseudocode to implement an anisotropic kernel

Begin: The model inputs

Specify the 4-element bandwidth matrix ($\mathbf{H}$):

smoothing in the E–W direction: xx

smoothing in the N–S direction: yy

bandwidth rotation to the East or to the West: rr

Specify the location (easting, northing) of each volcanic event: x_e, y_e

Specify the number of volcanic events: N

Specify the map boundaries: $west, east, south, north$

Specify the map grid spacing: $\Delta x, \Delta y$

Internal Variables

Initialize a variable to hold the spatial density at a grid location (x_g, y_g) : $\lambda_g = 0$

Initialize the sum of normalized spatial densities for the entire grid: $\Lambda = 0$

Find the determinant of matrix $\mathbf{H}$: $d_{\mathbf{H}} = |\mathbf{H}|$

Find the square root of $|\mathbf{H}|$: $(d_{\mathbf{H}})^{\frac{1}{2}}$

Find the square root of matrix $\mathbf{H}$: $\mathbf{H}^{\frac{1}{2}}$

Find the inverse of $\mathbf{H}^{\frac{1}{2}}$: $\mathbf{H}^{-\frac{1}{2}}$

Loop: Do for each grid location (x_g, y_g)

For ($y_g = south$ up to $y_g = north$) for each northing

For ($x_g = west$ up to $x_g = east$) for each easting

A. Initialize variables for summation and spatial density

$\Lambda = 0$

$\lambda_g = 0$

B. Loop: Do for each volcanic event (e)

For ($e = 0$ up to $e = N$)

1. Calculate the distance between the current grid location (x_g, y_g)
and each volcano event (x_e, y_e) :

a. Create a 2-element distance matrix: $\mathbf{d} = \begin{bmatrix} d_x & d_y \end{bmatrix}$,

where $d_x = x_g - x_e, d_y = y_g - y_e$

b. Calculate the cross product (*i.e.*, matrix multiply):

$\mathbf{b} = \mathbf{H}^{-\frac{1}{2}} \times \mathbf{d}$

c. Find the transpose of the cross product: $\mathbf{b}^T$

d. Calculate a *new* weighted distance: $d_w = \mathbf{b}^T \times \mathbf{b}$

2. Calculate the spatial density for this volcanic event:

$\lambda_e = \frac{1}{(2\pi\mathbf{H}^{\frac{1}{2}}N)} \times \exp(-0.5 \times d_w)$

3. Normalize this spatial density by the grid spacing: $\lambda_e = \lambda_e \times \Delta x \times \Delta y$

4. Sum the spatial density due to each volcanic event: $\lambda_g = \lambda_g + \lambda_e$

Next event (e)

C. Integrate across the entire grid space: $\Lambda = \Lambda + \lambda_g$

D. Print out the coordinate and spatial density at each grid location: print x_g, y_g, λ_g

Next grid easting x_g

Next grid northing y_g

End: Integration check: grid spatial densities should sum to 1: print Λ

A selection of automatic bandwidth estimators are found in the freely-available *ks* package, as supplemental to the *R* package of statistical computer codes. The PERL code provided in the supplementary material uses the SAMSE bandwidth estimator. Other variations exist, such as maximizing the Kullback-Leibler score,

$$S = \sum_k \log \hat{\lambda}(x_k, y_k) - \int \int \hat{\lambda}(x, y) dy dx, \quad (5)$$

in a *leave one out* cross-validation (Vere-Jones, 1992), where $\hat{\lambda}(x_k, y_k)$ is computed using all the locations except (x_k, y_k) itself. This technique better lends itself to a *hard boundary* (Bebbington, 2015) than the SAMSE or LSCV methods. Note that equation 5 is equivalent to the point process likelihood, and that the double integral should be unity.

A KDE implementation using the SAMSE bandwidth estimator on the set of volcanic vents located near Managua, Nicaragua, is shown in Figure 4c. The mathematically-derived, auto-generated bandwidth is entirely data-driven and eliminates subjective bias from the bandwidth selection process. This is a powerful process if one believes that *the location of previous events* is the sole driver of determining the next volcanic event location.

Step 4: Assessing the map

Since spatial density is a probability density function, the integral of spatial density across the grid space or map domain should approach a value of 1. This approximation assumes that a sufficiently large grid is used (*i.e.*, one that includes all of the volcanic events and extends a sufficient distance beyond all events, in all directions, such that the spatial density will be negligible at the map borders) and that the grid spacing is sufficiently small. The output of the KDE can be verified by summing the spatial density values over the map domain or grid space and confirming that the resulting value is very close to 1 (equation 3). If the map values sum to a number much less than 1, then increasing the map boundaries and trying again may give a better result; when the map domain is too small the spatial density contours will extend beyond the borders of the map with $\Lambda < 1$ (equation 3).

Example: Arsia Mons, Mars

KDE is also a valuable method to quantify the size and shape of volcanic fields on other planets with an objective to better understand how magma ascended and erupted in ancient volcanic terrains. A relatively young cluster of 29 low shield volcanoes comprises one such volcanic terrain on Mars, in the 110 km diameter caldera at the summit of Arsia Mons (Bleacher *et al.*, 2010). Volcanism at this site has likely been extinct for tens of millions of years (Richardson *et al.*, 2017), so the purpose of estimating the spatial density of vents is not to estimate locations of future vents. Instead, modes of vent spatial density indicate increased ability for the region to accumulate and/or erupt magma from given locations, which can provide insight into magma-tectonic interactions during epochs of volcanic activity.

Following the **anisotropic kernel pseudocode** algorithm, the spatial density map of the Arsia Mons field is determined with the SAMSE pilot bandwidth (Figure 6a). The resulting map shows two linear trends of vent spatial density, which are each parallel to graben sets on the flanks of Arsia Mons. This indicates that magma pathways during the formation of this field exploited pre-existing fractures associated with the graben, or were developed coincident with the graben.

Volumes have additionally been modeled for the lavas associated with each vent in the Arsia Field using interpolated subsurface models for each lava flow (Richardson, 2016). These volumes can be used to construct a volume spatial density model to identify loci of increased magma flux to the surface, as opposed to focusing on vent production in the region. This model is constructed by altering the code to weight each vent (e in equation 4) by its corresponding effused volume, v_e (*i.e.*, multiply the exponent on the right of the summation in equation 4 by the event volume), and normalize the spatial density model to unity by dividing the model by the entire volume erupted instead of the number of vents (*i.e.*, replace N with the total field volume to the left of the summation in equation 4)(also see (Martin *et al.*, 2004)):

Figure 6: Vent density (equation 4) (a) and volume density (equation 6) (b) maps of the Martian Arsia Mons volcanic field. The two lineaments seen in the vent density model are parallel with large rift graben that cut the flanks of the volcano. The volume density plot shows that lavas from the east lineament of vents, also visible in the vent density model, are more voluminous than those from the west lineament of vents. The total area enclosing 99.5% of the vent density in (a) is 10 300 km². Color bars are annotated with density contours (standard deviations) and non-normalized kernel smoothing values (vents or volume lava (km³) erupted per unit area (km²)) for both models.

$$\hat{\lambda}_{vol}(\mathbf{g}) = \frac{1}{2\pi\sqrt{|\mathbf{H}|} \left[\sum_{e=1}^N v_e \right]} \sum_{e=1}^N v_e \exp \left[-\frac{1}{2} \mathbf{b}^T \mathbf{b} \right], \quad (6)$$

where all notation is the same as in equation 4 and v_e is the volume of volcanic event e , including lavas. Incorporating this volume weighting method would occur in Step B.2 in the [anisotropic kernel pseudocode](#). Without normalization, this is *kernel smoothing* of the volume, $\hat{V}(\mathbf{g})$ (see [volume kernel smoothing pseudocode](#)), a generalization of the kernel density estimation:

$$\hat{V}(\mathbf{g}) = \frac{1}{2\pi\sqrt{|\mathbf{H}|}} \sum_{e=1}^N v_e \exp \left[-\frac{1}{2} \mathbf{b}^T \mathbf{b} \right]. \quad (7)$$

The resulting model (Figure 6b) shows more clearly how the eastern trend of volcanic vents was the dominant pathway for magma to reach the surface. The western trend delivered far less magma to the surface during development of this volcanic field. Thus, our understanding of Arsia Mons magmatism is improved by creating both maps (Figures 6a and b).

Pseudocode to implement anisotropic volume kernel smoothing

Begin: The model inputs

Specify the 4-element bandwidth matrix ($\mathbf{H}$):

smoothing in the E–W direction: xx

smoothing in the N–S direction: yy

bandwidth rotation to the East or to the West: rr

Specify the location (easting, northing) and eruptive volume of each volcanic event: x_e, y_e, v_e

Specify the number of volcanic events: N

Specify the map boundaries: $west, east, south, north$

Specify the map grid spacing: $\Delta x, \Delta y$

Internal Variables

Initialize a variable to hold the volume density at a grid location (x_g, y_g): $\hat{V}_g = 0$

Initialize the sum of volume densities for the entire grid: $\Lambda_V = 0$

Find the determinant of matrix $\mathbf{H}$: $d_{\mathbf{H}} = |\mathbf{H}|$

Find the square root of $|\mathbf{H}|$: $(d_{\mathbf{H}})^{\frac{1}{2}}$

Find the square root of matrix $\mathbf{H}$: $\mathbf{H}^{\frac{1}{2}}$

Find the inverse of $\mathbf{H}^{\frac{1}{2}}$: $\mathbf{H}^{-\frac{1}{2}}$

Loop: Do for each grid location (x_g, y_g)

For ($y_g = south$ up to $y_g = north$) for each northing

For ($x_g = west$ up to $x_g = east$) for each easting

A. Initialize variables for summation and spatial density

$\Lambda_V = 0$

$\hat{V}_g = 0$

B. Loop: Do for each volcanic event (e)

For ($e = 0$ up to $e = N$)

1. Calculate the distance between the current grid location (x_g, y_g)
and each volcano event (x_e, y_e):

a. Create a 2-element distance matrix: $\mathbf{d} = \begin{bmatrix} d_x & d_y \end{bmatrix}$,

where, $d_x = x_g - x_e, d_y = y_g - y_e$

b. Calculate the cross product (*i.e.*, matrix multiply):

$\mathbf{b} = \mathbf{H}^{-\frac{1}{2}} \times \mathbf{d}$

c. Find the transpose of the cross product: $\mathbf{b}^T$

d. Calculate a *new* weighted distance: $d_w = \mathbf{b}^T \times \mathbf{b}$

2. Calculate the volume density for this volcanic event:

$\lambda_e = \frac{1}{(2\pi\mathbf{H}^{\frac{1}{2}})} \times v_e \exp(-0.5 \times d_w)$

3. Sum the volume density due to each volcanic event: $\hat{V}_g = \hat{V}_g + \lambda_e$

Next event (e)

C. Integrate across the entire grid space: $\Lambda_V = \Lambda_V + \hat{V}_g$

D. Print out the coordinate and spatial density at each grid location: print $x_g, y_g, \hat{V}_g$

Next grid easting x_g

Next grid northing y_g

End: Integration check: grid volume densities should sum to total erupted volume: print Λ_V

Example: Lassen Volcanic System, California

Quaternary volcanic activity in the Lassen segment of the Cascades volcanic arc comprises a spatio-temporal succession of long-lived volcanic centers intercalated within a 4 km-thick (*Berge & Stauber, 1987*) regional volcanic platform (*Clynne & Muffler, 2010*). Volcanic activity in this region has been separated into two groups, regional mafic volcanism consisting of primitive magmas and the Lassen volcanic center (LVC) vents that erupt differentiated and contaminated / hybrid andesitic magmas. The regional mafic group contains both calc-alkaline and tholeiitic magmas, whereas the LVC erupts only calc-alkaline magmas (*Germa et al., 2019*).

Volcanoes belonging to the regional volcanism are small to medium size edifices, exhibit restricted compositional variations (basalts to andesite) during their short lifetime (a few millennia), and lack complex geochemistry, indicating direct ascent of primitive melts from the lower crust or the mantle. In contrast, LVC products are compositionally diverse (basalt to rhyolite), voluminous (hundreds of km³) and long-lived (0.5–1 Myr). Regional mafic vents are fed by independent primitive magma batches that result from mantle or lower crust melting and ascend without entrapment through the crust, whereas LVC vents develop from large upper crustal magma bodies that grow incrementally and persistent for extended eruptive periods (*Hildreth, 2007; Clynne & Muffler, 2010*). Interaction between mafic magma and country rock promotes crustal melting, which, along with mixing among andesitic and dacitic batches, leads to eruption of intermediate to silicic hybrid magmas at LVC vents (*Borge et al., 1997; Borge & Clynne, 1998*). No regional primary compositions are found at LVC vents (*Hildreth, 2007*).

We use mapped vents from the USGS Map 2899 (*Clynne & Muffler, 2010*) as well as vents digitized based on *Clynne & Muffler (2010)* and *Hildreth (2007)* to build spatial density models. A total of 306 vents are considered, including 253 regional mafic vents and 53 LVC vents. Each regional mafic vent is assumed to represent an individual magma batch that ascended directly from the mantle or the base of the crust with minimal entrapment. In contrast, LVC vents have erupted evolved and hybridized melts that have spent significant amount of time entrapped in crustal storage. Thus, the two spatial density models illustrate different magmatic processes.

We have not attempted to group vents into independent events in this model, even though some vents are very likely constructed during the same eruptive episode. For estimating volcanic hazards, the grouping of vents into events is important, but awaits additional data on the ages of these vents. Using the 253 and 53 vents, respectively, we find that the SAMSE estimated bandwidth matrix for LVC vents has determinant $|\mathbf{H}| = 12 \text{ km}^2$ and for regional mafic vents $|\mathbf{H}| = 114 \text{ km}^2$. This difference in the determinants of the best-fit bandwidth shows that LVC vents are much more tightly clustered than the regional mafic volcanic vents, a finding consistent with the geochemical model for these magmas.

Comparison of the spatial density maps for regional mafic vents (Figure 7a) and LVC vents (Figure 7b) illustrates that the LVC vents are strongly clustered. Furthermore, the LVC vents fill a gap in the distribution of regional mafic vents. This map pattern is consistent with the geochemical differences between the two groups and is consistent with the idea that localized magma flux is sufficient in the LVC area to produce the differentiated and contaminated magmas that erupt from these vents. Geophysical anomalies are mapped in the shallow crust corresponding to the potential mid-crustal magma source region for this cluster (*Blakely et al., 1997; Park & Ostos, 2013*). LVC volcanic hazards are distinct from the regional mafic vents *Clynne et al. (2012)*. Thus, mapping spatial density by geochemical trends is a powerful approach to understanding magmatism and improving hazard assessments.

Discussion

Uncertainty exists in estimates of spatial density. This uncertainty stems from: (i) ambiguity in the event data sets used to develop kernel estimates, including the degree to which the past predicts the future, (ii) uncertainty in the bandwidth estimate used in the kernel density estimation, (iii) too few event data, a common problem in hazard assessment, and to a lesser extent, (iv) choice of the kernel density function.

Event definition affects the total number of vents used to estimate spatial density. There is uncertainty about which mapped volcanoes should be included as a single event, due to eruption age or type, which volcanic vents formed at the same time, during the same episode of activity, and uncertainty about the completeness of the geologic record. For

Figure 7: Vent density for (a) regional mafic vents (b) LVC vents. Note that the LVC vents, with erupt differentiated and contaminated / hybrid andesitic magmas form a single cluster that fills a gap in the regional density model.

example, loci of activity may wax and wane with time, such that past vent patterns may not accurately forecast future volcano locations (*Condit & Connor, 1996; Tadini et al., 2014*).

Vent morphologies sometimes do not easily fit into a point process model on the scale of hazard maps. For example, lava flows sometimes erupt from highly elongate fissures, which only later might localize into vents (*Wylie et al., 1999; Valentine & Gregg, 2008*), as occurred for some regional mafic vents in the Lassen region (*Clynne & Muffler, 2010*).

Are temporal patterns present in the distribution of past events? If so, an appropriate time interval can be selected for the analysis (i.e., use only those vents that represent likely future patterns of activity, not patterns that are based on the locations of older volcano distributions) (*Conway et al., 1998*). These factors play a major role in the uncertainty in spatial density maps of future volcanic vent locations.

Bandwidth selection is a key feature of kernel density estimation, and is particularly relevant to volcanic hazard studies (*Bebbington, 2013, 2015; Jaquet et al., 2008*). Bandwidths that are small focus density near past events. Conversely, a large bandwidth may over-smooth the density estimate, resulting in unreasonably low density estimates near clusters of past events, and overestimate density far from past events. Bivariate bandwidth selectors like SAMSE methods remove some of the guess-work in bandwidth estimation since these selectors are based on the distribution of the input data.

One disadvantage of these bandwidth selectors is that they estimate a single *best-fit* bandwidth for the entire data set. Spatial density of volcanic vents might vary substantially in zones, such as inside or outside volcanic rifts, as in the Arsia Mons (Mars) example. An alternative method is to use adaptive kernel estimates, in which case the bandwidth changes with event density (*Weller et al., 2006*). An example is the *m*th nearest neighbor kernel estimate (*Connor & Hill, 1995; Bebbington, 2013*) where the bandwidth at any point in space is the distance to the *m*th nearest vent. This can be generalized to asymmetric kernels by having the kernel at a point be the bivariate Gaussian density best fitting the *m* nearest vents. Current and future research will likely involve further development and interpretation of bandwidth selectors that are adaptive across the map region.

Often in hazard assessment there is the *problem* that there are few data available from which to forecast future events. That is, often hazard assessments are needed for places where events are not so frequent that the geologic hazards are completely obvious. Instead, hazard analysis is most often required where few geologically hazardous events have occurred in the past. This is paradoxical because, by definition, uncertainty in hazard assessments must be comparatively high in these regions. If a spatial density is estimated using hundreds of volcanoes, we can assume that the true density is well-represented by this model. Conversely, if the spatial intensity estimate is based on a handful of events, we might

expect high uncertainty in the estimate. For example, the discovery of a single additional volcano, buried in sediment, might alter the shape of the estimated regional spatial density. Using the SAMSE or Kullback-Liebler bandwidth estimator protects against unusual location configurations that might result in large amounts of probability at great distances from the nearest vent.

One might hope that a complete understanding of the geology would result in a modification of the density estimate derived from a mathematical function. The Lassen example illustrates an approach of sub-dividing the vent dataset based on a geochemical model. The Arsia Mons example illustrates weighting the spatial density model by eruption volume. A number of authors have noted the correlation between geophysical anomalies and changes in spatial density (*Martin et al., 2004; Kiyosugi et al., 2010*), and have attempted to model this correlation (*Deng et al., 2017*). In the Lassen region, it appears that there is broad correlation between gravity (*Blakely et al., 1997*) and magnetotelluric (*Park & Ostos, 2013*) anomalies and the LVC cluster, consistent with the geochemical model.

One approach to incorporating such geologic data is to weight KDE using other data, such as fault maps or gravity anomalies associated with faults (e.g., *Connor et al., 2000; Galindo et al., 2016; Bevilacqua et al., 2017*). While definitely interesting and promising, these approaches can be quite sensitive to the weighting method used to develop the probability density function using additional geological data.

Hazards associated with the opening of new vents may be exasperated by the topography of volcanic systems, which is often complex and characterized by steep slopes. For example, small variations in vent location may cause lava to flow in a completely different direction down the flanks of a volcano (*Favalli et al., 2012; Cappello et al., 2012; Connor et al., 2012; Pérez & Walter, 2016*). There is no doubt that probabilistic models of lava flow inundation, like other geophysical flows, are quite sensitive to models of vent location.

Although kernel densities are usually unimodal, the mathematical machinery can be adapted to other forms. For example, if one wishes to test the hypothesis that existing vents are unlikely to re-erupt, or that the locus of volcanism has shifted away from older vent clusters, these may be cases where the kernel density maximum is not located at (x_e, y_e) . In other words, feel free to experiment with KDE!

Conclusions

The authors hope that by detailing their methodology for using KDE for spatial density analysis of volcanic vents others will be encouraged to use this powerful tool.

Acknowledgments

Software development and the analysis of Nicaragua vent distributions in the Managua Graben were funded through NASA grant 16-ESI-0030. Work performed for the Lassen volcano example was supported by a grant from the U. S. National Science Foundation (EAR 1347899). Work performed for the Arsia Mons (Mars) example was supported by a grant from NASA Mars Data Analysis Program (MDAP). We thank Antonio Costa for handling the editorial and two anonymous reviewers. Their comments helped improve this manuscript!

Supplementary material

Our computer code for calculating the spatial density of distributed volcanic events using KDE is included as supplementary material. The code is written in PERL, uses R (*R Core Team, 2019*) to calculate the SAMSE bandwidth, and depends on GMT (*contributors, 2019*), PROJ (*PROJ contributors, 2019*), and GDAL (*GDAL/OGR contributors, 2019*) to plot results. This code can be found at:

https://github.com/geoscience-community-codes/spatial_density

along with instructions for installation on a Linux computer, and a test dataset of volcano vent locations of the Nejapa volcano alignment near Managua, Nicaragua.

References

- AVELLÁN, D. R., MACÍAS, J. L., PARDO, N., SCOLAMACCHIA, T. & RODRIGUEZ, D. (2012) Stratigraphy, geomorphology, geochemistry and hazard implications of the Nejapa Volcanic Field, Western Managua, Nicaragua. *Journal of Volcanology and Geothermal Research*, **213–214**:51–71. DOI 10.1016/j.jvolgeores.2011.11.002. 5
- BARTOLINI, S., BOLÓS, X., MARTÍ, J., PEDRA, E. R. & PLANAGUMÀ, L. (2015) Hazard assessment at the Quaternary La Garrotxa Volcanic Field (NE Iberia). *Natural Hazards*, **78**(2):1349–1367. DOI 10.1007/s11069-015-1774-y. 2
- BARTOLINI, S., CAPPELLO, A., MARTÍ, J. & NEGRO, C. DEL (2013) QVAST: a new Quantum GIS plugin for estimating volcanic susceptibility. *Natural Hazards and Earth System Sciences*, **13**(11):3031. DOI 10.5194/nhess-13-3031-2013. 3
- BEBBINGTON, M. S. (2013) Assessing spatio-temporal eruption forecasts in a monogenetic volcanic field. *Journal of Volcanology and Geothermal Research*, **252**:14–28. DOI 10.1016/j.jvolgeores.2012.11.010. 5, 11, 17
- BEBBINGTON, M. S. (2015) Spatio-volumetric hazard estimation in the Auckland volcanic field. *Bulletin of Volcanology*, **77**(5):39. DOI 10.1007/s00445-015-0921-. 6, 13, 17
- BEBBINGTON, M. S. & CRONIN, S. J. (2011) Spatio-temporal hazard estimation in the Auckland Volcanic Field, New Zealand, with a new event-order model. *Bulletin of Volcanology*, **73**(1):55–72. DOI 10.1007/s00445-010-0403-6. 3, 6
- BERGE, P. A. & STAUBER, D. A. (1987) Seismic refraction study of upper crustal structure in the Lassen Peak area, Northern California. *Journal of Geophysical Research*, **92**(B10):10571–10579. DOI 10.1029/JB092iB10p10571. 16
- BERTIN, D., LINDSAY, J. M., BECERRIL, L., CRONIN, S. J. & BERTIN, L. J. (2019) MatHaz: a Matlab code to assist with probabilistic spatio-temporal volcanic hazard assessment in distributed volcanic fields. *Journal of Applied Volcanology*, **8**(1):4. DOI 10.1186/s13617-019-0084-6. 3
- BEVILACQUA, A., BURSIK, M., PATRA, A., PITMAN, E. B. & TILL, R. (2017) Bayesian construction of a long-term vent opening probability map in the Long Valley volcanic region (CA, USA). *Statistics in Volcanology*, **3**(1):1–36. DOI 10.5038/2163-338X.3.1. 3, 5, 18
- BEVILACQUA, A., ISAIA, R., NERI, A., VITALE, S., ASPINALL, W. P., BISSON, M., FLANDOLI, F., BAXTER, P. J., BERTAGNINI, A., ONGARO, T. ESPOSTI *et al.* (2015) Quantifying volcanic hazard at Campi Flegrei caldera (Italy) with uncertainty assessment: 1. Vent opening maps. *Journal of Geophysical Research: Solid Earth*, **120**(4):2309–2329. DOI 10.1002/2014JB011775. 2
- BLAKELY, R. J. R. L. CHRISTIANSEN, GUFFANTI, M., WELLS, R. E., DONNELLY-NOLAN, J. M., MUFFLER, L. J. P., CLYNNE, M. A. & SMITH, J. G. (1997) Gravity anomalies, quaternary vents, and quaternary faults in the southern Cascade Range, Oregon and California: Implications for arc and backarc evolution. *Journal of Geophysical Research*, **102**(B10):22513–22527. DOI 10.1029/97JB01516. 16, 18
- BLEACHER, J. E., RICHARDSON, J. A., RICHARDSON, P. W., GLAZE, L. S., BALOGA, S. S. M., GREELEY, R., HAUBER, E. & LILLIS, R. R. J. (2010) Updates to the catalog of Tharsis Province Small Volcanic Vents, Mars. In *41st Lunar and Planetary Science Conference*, LPI 1533, 1615. The Woodlands, Texas, US. ID 2010LPI....41.1615B. 13
- BORGE, L. E. & CLYNNE, M. A. (1998) The petrogenesis of felsic calc-alkaline magmas from the southernmost Cascades, California: Origin by partial melting of basaltic lower crust. *Journal of Petrology*, **39**(6):1197–1222. DOI /10.1093/petroj/39.6.1197. 16

- BORGE, L. E., CLYNNE, M. A. & BULLEN, T. D. (1997) The variable role of slab-derived fluids in the generation of a suite of primitive calc-alkaline lavas from the southernmost Cascades, California. *The Canadian Mineralogist*, **35**:425–452. **ARTICLE**. 16
- CAPPELLO, A., NERI, M., ACOCELLA, V., GALLO, G., VICARI, A. & NEGRO, C. DEL (2012) Spatial vent opening probability map of Etna volcano (Sicily, Italy). *Bulletin of Volcanology*, **74**(9):2083–2094. DOI 10.1007/s00445-012-0647-4. 3, 18
- CLYNNE, M. A. & MUFFLER, L. J. P. (2010) Geologic Map of Lassen Volcanic National Park and Vicinity, California. Scientific investigations map 2899, scale 1:50,000, U. S. Department of the Interior, U. S. Geological Survey, Reston, VA. **MAP 2899**. 16, 17
- CLYNNE, M. A., ROBINSON, J. E., NATHENSON, M. & MUFFLER, L. J. P. (2012) Volcano hazards assessment for the Lassen region, Northern California. Scientific investigations report 2012–5176–a, 1 plate, scale 1:200,000, U. S. Department of the Interior, U. S. Geological Survey, Reston, VA. **REPORT 2012–5176–A**. 16
- CONDIT, C. D. & CONNOR, C. B. (1996) Recurrence rate of basaltic volcanism in volcanic fields: An example from the Springerville Volcanic Field, AZ, USA. *Geological Society of America, Bulletin*, **108**(10):1225–1241. DOI 10.1130/0016-7606(1996)108<1225:RROVIB>2.3.CO;2 . 5, 17
- CONNOR, C. B., BEBBINGTON, M. S. & MARZOCCHI, W. (2015) Probabilistic volcanic hazard assessment. In H. SIGURDSSON (ed.) *The Encyclopedia of Volcanoes (Second Edition)*, 897–910. Elsevier. DOI 10.1016/B978-0-12-385938-9.00051-1. 2
- CONNOR, C. B. & CONNOR, L. J. (2009) Estimating spatial density with kernel methods. In C.B. CONNOR, N.A. CHAPMAN & L.J. CONNOR (eds.) *Volcanic and Tectonic Hazard Assessment for Nuclear Facilities*, 331–343. Cambridge University Press. DOI 10.1017/CBO9780511635380.015. 2
- CONNOR, C. B. & HILL, B. E. (1995) Three nonhomogeneous Poisson models for the probability of basaltic volcanism: Application to the Yucca Mountain region. *Journal of Geophysical Research: Solid Earth*, **100**(B6):10 107–10 125. DOI 10.1029/95JB01055. 3, 11, 17
- CONNOR, C. B., STAMATAKOS, J. A., FERRILL, D. A., HILL, B. E., OFOEGBU, G. I., CONWAY, F. M., SAGAR, B. & TRAPP, J. (2000) Geologic factors controlling patterns of small-volume basaltic volcanism: Application to a volcanic hazards assessment at Yucca Mountain, Nevada. *Journal of Geophysical Research: Solid Earth*, **105**(B1):417–432. DOI 10.1029/1999JB900353. 5, 18
- CONNOR, L. J., CONNOR, C. B., MELIKSETIAN, K. & SAVOV, I. (2012) Probabilistic approach to modeling lava flow inundation: a lava flow hazard assessment for a nuclear facility in Armenia. *Journal of Applied Volcanology*, **1**(3):doi:10.1186/2191–5040–1–3. DOI 10.1186/2191-5040-1-3. 3, 18
- CONTRIBUTORS, GMT (2019) *GMT The Generic Mapping Tools*. School of Ocean and Earth Science and Technology of the University of Hawai’i at Manoa. URL <https://www.generic-mapping-tools.org/>. 2, 18
- CONWAY, M. F., CONNOR, C. B., HILL, B. E., CONDIT, C. D., MULLANEY, K. & HALL, C. M. (1998) Recurrence rates of basaltic volcanism in the SP cluster, San Francisco volcanic field, Arizona. *Geology*, **26**(7):655–658. DOI 10.1130/0091-7613(1998)026<0655:RROBVI>2.3.CO;2. 17
- COPPERSMITH, K. J., JENNI, K. E., PERMAN, R. C. & YOUNGS, R. R. (2009) Formal expert assessment in probabilistic seismic and volcanic hazard analysis. In C.B. CONNOR, N.A. CHAPMAN & L.J. CONNOR (eds.) *Volcanic and Tectonic Hazard Assessment for Nuclear Facilities*, 593–611. Cambridge University Press. DOI 10.1017/CBO9780511635380.027. 3
- DE LA CRUZ-REYNA, S., MENDOZA-ROSAS, A. T., BORSELLI, L. & SAROCCHI, D. (2019) Volcanic hazard estimations for Volcán de Colima. In N. VARLEY, C. CONNOR & J.C. KOMOROWSKI (eds.) *Volcán de Colima: Portrait of a Persistently Hazardous Volcano*, 267–289. Springer, Berlin, Heidelberg. DOI 10.1007/978-3-642-25911-1_6. 3

- DENG, F., CONNOR, C. B., MALSERVISI, R., CONNOR, L. J., WHITE, J. T., GERMA, A. & WETMORE, P. H. (2017) A geophysical model for the origin of volcano vent clusters in a Colorado Plateau volcanic field. *Journal of Geophysical Research: Solid Earth*, **122**:8910–8924. DOI 10.1002/2017JB014434. 3, 18
- DIEZ, M., LA FEMINA, P. C., CONNOR, C. B., STRAUCH, W. & TENORIO, V. (2005) Evidence for static stress changes triggering the 1999 eruption of Cerro Negro Volcano, Nicaragua and regional aftershock sequences. *Geophysical Research Letters*, **32**(4). DOI 10.1029/2004GL021788. 5
- DUONG, T. & HAZELTON, M. (2003) Plug-in bandwidth matrices for bivariate kernel density estimation. *Journal of Nonparametric Statistics*, **15**(1):17–30. DOI 10.1080/1048525030606039. 11
- DUONG, T. *et al.* (2007) ks: Kernel density estimation and kernel discriminant analysis for multivariate data in R. *Journal of Statistical Software*, **21**(7):1–16. DOI 10.18637/jss.v021.i07. 2, 11
- EL DIFRAWY, M. A., RUNGE, M. G., MOUFTI, M. R., CRONIN, S. J. & BEBBINGTON, M. (2013) A first hazard analysis of the Quaternary Harrat Al-Madinah volcanic field, Saudi Arabia. *Journal of Volcanology and Geothermal Research*, **267**:39–46. DOI 10.1016/j.jvolgeores.2013.09.006. 2
- FAVALLI, M., KARÁTON, D., MAZZARINI, F., PARESCHI, M. T. & BOSCHI, E. (2009) Morphometry of scoria cones located on a volcano flank: a case study from Mt. Etna (Italy), based on high-resolution LiDAR data. *Journal of Volcanology and Geothermal Research*, **186**(3–4):320–330. DOI 10.1016/j.jvolgeores.2009.07.011. 3
- FAVALLI, M., TARQUINI, S., PAPALE, P., FORNACIAI, A. & BOSCHI, E. (2012) Lava flow hazard and risk at Mt. Cameroon volcano. *Bulletin of Volcanology*, **74**(2):423–439. DOI 10.1007/s00445-011-0540-6. 3, 18
- FELPETO, A., MARTÍ, J. & ORTIZ, R. (2007) Automatic GIS-based system for volcanic hazard assessment. *Journal of Volcanology and Geothermal Research*, **166**(2):106–116. DOI 10.1016/j.jvolgeores.2007.07.008. 3
- FREUNDT, A., HARTMANN, A., KUTTEROLF, S. & STRAUCH, W. (2010) Volcaniclastic stratigraphy of the Tiscapa maar crater walls (Managua, Nicaragua): implications for volcanic and seismic hazards and Holocene climate changes. *International Journal of Earth Sciences*, **99**(6):1453–1470. DOI 10.1007/s00531-009-0469-6. 3
- GALINDO, I., ROMERO, M. C., SÁNCHEZ, N. & MORALES, J. M. (2016) Quantitative volcanic susceptibility analysis of Lanzarote and Chinijo Islands based on kernel density estimation via a linear diffusion process. *Scientific Reports*, **6**:27381. DOI 10.1038/srep27381. 2, 18
- GALLANT, E., RICHARDSON, J., CONNOR, C., WETMORE, P. & CONNOR, L. J. (2018) A new approach to probabilistic lava flow hazard assessments, applied to the Idaho National Laboratory, eastern Snake River Plain, Idaho, USA. *Geology*, **46**(10):895–898. DOI 10.1130/G45123.1. 5
- GDAL/OGR CONTRIBUTORS (2019) *GDAL/OGR Geospatial Data Abstraction software Library*. Open Source Geospatial Foundation. URL <https://gdal.org>. 18
- GEORGE, O. A., MCILRATH, J., FARRELL, A., GALLANT, E., TAVAREZ, S., MARSHALL, A., MCNIFF, C., NJORGE, M., WILSON, J., CONNOR, C. B., CONNOR, L. J. & KRUSE, S. (2015) High-resolution ground-based magnetic survey of a buried volcano: Anomaly B, Amargosa Desert, NV. *Statistics in Volcanology*, **1**(1):1–23. DOI 10.5038/2163-338X.1.3. 5
- GERMA, A., CONNOR, L. J., CANON-TAPIA, E. & LECORVEC, N. (2013) Tectonic and magmatic controls on the location of post-subduction monogenetic volcanoes in Baja California, Mexico, revealed through spatial analysis of eruptive vents. *Bulletin of Volcanology*, **75**(12):782. DOI 10.1007/s00445-013-0782-6. 2, 3, 5
- GERMA, A., PERRY, C., QUIDELLEUR, X., CALVERT, A., CLYNNE, M., CONNOR, C. B., CONNOR, L. J., MALSERVISI, R. & CHARBONNIER, S. (2019) Temporal relationship between the Lassen volcanic center and mafic regional volcanism. *Bulletin of Volcanology*, **81**:38. DOI 10.1007/s00445-019-1296-7. 3, 16

- GONZÁLEZ, M., BRETÓN, RAMÍREZ, J. J. & NAVARRO, C. (2002) Summary of the historical eruptive activity of Volcán de Colima, Mexico 1519–2000. *Journal of Volcanology and Geothermal Research*, **117**(1-2):21–46. DOI 10.1016/S0377-0273(02)00233-0. 3
- HASENAKA, T. & CARMICHAEL, I. S. E. (1985) The cinder cones of Michoacán-Guanajuato, central Mexico: Their age, volume and distribution, and magma discharge rate. *Journal of Volcanology and Geothermal Research*, **25**(1-2):105–124. DOI 10.1016/0377-0273(85)90007-1. 2
- HILDRETH, W. (2007) Quaternary magmatism in the Cascades: Geologic perspectives. Professional Paper, USGS Numbered Series 1744, U. S. Department of the Interior, US Geological Survey, Reston, VA. DOI 10.3133/pp1744. 16
- JAQUET, O., CONNOR, C. B. & CONNOR, L. J. (2008) Probabilistic modeling for long-term assessment of volcanic hazards. *Nuclear Technology*, **163**(1):180–189. DOI 10.13182/NT08-A3980. 3, 17
- JONES, M. C., MARRON, J. S. & SHEATHER, S. J. (1996) A brief survey of bandwidth selection for density estimation. *Journal of the American Statistical Association*, **91**(433):401–407. DOI 10.1080/01621459.1996.10476701. 2, 11
- KIYOSUGI, K., CONNOR, C. B., WETMORE, P. H., FERWERDA, B. P., GERMA, A. M., CONNOR, L. J. & HINTZ, A. R. (2012) Relationship between dike and volcanic conduit distribution in a highly eroded monogenetic volcanic field, San Rafael, Utah, USA. *Geology*, **40**(8):695 – 698. DOI 10.1130/G33074.1. 5
- KIYOSUGI, K., CONNOR, C. B., ZHAO, D., CONNOR, L. J. & TANAKA, K. (2010) Relationships between temporal-spatial distribution of monogenetic volcanoes, crustal structure, and mantle velocity anomalies: An example from the Abu Monogenetic Volcano Group, SW Japan. *Bulletin of Volcanology*, **72**(3):331–340. DOI 10.1007/s00445-009-0316-4. 18
- KUTTEROLF, S., FREUNDT, A. & BURKERT, C. (2011) Eruptive history and magmatic evolution of the 1.9 kyr Plinian dacitic Chiltepe Tephra from Apoyeque volcano in west-central Nicaragua. *Bulletin of Volcanology*, **73**(7):811–831. DOI 10.1007/s00445-011-0457-0. 3
- LA FEMINA, P.C., DIXON, T. H. & STRAUCH, W. (2002) Bookshelf faulting in Nicaragua. *Geology*, **30**(8):751–754. DOI 10.1130/0091-7613(2002)030<0751:BFIN>2.0.CO;2 . 5
- LUTZ, T. M. & GUTMANN, J. T. (1995) An improved method for determining and characterizing alignments of pointlike features and its implications for the Pinacate volcanic field, Sonora, Mexico. *Journal of Geophysical Research: Solid Earth*, **100**(B9):17659–17670. DOI 10.1029/95JB01058. 3
- MAGILL, C. R., MCANENEY, K. J. & SMITH, I. E. M. (2005) Probabilistic assessment of vent locations for the next Auckland volcanic field event. *Mathematical Geology*, **37**(3):227–242. DOI 10.1007/s11004-005-1556-2. 6
- MARTIN, A. J., UMEDA, K., CONNOR, C. B., WELLER, J. N., ZHAO, D. & TAKAHASHI, M. (2004) Modeling long-term volcanic hazards through Bayesian inference: example from the Tohoku volcanic arc, Japan. *Journal of Geophysical Research: Solid Earth*, **109**(B10208). DOI 10.1029/2004JB003201. 6, 13, 18
- MAZZARINI, F., CORVEC, N. LE, ISOLA, I., FAVALLI, M. & MASSIMILIANO, F. (2016) Volcanic field elongation, vent distribution, and tectonic evolution of a continental rift: The Main Ethiopian Rift example. *Geosphere*, **12**(3):706–720. DOI 10.1130/GES01193.1. 2
- MCBIRNEY, A. R. (1955) The origin of the Nejapa pits near Managua, Nicaragua. *Bulletin of Volcanology*, **17**(1):145–154. DOI 10.1007/BF02596053. 3
- NAGAOKA, S. & OKUNO, M. (2011) Tephrochronology and eruptive history of Kirishima volcano in southern Japan. *Quaternary International*, **246**(1-2):260–269. DOI 10.1016/j.quaint.2011.06.007. 3
- NAKADA, S., NAGAI, M., KANEKO, T., SUZUKI, Y. & MAENO, F. (2013) The outline of the 2011 eruption at Shinmoe-dake (Kirishima), Japan. *Earth, Planets and Space*, **65**(6):1. DOI 10.5047/eps.2013.03.016. 3

- NAKAMURA, K. (1977) Volcanoes as possible indicators of tectonic stress orientation – principle and proposal. *Journal of Volcanology and Geothermal Research*, **2**(1):1–16. DOI 10.1016/0377-0273(77)90012-9. 3
- PARDO, N., AVELLÁN, D. R., MACÍAS, J. L., SCOLAMACCHIA, T. & RODRÍGUEZ, D. (2008) The 1245 yr BP Asososca maar: New advances on recent volcanic stratigraphy of Managua (Nicaragua) and hazard implications. *Journal of Volcanology and Geothermal Research*, **176**(4):493–512. DOI 10.1016/j.jvolgeores.2008.04.020. 3
- PARDO, N., MACIAS, J. L., GIORDANO, G., CIANFARRA, P., AVELLÁN, D. R. & BELLATRECCIA, F. (2009) The 1245 yr BP Asososca maar eruption: The youngest event along the Nejapa–Mirafleres volcanic fault, Western Managua, Nicaragua. *Journal of Volcanology and Geothermal Research*, **184**(3-4):292–312. DOI 10.1016/j.jvolgeores.2009.04.006. 5
- PARK, S. K. & OSTOS, L. C. (2013) Constraints from magnetotelluric measurements on magmatic processes and upper mantle structure in the vicinity of Lassen volcanic center, Northern California. *Geosphere*, **9**(3):382–393. DOI 10.1130/GES00799.1. 16, 18
- PÉREZ, N. M. & WALTER, T. R. (2016) Lava flow hazard at Fogo Volcano, Cabo Verde, before and after the 2014–2015 eruption. *Natural Hazards and Earth System Sciences*, **16**(8):1925. DOI 10.5194/nhess-16-1925-2016. 18
- POLACCI, M. & PAPALE, P. (1997) The evolution of lava flows from ephemeral vents at Mount Etna: insights from vesicle distribution and morphological studies. *Journal of Volcanology and Geothermal Research*, **76**(1-2):1–17. DOI 10.1016/S0377-0273(96)00070-4. 2
- PROJ CONTRIBUTORS (2019) *PROJ coordinate transformation software library*. Open Source Geospatial Foundation. URL <https://proj.org/>. 5, 18
- R CORE TEAM (2019) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org>. 18
- RAUSCH, J. & SCHMINCKE, H.-U. (2010) Nejapa Tephra: the youngest (c. 1 ka BP) highly explosive hydroclastic eruption in Western Managua (Nicaragua). *Journal of Volcanology and Geothermal Research*, **192**(3-4):159–177. DOI 10.1016/j.jvolgeores.2010.02.010. 5
- RECHES, Z. & FINK, J. (1988) The mechanism of intrusion of the Inyo dike, Long Valley Caldera, California. *Journal of Geophysical Research: Solid Earth*, **93**(B5):4321–4334. DOI 10.1029/JB093iB05p04321. 5
- RICHARDSON, J. A. (2016) Modeling the construction and evolution of distributed volcanic fields on Earth and Mars. *Graduate Theses and Dissertations*. THESIS 6136. 13
- RICHARDSON, J. A., A., J., CONNOR, C. B., BLEACHER, J. E. & KIYOSUGI, K. (2017) Recurrence rate and magma effusion rate for the latest volcanism on Arsia Mons, Mars. *Earth and Planetary Science Letters*, **458**:170–178. ISSN 0012821X. DOI 10.1016/j.epsl.2016.10.040. 13
- RICHARDSON, J. A., BLEACHER, J. E., CONNOR, C. B. & CONNOR, L. J. (2012) Using spatial density to characterize volcanic fields on Mars. In 43rd Lunar and Planetary Science Conference, GSFC.CP.00129.2012, 2. NASA, The Woodlands, Texas, US. ID 20120009829. 2
- RIGGS, N., ORT, M., CONNOR, C. B., ALFANO, F. & CONWAY, M. (2019) Volcanology and associated hazards of the San Francisco volcanic field. In P.A. PEARTHREE (ed.) *Geological Excursions in Southwestern North America*, volume 55, 127–146. Geological Society of America. DOI 10.1130/2019.0055(05). 2
- RITTMANN, A. (1962) *Volcanoes and Their Activity*. Interscience Publishers, John Wiley & Sons, New York/London. ISBN 0471724505. 3
- RUNGE, M. G., BEBBINGTON, M. S., CRONIN, S. J., LINDSAY, J. M., KENEDI, C. L. & MOUFTI, M. R. H. (2014) Vents to events: determining an eruption event record from volcanic vent structures for the Harrat Rahat, Saudi Arabia. *Bulletin of Volcanology*, **76**(3):804. DOI 10.1007/s00445-014-0804-z. 3, 5

- SCAINIA, C., FELPETO, A., MARTÍ & CARNIEL, R. (2014) A GIS-based methodology for the estimation of potential volcanic damage and its application to Tenerife Island, Spain. *Journal of Volcanology and Geothermal Research*, **278–279**:40–58. DOI [10.1016/j.jvolgeores.2014.04.005](https://doi.org/10.1016/j.jvolgeores.2014.04.005). 3
- SILVERMAN, B. W. (1981) Using kernel density estimates to investigate multimodality. *Journal of the Royal Statistical Society. Series B (Methodological)*, **3**(1):97–99. URL. 2, 11
- TADINI, A., BEVILACQUA, A., NERI, A., CIONI, R., ASPINALL, W. P., BISSON, M., ISAIA, R., MAZZARINI, F., VALENTINE, G. A., VITALE, S. *et al.* (2017) Assessing future vent opening locations at the Somma-Vesuvio volcanic complex: 2. Probability maps of the caldera for a future Plinian/sub-Plinian event with uncertainty quantification. *Journal of Geophysical Research: Solid Earth*. DOI [10.1002/2016JB013860](https://doi.org/10.1002/2016JB013860). 2
- TADINI, A., BONALI, F. L., CORAZZATO, C., CORTÉS, J. A., TIBALDI, A. & VALENTINE, G. A. (2014) Spatial distribution and structural analysis of vents in the Lunar Crater Volcanic Field (Nevada, USA). *Bulletin of Volcanology*, **76**(11):877. DOI [10.1007/s00445-014-0877-8](https://doi.org/10.1007/s00445-014-0877-8). 5, 17
- TAKADA, A. (1997) Cyclic flank-vent and central-vent eruption patterns. *Bulletin of Volcanology*, **58**(7):539–556. doi [10.1007/s004450050161](https://doi.org/10.1007/s004450050161). 3
- VALENTINE, G. A. & CONNOR, C. B. (2015) Basaltic volcanic fields. In *The Encyclopedia of Volcanoes (Second Edition)*, 423–439. Elsevier. DOI [10.1016/B978-0-12-385938-9.00023-7](https://doi.org/10.1016/B978-0-12-385938-9.00023-7). 2
- VALENTINE, G. A. & GREGG, T. K. P. (2008) Continental basaltic volcanoes – processes and problems. *Journal of Volcanology and Geothermal Research*, **177**(4):857–873. DOI [10.1016/j.jvolgeores.2008.01.050](https://doi.org/10.1016/j.jvolgeores.2008.01.050). 17
- VALENTINE, G. A. & PERRY, F. V. (2006) Decreasing magmatic footprints of individual volcanoes in a waning basaltic field. *Geophysical Research Letters*, **33**(14). DOI [10.1029/2006GL026743](https://doi.org/10.1029/2006GL026743). 5
- VAN DEN HOVE, J., GROSE, L., BETTS, P. G., AILLERES, L., OTTERLOO, J. VAN & CAS, R. A. F. (2017) Spatial analysis of an intra-plate basaltic volcanic field in a compressional tectonic setting: South-eastern Australia. *Journal of Volcanology and Geothermal Research*, **335**:35–53. DOI [10.1016/j.jvolgeores.2017.02.001](https://doi.org/10.1016/j.jvolgeores.2017.02.001). 2
- VERE-JONES, D. (1992) Statistical methods for the description and display of earthquake catalogs. In A.T. WALDEN & P. GUTTORP (eds.) *Statistics in the Environmental and Earth Sciences: New Developments in Theory and Practice*, 220–246. Hodder Education Publishers. ISBN 0340545305. 13
- WALKER, J. A. (1984) Volcanic rocks from the Nejapa and Granada cinder cone alignments, Nicaragua, Central America. *Journal of Petrology*, **25**(2):299–342. DOI [10.1093/petrology/25.2.299](https://doi.org/10.1093/petrology/25.2.299). 3
- WAND, M. P. & JONES, M. C. (1993) Comparison of smoothing parameterizations in bivariate kernel density estimation. *Journal of the American Statistical Association*, **88**(422):520–528. DOI [10.1080/01621459.1993.10476303](https://doi.org/10.1080/01621459.1993.10476303). 9
- WAND, M. P. & JONES, M. C. (1994) *Kernel Smoothing*. Monographs on Statistics and Applied Probability. Chapman and Hall/CRC. ISBN 9780412552700. 2
- WELLER, J. N., MARTIN, A. J., CONNOR, C. B., CONNOR, L. J. & KARAKHANIAN, A. (2006) Modelling the spatial distribution of volcanoes: An example from Armenia. In H.M. MADER, S.G. COLES, C.B. CONNOR & L.J. CONNOR (eds.) *Statistics in Volcanology*, number 1 in Special Publications of IAVCEI, 77–88. Geological Society of London. DOI doi.org/10.1144/IAVCEI001.7. 17
- WESSEL, P., LUIS, J., UIEDA, L., SCHARROO, R., WOBBE, F., SMITH, W.H.F. & TIAN, D. (2019) The Generic Mapping Tools version 6. *Geochemistry, Geophysics, Geosystems*, **20**:1–9. [10.1029/2019GC008515](https://doi.org/10.1029/2019GC008515). 2
- WETMORE, P. H., HUGHES, S. S., CONNOR, L. J. & CAPLINGER, M. (2009) Spatial distribution of eruptive centers about the Idaho National Laboratory. In C.B. CONNOR, N.A. CHAPMAN & L.J. CONNOR (eds.) *Volcanic and Tectonic Hazard Assessment for Nuclear Facilities*, 385–405. Cambridge University Press. DOI [10.1017/CBO9780511635380.017](https://doi.org/10.1017/CBO9780511635380.017). 5

- WILLIAMS, H. (1950) Volcanoes of the Parícutin Region, Mexico. USGS Numbered Series, Bulletin 965-B, U.S. Government Printing Office. DOI [10.3133/b965B](https://doi.org/10.3133/b965B). 3
- WYLIE, J. J., HELFRICH, K. R., DADE, B., LISTER, J. R. & SALZIG, J. F. (1999) Flow localization in fissure eruptions. *Bulletin of Volcanology*, **60**(6):432–440. DOI [10.1007/s004450050243](https://doi.org/10.1007/s004450050243). 17