Bulletin of Volcanology (2019) 81:51
https://doi.org/10.1007/s00445-019-1310-0

RESEARCH ARTICLE

A new method to identify the source vent location of tephra fall deposits: development, testing, and application to key Quaternary eruptions of Western North America

Qingyuan Yang[1] · Marcus Bursik[1] · E. Bruce Pitman[2]

Received: 19 April 2018 / Accepted: 11 July 2019

Abstract
A new method to identify the source vent location of tephra fall deposits based on thickness or maximum clast size measurements is presented in this work. It couples a first-order gradient descent method with either one of two commonly used semi-empirical models of tephra thickness distribution. The method is applied to three tephra thickness and one maximum clast size datasets of the North Mono and Fogo A tephra deposits. Randomly selected and localized subsets of these datasets are used as input to evaluate its performance. The results suggest the utility of the method and show that estimating the dispersal axis is a more robust way to constrain the vent location compared with directly estimating the vent coordinates given sparse observations. Local change in dispersal direction can be detected given localized observations. Bootstrap aggregating and visualizing the surface of the cost function are used to analyze epistemic uncertainty for the method. Our discussion focuses on how different features of tephra deposits and technical aspects of the method would affect the performance of the method. Suggestions on how to use the method given limited observations are listed. One subset of the North Mono Bed 1 thickness dataset and thickness datasets of the Trego Hot Springs and Rockland tephras are used as case studies. The method is then applied to the well-correlated tephra subunits within the Wilson Creek Formation to estimate their vent location and volume. The simplicity and flexibility of the method make it a potentially useful tool for analyzing tephra fall deposits.

Keywords Tephra source vent · Thickness and maximum clast size distribution · Tephra correlation · Eruption parameter estimation · Inverse method

Introduction

Identifying the source vent location of tephra fall deposits is critical to the reconstruction of volcanic eruptions. Given the source vent location, characteristics of fall deposits at different sample sites can be integrated in a systematic way (e.g., Walker and Croasdale 1971b; Self 1983; Sieh and Bursik 1986; Engwell et al. 2013; Klawonn et al. 2014), which is a necessary step towards further interpretation and quantification (e.g., Suzuki 1983; Carey and Sparks 1986; Burden et al. 2011; Pyle 1989; Bursik et al. 1992b; Koyaguchi 1994; Connor and Connor 2006; Biass and Bonadonna 2011).

For fall deposits near vent, the source location is commonly identified based on isopach and isopleth mapping and direct observation (e.g., Walker and Croasdale 1971b; Sieh and Bursik 1986; Miller 1985; Bonadonna et al. 2002; Bursik et al. 2014). This is subjective and may fail given insufficient sample sites. In such cases, identifying the source vent of a tephra deposit requires extra knowledge of the specific deposit and its potential vents (e.g., Green et al. 2014; Kawabata et al. 2015, 2016). Variable data quality and format hamper the construction of a universal solution. A generalized way of estimating eruption parameters, including the vent position, is to combine inverse methods with tephra transport models

Editorial responsibility: J. Dufek

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s00445-019-1310-0) contains supplementary material, which is available to authorized users.

✉ Qingyuan Yang
qyang5@buffalo.edu

[1] Department of Geology, University at Buffalo, 126 Cooke Hall, Buffalo, NY 14260, USA

[2] Department of Materials Design and Innovation, University at Buffalo, 120 Bonner Hall, Buffalo, NY 14260, USA

Published online: 13 August 2019

(e.g., Connor and Connor 2006; Volentik et al. 2010; Bonasia et al. 2010; Klawonn et al. 2012; Burden et al. 2013; Kawabata et al. 2013; White et al. 2017). Most of these models require the vent coordinates as input or are very sensitive to a prescribed vent location. Although sensitivity analyses have been conducted on these models (Scollo et al. 2008; Bonadonna et al. 2015; Pouget et al. 2016), their high-dimensional parameter space can still introduce uncertainties to the problem. This stresses the importance of uncertainty quantification (Biass and Bonadonna 2011; Biasse et al. 2014; Green et al. 2016; White et al. 2017), which cannot be easily done given sparse observations.

For tephra found at distal sites, the identification of the source vent relies heavily on geochemical analyses and geologic records (e.g., Wood 1977; Turner et al. 2009; Pouget et al. 2014a; Marcaida 2015; Sigl et al. 2015; Marcaida et al. 2019). The correlation with a certain vent is determined by integrating different lines of evidence, such as similarity in age and composition.

Two semi-empirical models (Gonzalez-Mellado and Cruz-Reyna 2010; Yang and Bursik 2016) have been developed to analyze tephra thickness distributions for different purposes (e.g., Rhoades et al. 2002; Kawabata et al. 2013, 2016; Yang and Bursik 2016). Kawabata et al. (2013, 2015) have used the semi-empirical model proposed by Gonzalez-Mellado and Cruz-Reyna (2010) together with statistical approaches to recognize the number of lobes and identify the source of each lobe for different deposits, with prior knowledge of potential vent locations. The semi-empirical model proposed by Yang and Bursik (2016) is used for isopach mapping. These models describe the tephra thickness distribution as a function of location with respect to the source vent and are similar to a cone that is stretched and rotated along the dispersal axis. This particular geometry is mainly caused by turbulent diffusion, wind advection, and horizontal spreading of the volcanic plume at the neutral buoyancy level (Csanady 1973; Sparks et al. 1991; Bursik et al. 1992a; Rhoades et al. 2002; Bonadonna and Phillips 2003; Gonzalez-Mellado and Cruz-Reyna 2010; Costa et al. 2013). The two semi-empirical models have distinct functional forms describing how tephra deposits thin with distance from the vent.

In this study, we present a new algorithm that can be used to identify the source vent location and dispersal axis of tephra fall deposits based on thickness or maximum clast size measurements. It couples either of the two semi-empirical models of tephra thickness distribution (Gonzalez-Mellado and Cruz-Reyna 2010; Yang and Bursik 2016) with a first-order gradient descent method. Our working hypotheses are:

1. This method can be used to identify the vent location of tephra deposits based on thickness or maximum clast size measurements.
2. Given sparse observations, compared with estimating the vent location, estimating the dispersal axis is a more robust way to constrain the source vent location for tephra deposits whose transport was affected by wind. If a deposit was affected by wind, the direction along which the deposit has the lowest thinning rate (the downwind direction) should be easier to estimate compared with the vent location.

In the following text, we briefly introduce the methodology and apply the method to three different thickness plus one maximum clast size datasets of North Mono Beds 1 and 2 (NMB1 and NMB2) from the North Mono eruptions (Sieh and Bursik 1986) and the Fogo A tephra deposit (Walker and Croasdale 1971b) to demonstrate its utility. Randomly selected and localized subsets of data are used as input to evaluate its performance in cases of sparse observations. With randomly selected subsets of different sizes, we examine the vent locations and dispersal axes predicted from the method. Using localized subsets, we evaluate how locations of sample sites affect the estimated results. This experiment also serves to test how the method responds when a deposit was affected by a change in wind direction during sedimentation. Bootstrap aggregating (bagging; Breiman 1996) and examining surface of the cost function, together with using localized subsets as input, are proposed to characterize the epistemic uncertainty. Surface of the cost function also provides inferences on how different thinning patterns of tephra fall deposits would affect the performance of the method.

We discuss how different features (i.e., thinning rate, change in dispersal direction, local variation, and multiple lobes) of tephra thickness or maximum clast size distribution affect the performance of the method, and we analyze the benefits and limitations of this approach. Suggestions on how to use the method given dataset of different sizes and qualities are listed. One subset of the NMB1 thickness dataset and thickness datasets of the Trego Hot Springs (THS) and Rockland tephras are applied to the method to test its performance in dealing with sparse measurements. The method is then applied to constrain the vent location and estimate the volume of a series of well-correlated tephra sub-units preserved within the Wilson Creek Formation (WCF) of Mono Lake, CA, USA (Lajoie 1968).

The well-constrained near-field dispersal pattern of single-unit tephra deposit indicates the direction towards which most of the finer ash would travel and therefore provides guidance on where the tephra is likely to be preserved at locations much farther from the vent. This can be used as an additional and independent constraint for the correlation of marker tephras, such as the ones preserved in the Great Basin (Davis 1978; Madsen et al. 2002; Bursik et al. 2014). The method presented here is simple to implement and does not necessarily require additional information about the deposit, which makes it promising to be integrated into quantitative strategies for tephra characterization and correlation (e.g., Bursik and Rogova

2006; Rogova et al. 2007; Pouget et al. 2014b; Rogova et al. 2015; Kawabata et al. 2016).

Data

We build and test the method against thickness datasets of NMB1 and NMB2 (digitized from Sieh and Bursik 1986 and Bursik and Sieh 2013, respectively), the Fogo A tephra deposit (digitized from Walker and Croasdale 1971b), and maximum clast size measurements of NMB1 (Bursik 1993). These deposits and their thickness and maximum clast size measurements are well-studied (e.g., Walker and Croasdale 1971a; Sieh and Bursik 1986; Sparks et al. 1992; Bursik 1993; Engwell et al. 2013), have known vent location, and display distinct features that are typical of tephra thickness or maximum clast size distribution (Yang and Bursik 2016). The method is also applied to tephra deposits with sparse measurements, including the THS and Rockland tephras, and several tephra sub-units within the WCF. Brief description of the deposits and their datasets is given in this section. The observed thickness and maximum clast size at each sample site for NMB1, NMB2, and Fogo A deposit are shown in Fig. 1a–d (see supplementary material for raw data and hand-drawn isopachs). Observed thickness and maximum clast size in log scale are also plotted against distance to the source vent for each sample site of the deposits (Fig. 1e). Key characteristics of the deposits and their associated datasets are summarized in Table 1.

North Mono Beds 1 and 2

NMB1 and NMB2 are two sub-Plinian tephra beds produced from the most recent eruptions from the Mono Craters, eastern central California. The eruptions took place during the fourteenth century A.D. (Sieh and Bursik 1986; Bursik 1993). The rhyolitic deposits were erupted from vents at the northern end of the volcanic chain (Fig. 1a) and hence named North Mono by Sieh and Bursik (1986). The eruption produced eight distinct and widely dispersed air fall beds, as well as pyroclastic flow and surge deposits, and lava domes and flows. Isopach and isopleth maps of them are presented in the studies of Sieh and Bursik (1986) and Bursik (1993), respectively. The datasets used in this work are from Sieh and Bursik (1986) and Bursik and Sieh (2013).

NMB1 is the basal unit of the North Mono tephra and has a total volume of 0.042 km^3. It was erupted from the southwestern portion of North Coulee, and the vent is oriented along the arch of the Mono Craters (Sieh and Bursik 1986; Bursik and Sieh 1989, Fig. 1a inset figure). The deposit is characterized by normal grading and was deposited in south-southwesterly winds (dispersal direction ~ 18° from north clockwise) throughout the course of the eruption, but the proximal material was transported more towards the northeast (~ 40° from north clockwise; Sieh and Bursik 1986; Table 1). A hundred and eighteen thickness and forty-six maximum clast size measurements (non-zero) without signs of reworking were made for this deposit (Fig. 1a, b).

NMB2 is in sharp contact with NMB1. Its vent is the Upper Dome of Northwest Coulee (Fig. 1c). The lower part of the deposit is rich in obsidian and hence presents a darker color. This bed was blown towards the north-northwest (334°; Sieh and Bursik 1986; Bursik 1993). However, distal (> ~ 20 km from the vent) portion of the deposit was dispersed towards the north-northeast (Sieh and Bursik 1986). The isopach map suggests that the dispersal direction for the distal portion is ~ 345°. The deposit is subject to limited or no reworking, and was sampled at seventy-five sites (Sieh and Bursik 1986; Fig. 1c).

Fogo Member A

The Fogo A Plinian eruption on São Miguel, Azores, took place about 5000 years ago (Moore 1990). It has a total volume of 1.2 km^3 on land and originated from Lagoa do Fogo (Walker and Croasdale 1971b). Thickness and maximum pumice and lithic size of this deposit were sampled at 249 sites (Walker and Croasdale 1971b), but these define a lower bound on thickness due to severe erosion. There are 184 valid thickness measurements (Fig. 1d) that can be directly used as input (primary deposit, excluding sample sites with no tephra observed) for the Fogo A deposit. The deposit was further divided into syenite-poor lower (< 20 vol%) and syenite-rich upper divisions (> 20 vol%) by Bursik et al. (1992b) with the lower dispersed predominantly to the south and the upper transported to the east. The presence of two lobes is not evident based on thickness measurements (Walker and Croasdale 1971b; Engwell et al. 2015; Yang and Bursik 2016). The dataset used in this work uses the summed thickness of the two divisions.

THS and Rockland tephras

The THS tephra layer was first described by Davis (1978). It was erupted from Mt. Mazama 23.21 ± 0.30 ka, which can be corrected to ~ 26 ka BP using the Kitagawa and van der Plicht calibration curve (Benson et al. 1997; King et al. 2001; Benson et al. 2003). A GISP2 (Greenland Ice Sheet Project 2) model age has been estimated to be 29.9 ka (Benson et al. 2013). The Rockland tephra (Sarna-Wojcicki et al. 1985), aged 570–610 ka (Sarna-Wojcicki et al. 1985; Lanphere et al. 1999, 2004), was produced from the Rockland caldera complex now infilled by Brokeoff Volcano, a part of the Lassen volcanic complex (Clynne 1984; Lanphere et al. 1999; Clynne and Muffler 2010). Wide dispersal of these two tephra deposits makes them critical markers for climate

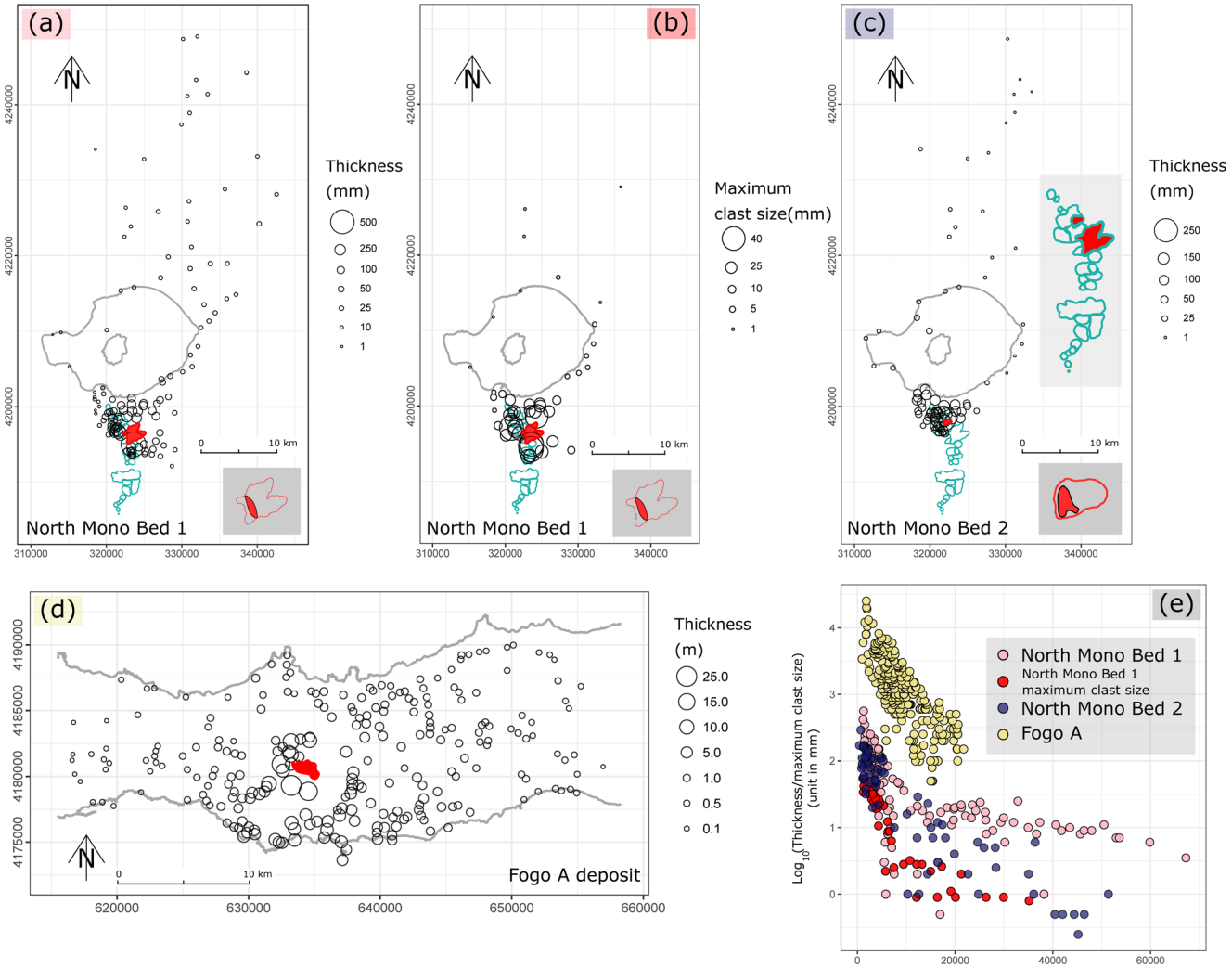


Fig. 1 **a**, **b** NMB1 thickness and maximum clast size measurements. **c** NMB2 thickness measurements. **d** Fogo A thickness measurements. Corresponding source vents for these deposits are marked as red polygons in **a**–**d**. In the lower-right corner in **a**–**c**, the vent location is marked with a greater resolution. The contour of the Mono Craters is highlighted in light green polygons in **a**–**c** and is amplified in **c**. Raw data and hand-drawn isopach and isopleth maps can be found in supplementary material. The data are digitized from Walker and Croasdale (1971b) for the Fogo A deposit and from Sieh and Bursik (1986) and Bursik (1993) for all North Mono Beds. **e** Observed thickness (mm) and maximum clast size (mm) under log-scale are plotted against distance (m) to the source vent for each measurement and for each dataset

and stratigraphic reconstructions of the geologic history of the western USA (Davis 1978; Sarna-Wojcicki et al. 1985, 1991; Alloway et al. 1992; Benson et al. 1997; Lanphere et al. 1999; Sarna-Wojcicki 2000; Benson et al. 2003; Adams 2010). Estimated volumes of the THS and Rockland tephras are 15–56 km^3 (Pouget et al. 2014a) and 50–248 km^3 (Sarna-Wojcicki et al. 1985; Pouget et al. 2014b), respectively. In this work, we use the thickness datasets compiled by Pouget et al. (2014b) from previous studies as input (Sarna-Wojcicki et al. 1985, 1987; Negrini et al. 1988; Alloway et al. 1992; Rieck et al., 1992; Benson et al. 1997; King et al. 2001; Bowers 2009; Adams 2010; Benson et al. 2013). Each dataset contains eight observations, and two measurements for the Rockland tephra may have been reworked (Pouget et al. 2014a).

Tephra sub-units in the Wilson Creek Formation

The WCF is composed of lacustrine deposits formed in Pleistocene Lake Russell (the present-day Mono Lake), interbedded with tephra layers (Lajoie, 1968), which were mostly produced from the Mono Craters. Little is known about their vent location and total volume (Lajoie 1968; Bursik and Sieh 1989; Marcaida et al. 2014). Among them, sub-unit correlation of ashes B7, A4, A3, and A2 (from Black Point volcano) has been established based on outcrops sampled near Mono Lake (Yang et al. 2019). The sub-unit A4-d has been observed at six sites, and its stratigraphic features suggest that it was produced from the northern part of the Mono Craters (Yang et al. 2019).

Table 1 Summary of the NMB1, NMB2, Fogo A deposits, and the corresponding datasets

Tephra bed	North Mono Bed 2(NMB2; sub-Plinian)		Fogo A (Plinian)	North Mono Bed 1 (NMB1; sub-Plinian)
Data type	Thickness	Maximum clast size	Thickness	Thickness
Number of measurements used	118	46	75	184
Thinning rate	Rapid change in thinning rate (Fig. 1e)	Rapid change in fining rate (Fig. 1e)	Relatively stable thinning rate	Relatively stable thinning rate
Number of layer/sub-layers within the deposit	One	One	One	Two
Source vent	North Coulee	North Coulee	Upper Dome	Lagoa do Fogo Crater Lake
Source vent geometry (km×km)	2.6 × 1.9	2.6 × 1.9	1.0 × 0.8	2.1 × 1.0
Effect by wind	Strongly affected by wind	Strongly affected by wind	Strongly affected by wind	Slightly affected by wind
Dispersal direction (° from north clockwise)	40 (< 6 km from the source) 18 (> 6 km from the source)	40 (< 6 km from the source) 18 (> 6 km from the source)	334 (< 20 km from the source) ~ 345 (> 20 km from the source)	Southwards (lower sub-layer) Eastwards (upper sub-layer)
Note	–	–	–	Severe local thickness variation
Semi-empirical model used for different experiments (e.g., $p + e$ if both power-law and exponential models are used)	Method illustration ($p + e$) Results from random subsetting Vent location (p) Dispersal axis ($p + e$, visualized) Results from localized subsets ($p + e$) – Comparison between semi-empirical models ($p + e$)	Method illustration ($p + e$) Results from random subsetting Vent location (p) Dispersal axis ($p + e$, not visualized) Results from localized subsets ($p + e$) – –	Method illustration ($p + e$) Results from random subsetting Vent location ($p + e$) Dispersal axis ($p + e$, visualized) Results from localized subsets ($p + e$) – –	Method illustration ($p + e$) Results from random subsetting Vent location (p) – – Statistical bagging (p) –

The information takes reference from Sieh and Bursik (1986) for the North Mono tephra beds and from Walker and Croasdale (1971b) for the Fogo A deposit. The semi-empirical model(s) used for each experiment is also listed

Ash A4 is possibly correlative with the Lowder Creek ash found in the Markagunt Plateau, Utah (Madsen et al. 2002). Sub-units B7-a and A3-f have been correlated at four to five sites. Since the method allows for the drawing of isopachs in an objective way (including estimating the wind direction objectively, given a potential vent location), it is applied to these datasets to estimate their volumes.

Method

The main idea upon which the method is built is a gradient descent method applied to a semi-empirical model of tephra thickness distribution. The semi-empirical models and gradient descent method are introduced separately in this section. The algorithm is programmed in the R language (R Core Team 2017), and the code is made public on Vhub (https://vhub.org/resources/4377; Yang et al. 2018). Workflow of the method is shown in Fig. 2.

Semi-empirical models of tephra thickness distribution

The semi-empirical models proposed by Gonzalez-Mellado and Cruz-Reyna (2010) and Yang and Bursik (2016) are used here to characterize thickness and maximum clast size distribution of tephra deposits. In this section, we only refer to thickness in describing the method to avoid redundancy.

Given a known vent location (x_s, y_s) and downwind direction (φ), the two semi-empirical models describe the tephra thickness distribution (t_i) at any coordinates (x_i, y_i) as (Gonzalez-Mellado and Cruz-Reyna 2010; Yang and Bursik 2016):

$$t_i(r_i, \theta_i | x_s, y_s, \varphi) = \gamma e^{\{-\beta U r_i [1-\cos(\theta_i - \varphi)]\}} r_i^{-\alpha} \quad (1)$$

and

$$t_i(r_i, \theta_i | x_s, y_s, \varphi) = e^{[\beta_z + \beta_{dd} r_i \cos(\theta_i - \varphi) + \beta_r r_i]} \quad (2)$$

where (r_i, θ_i) is the polar coordinates with respect to the source vent (x_s, y_s) for the arbitrary point (x_i, y_i). Other parameters in

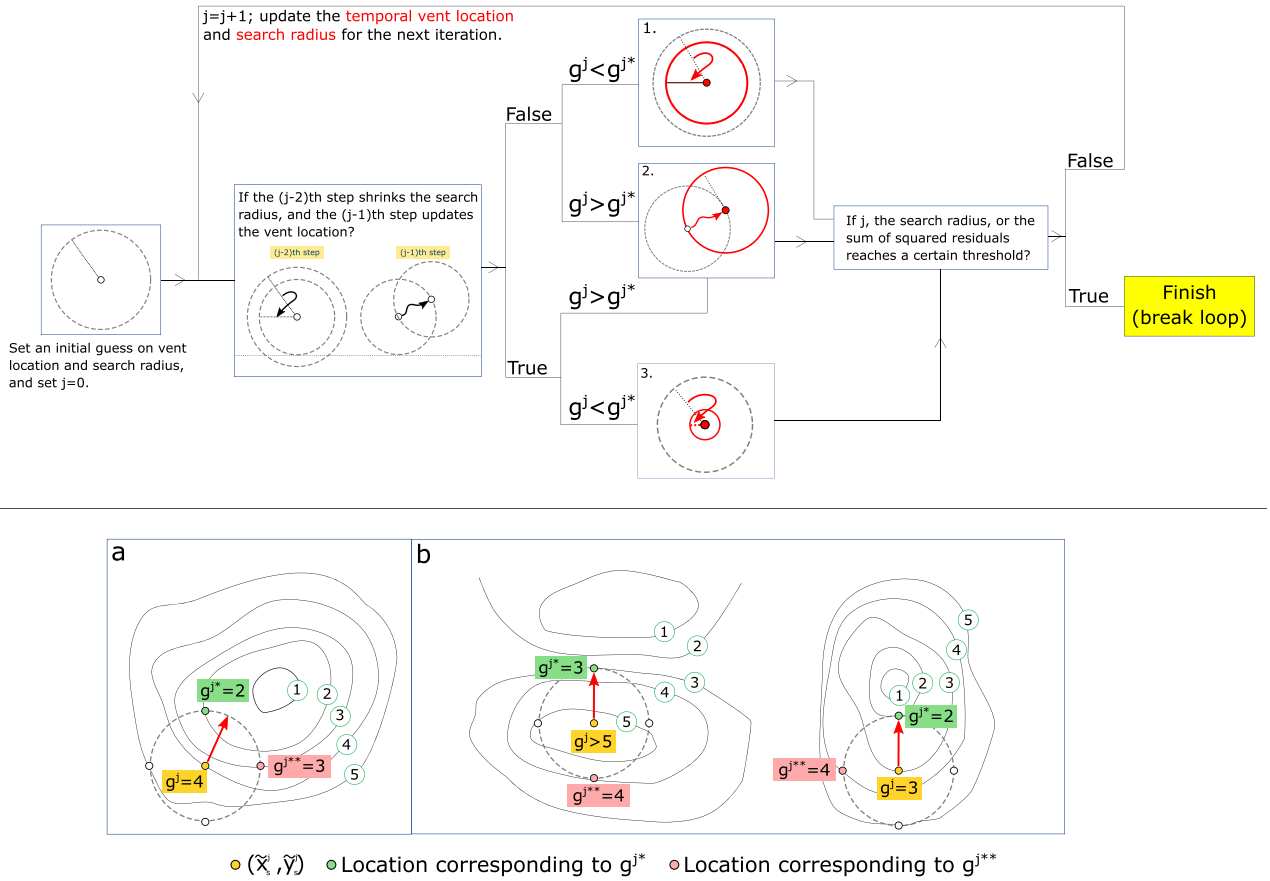


Fig. 2 Workflow of the method (upper portion). The numbers (1, 2, and 3) marked in boxes correspond to method description in text. Lower portion of the figure shows how the method moves in the *x*-*y* plane given case 2. a and b correspond to the cases listed in text. Black lines denote the isopachs of the cost function (value marked in green circle) in the *x*-*y* plane. The figure briefly shows the direction (red arrow) towards which the next iteration would move based on values of g^j, g^{j*}, and g^{j**}

the two equations (i.e., γ, β, U, α in Eq. 1 and β_r, β_{dd}, and β_z in Eq. 2) denote how tephra thickness changes with distance and azimuth with respect to the source vent: γ and β_z are used to characterize the thickness of the tephra fall deposit near the vent; α and β_r describe how tephra thickness decays with distance to the source vent; and βU (combined to a single parameter) and β_{dd} determine how much the tephra fall deposit is stretched along the downwind direction and shortened in the upwind direction due to the advection of wind. See Gonzalez-Mellado and Cruz-Reyna (2010) and Yang and Bursik (2016) for more details on the interpretation of these parameters.

The major distinction between the two semi-empirical models is that, given a fixed wind direction (φ) with respect to the source (also fixed θ_i), the thickness in Eq. 1 changes with both $e^{\{-\beta U r_i[1-\cos(\theta_i-\varphi)]\}}$ and $r_i^{-\alpha}$, but only decays with $e^{[\beta_{dd} r_i \cos(\theta_i-\varphi)+\beta_r r_i]}$ in Eq. 2. That is to say, Eq. 1 is the product of exponential and power-law functions of r_i, whereas Eq. 2 only thins exponentially. The thinning rate of Eq. 1 using the log scale is not constant (with fixed θ_i): it has a relatively greater thinning rate closer to the vent and thins more slowly for a distal portion of a deposit. Eq. 2 has a constant thinning rate using the log scale (with fixed θ_i). For simplicity, we call them the power-law (note the term $r_i^{-\alpha}$ in Eq. 1) and exponential models, respectively, in this study.

Gradient method

The semi-empirical models show that to estimate the vent location, another four variables, namely, wind direction φ and coefficients in Eqs. 1 and 2, need to be estimated. We construct the cost function as the sum-of-squared residuals under log scale:

$$f(x_s, y_s, \varphi) = \sum_{i=1}^{n} \left[\ln\left(\tau_i\right) - \ln(t_i)\right]^2 \quad (3)$$

where τ_i denotes the ith prediction from the semi-empirical model, and n is the number of observations. The occurrence of wind direction, an angular variable, in Eqs. 1 and 2 indicates that the problem cannot be solved directly through pseudo-inverse. However, for a given vent position (x_s, y_s) and wind direction φ, a linear system can be constructed for each semi-empirical model, and $t_i(r_i, \theta_i)$ can be written in matrix form (Gonzalez-Mellado and Cruz-Reyna 2010; Yang and Bursik 2016):

$$A_P b_p = T \quad (4)$$

and

$$A_e b_e = T \quad (5)$$

where T is a column vector with logarithmic thickness at sites (x_i, y_i), $i = 1, 2, \ldots, n$, $b_p = (\ln(\gamma), -\beta U, -\alpha)^T$, $b_e = (\beta_z, \beta_{dd}, \beta_r)^T$. A_P and A_e are n-by-3 matrices with the ith row being $(1, r_i[1-\cos(\theta_i-\varphi)], \ln(r_i))$ and $(1, r_i\cos(\theta_i-\varphi), r_i)$, respectively.

With known vent location and wind direction, other coefficients in Eqs. 1 and 2 (equivalently, b_p and b_e in Eqs. 4 and 5) can be determined by solving a standard least square problem through pseudo-inverse. Estimating the vent location is thus simplified as finding the combination of vent location and wind direction that minimizes Eq. 3. This is done by a gradient descent method, which is divided into two steps.

First, a standard one-dimensional gradient descent method is constructed to find the optimum wind direction with a fixed vent location $(\text{𝓍}_s, \text{𝓎}_s)$, which minimizes $f(x_s, y_s, \varphi | x_s = \text{𝓍}_s, y_s = \text{𝓎}_s)$. This one-dimensional gradient descent method can be regarded as a function $s(\text{𝓍}_s, \text{𝓎}_s)$ whose output is the optimum wind direction φ_{opt}.

Another two-dimensional gradient descent method is adopted to minimize:

$$g\left(\text{𝓍}_s, \text{𝓎}_s\right) = f\left(x_s, y_s, \varphi | x_s = \text{𝓍}_s, y_s = \text{𝓎}_s, \varphi = s\left(\text{𝓍}_s, \text{𝓎}_s\right)\right) \quad (6)$$

The gradient descent method finds the optimum vent location by moving iteratively in the x-y plane towards the direction that leads to a smaller value of $g(\text{𝓍}_s, \text{𝓎}_s)$. Note that the wind direction φ is updated for each iteration (Eq. 6). Starting from an initial guess of the vent position $(\text{𝓍}_s^j, \text{𝓎}_s^j)$, and a specified search radius h^j with $j = 0$, the method compares the value of $g^j = g\left(\text{𝓍}_s^j, \text{𝓎}_s^j\right)$ and the value of g at locations near $\left(\text{𝓍}_s^j, \text{𝓎}_s^j\right)$. These locations are $(\text{𝓍}_s^j \pm h^j, \text{𝓎}_s^j \pm h^j)$, where h^j can be regarded as search radius. By defining $g^{j*} = \min\left(g\left(\text{𝓍}_s^j \pm h^j, \text{𝓎}_s^j \pm h^j\right)\right)$ and comparing it with g^j, the method determines what to do for the next iteration:

1. If $g^j < g^{j*}$, then the true vent position is closer to $\left(\text{𝓍}_s^j, \text{𝓎}_s^j\right)$, and therefore, we keep $\left(\text{𝓍}_s^{j+1}, \text{𝓎}_s^{j+1}\right) = \left(\text{𝓍}_s^j, \text{𝓎}_s^j\right)$ and shrink the search radius $h^{j+1} = 0.7h^j$. The value 0.7 can be changed to any value from 0.5 to 1. The range has been chosen such that at the initial stage of the iteration, h^j would not decrease too rapidly. This allows the method to move $\left(\text{𝓍}_s^j, \text{𝓎}_s^j\right)$ to the area close to the global or local minimum with relatively fewer iterations even if the search radius has been shrunk once or twice at the very beginning of the iteration.
2. If $g^j > g^{j*}$, then the vent position is closer to one of $\left(\text{𝓍}_s^j \pm h^j, \text{𝓎}_s^j \pm h^j\right)$, rather than $\left(\text{𝓍}_s^j, \text{𝓎}_s^j\right)$. That means a new location provides a better fit between observations and model predictions. Thus, we need to update the vent position, which is described in the following text.
3. For the jth iteration ($j > 1$), if (i) the $(j-2)$th iteration did not update the vent to a new location (i.e., $\left(\text{𝓍}_s^{j-2}, \text{𝓎}_s^{j-2}\right)$

$= \left(x_s^{j-1}, y_s^{j-1}\right)$); (ii) the $(j-1)$th iteration updates the vent position; and (iii) $g^j < g^{j*}$, the search radius for the next iteration is updated as $h^{j+1} = (1-0.7)h^j$, instead of $h^{j+1} = 0.7h^j$, and we keep $\left(x_s^{j+1}, y_s^{j+1}\right) = \left(x_s^j, y_s^j\right)$ for the next iteration. This is because the $(j-2)$th and $(j-1)$th steps imply that location $\left(x_s^j, y_s^j\right)$ is close to the true vent location, and it is thus justified to shrink the search radius with a greater rate to improve the efficiency of the method (there is no need to decrease the search radius gradually).

This setup allows the method to be in low resolution but more efficient during the initial iterations. As $\left(x_s^j, y_s^j\right)$ approximates the true vent, the length of the search radius shrinks faster, and the decreased search radius helps improve the resolution of the method during the last iterations. The iteration will be terminated as it reaches certain thresholds on prespecified search radius, number of iterations, or value of the cost function (workflow shown in upper portion of Fig. 2).

The location for the next iteration $\left(x_s^{j+1}, y_s^{j+1}\right)$ is updated based on the difference between g^j and the two smallest values among $g\left(x_s^j \pm h^j, y_s^j \pm h^j\right)$, defined as g^{j*} and g^{j**}:

(i) If $g^j > g^{j*}$ and $g^j > g^{j**}$ and g^{j*} and g^{j**} are produced due to the change in both x and y-directions:

The quadrant which $\left(x_s^{j+1}, y_s^{j+1}\right)$ will be in with respect to $\left(x_s^j, y_s^j\right)$ is determined. For example, if $g^{j*} = g\left(x_s^j + h^j, y_s^j\right)$ and $g^{j**} = g\left(x_s^j, y_s^j + h^j\right)$, then the next step will move to the first quadrant with respect to $\left(x_s^j, y_s^j\right)$. The direction of the shift vector is also determined by $\Delta g^{j*} = g^{j*} - g^j$ and $\Delta g^{j**} = g^{j**} - g^j$. Following that $\left(x_s^{j+1}, y_s^{j+1}\right)$ will go to the first quadrant, the next iteration is $\left(x_s^{j+1}, y_s^{j+1}\right) = \left(x_s^j + |\Delta x^j|,\ y_s^j + |\Delta y^j|\right)$, and $(|\Delta x^j|, |\Delta y^j|)$ can be written as $\left(\frac{-h^j \Delta g^{j*}}{\sqrt{(\Delta g^{j*})^2 + (\Delta g^{j**})^2}}, \frac{-h^j \Delta g^{j**}}{\sqrt{(\Delta g^{j*})^2 + (\Delta g^{j**})^2}}\right)$.

(ii) If $g^j > g^{j*}$ and $g^j > g^{j**}$ and they are produced due to the change only in the x- or y-direction (ridge), or, $g^j > g^{j*}$ and $g^j < g^{j**}$:

We move the point towards the direction that is consistent with g^{j*} and fix the other coordinate. For example, if $g^{j*} = g\left(x_s^j + h^j, y_s^j\right)$, the next iteration will move towards the positive direction of the x-axis, and the absolute value of the shift vector for the next step is $(|\Delta x^j|,\ |\Delta y^j|) = (h^j, 0)$.

Note that an explicit formulation of the above arguments is not given, because the method considers the comparison between $g\left(x_s^j, y_s^j\right)$ and $g\left(x_s^j \pm h^j, y_s^j \pm h^j\right)$ instead of approximating the gradient of the cost function. This approach is intentionally designed to reduce the potential of a local minimum solution. The gradient descent method described above is independent of the semi-empirical model and could be applied flexibly to other functional forms.

The output of the method includes the predicted vent coordinates and dispersal direction and the search radius (h) and fitted coefficients (γ, α, βU for the power-law model, and β_{dd}, β_r, β_z for the exponential model) at the last iteration. Sum-of-squared residuals and coefficient of determination from the fitting (conditioned on the predicted vent location and dispersal direction) are also calculated.

Use of the gradient method can introduce local minima, as well as unphysical and non-convergent predictions. These occur naturally in different kinds of optimization problems and are not exclusive to one another. Local minima can be detected by starting from different initial locations and assessing their convergence towards a similar or dissimilar location (indicating a local minimum). To reduce the possibility of their occurrence, it is necessary to specify the starting point based on prior knowledge about the deposit. The starting point should be within an area that includes all potential vents or that is close to the sample site with the greatest thickness or maximum clast size measurement. With limited observations, it is suggested to plot the surface of the cost function (Eq. 6) to examine local minima explicitly.

Unphysical predictions, namely, distributions that thicken or coarsen with distance, might occur, given sparse observations. They can be recognized by examining values of the fitted coefficients: the results are unphysical if the predicted $\alpha < 0$ for the power-law model (Eq. 1), or the predicted $(\beta_{dd} + \beta_r) > 0$ for the exponential model (Eq. 2). The occurrence of unphysical results suggests that the current measurements cannot be well described by the semi-empirical model being used. In our experience, unphysical predictions occur rarely.

By specifying a sufficiently large number of iterations for the method, it will lead to (1) convergence at local or global minimum or (2) lack of decrease of the search radius (divergence). If the latter occurs, the method could continue to "walk" in the x-y plane at the initial search radius. After a sufficiently large number of iterations, the predicted vent location could be at a point that is far from the potential source region and sampled area. Therefore, if the latter occurs, it is certain that the prediction is incorrect.

Here the criterion for "sufficiently large" is based on the size of the potential vent area and length of the initial specified search radius. Consider the worst case scenario wherein the initial guess and true vent location are two points that have the greatest distance from each other within the potential source region. The longest path possible to reach the true vent can be assumed to be the semicircle centered at the midpoint between the initial guess and true vent location. The iteration only migrates in 2D space and the semi-empirical models have simple functional forms, which make the surface of the cost function relatively simple, and the likelihood for a path that is longer than the semicircle is extremely small. Here we assume

that the iteration would not go away from the direction to the true vent. If the distance between the initial guess and true vent location is L, the longest path possible is $\frac{\pi L}{2}$. If the initial search radius is h_0, the rounding of $\frac{\pi L}{2h_0}$ plus one iterations would allow $\left(x_s^j, y_s^j\right)$ to approach the global or local minimum (distance to the true vent below h_0; if the search radius has shrunk) or lead to divergent results (the search radius remains unchanged). Then, the method should start to shrink the search radius if the result is not divergent. Again, consider the worst case scenario that the search radius shrinks at the lowest rate (0.7). To reach the resolution of $0.02 \times h_0$ (the prediction is within $0.02 \times h_0$ from the true vent location) takes ~ 11 iterations ($0.7^{11} \approx 0.198$). Therefore, the sufficiently large number of iterations could be $\frac{\pi L}{2l}$ plus 12. In the following experiments, the number of iterations for each run is specified as "sufficiently large" in the present sense.

Results

In this section, we present results from applying the method to datasets introduced above. First, we confirm that the approach determines accurate source locations given abundant data. More subsets with fewer input points are applied to the method to evaluate its performance given sparse observations. Since the forward models used for the method, namely, the power-law and exponential models, are simple, epistemic uncertainty cannot be neglected. Our key concerns are how the method reacts or if the method is able to detect (1) change in wind direction, (2) the presence of multiple lobes, and (3) severe local variation for a tephra deposit. The first concern is tested by applying localized subsets of the NMB1 and NMB2 datasets to the method, and the last two are addressed by applying statistical bagging with the method to the Fogo A dataset. Characteristics of these deposits match the concerns listed above (Table 1). Surfaces of the cost function for all tested datasets are plotted and visualized as another way of using the method, and the complete datasets are used as input such that systematic and intrinsic sources of uncertainty can be observed. The two semi-empirical models are considered equally important, but our results show that the performance of the method is greatly affected by the use of different semi-empirical models. It is thus not necessary to use both semi-empirical models all the time. Which semi-empirical model is applied to the below experiments is highlighted in Table 1, and the reasoning is also pointed out in text (based on results shown in "Method illustration").

The ideal initial condition for this approach should reflect our prior knowledge on the specific tephra deposit or the corresponding dataset. For NMB1 and NMB2, the prior knowledge is assumed to be that they were erupted from somewhere within the Mono Craters (Fig. 1). For the Fogo A deposit, the starting point is a random location within 250 m from the sample site with the greatest measurement (this sample site changes in each run as the input dataset is different).

The number of iterations is set to be 40 for all runs, and the initial search radius is set to be $h^0 = 1000$ m. This means that if the method fails to converge, the corresponding prediction will be ~ 40,000 m (40 × 1000 m) from the starting location. This is sufficient given the size of the potential vent area (Fig. 1). The randomness in specifying the starting point can be neglected because (1) the initial search radius is set to be 1000 m for all cases, greater than the randomness from specifying the starting point, and (2) the results are always presented as a series of predictions from using different subsets of data as input instead of a single prediction. Local minima could be detected if the predictions do not converge to the same point (each time the initial location is different, if local minima exist, they will be detected). In the following results, unphysical predictions occur rarely and are pointed out if present. Non-convergent results are not highlighted because it is certain that they are incorrect.

Method illustration

A different number of data points (NMB1 and NMB2 60, 40, and 20; Fogo A deposit 150, 100, 50, 30) are randomly drawn from the complete datasets for 25 times each and applied to the method. The same procedure was done for the maximum clast size data of NMB1 except with 30 input points drawn from the complete dataset for 10 times due to fewer measurements. Since this section is aimed at demonstrating the utility of the method, these subjective choices would not affect the motif herein. To avoid the uncertainty being amplified by data scarcity, it is preferred to have relatively more samples as input here.

For NMB1 (Fig. 3) with the power-law model, there are 25, 22, and 21 predictions that are within 2 km from the true vent center using 80, 40 and, 20 input points, respectively. Predictions with the exponential model form a linear pattern lying along the upwind portion of the dispersal axis. Predicted source vents derived from the maximum clast size data behave in similar fashions. When the power-law model is used, all ten predictions are located at the true vent location, whereas seven out of ten predictions (the other three are outside the extent of Fig. 3c) from the exponential model are again located in the upwind portion of the dispersal axis. The predicted wind directions using the NMB1 thickness data (Fig. 3d, e) are consistent with the general dispersal pattern of the deposit, and difference between results from the two models is negligible. Wind directions predicted from the method have two modes when the maximum clast size data is used (Fig. 3f). The two modes correspond to the dispersal direction for the proximal (northeast) and distal (north-northeast) portions of the deposit. This feature is more distinct when the maximum clast size

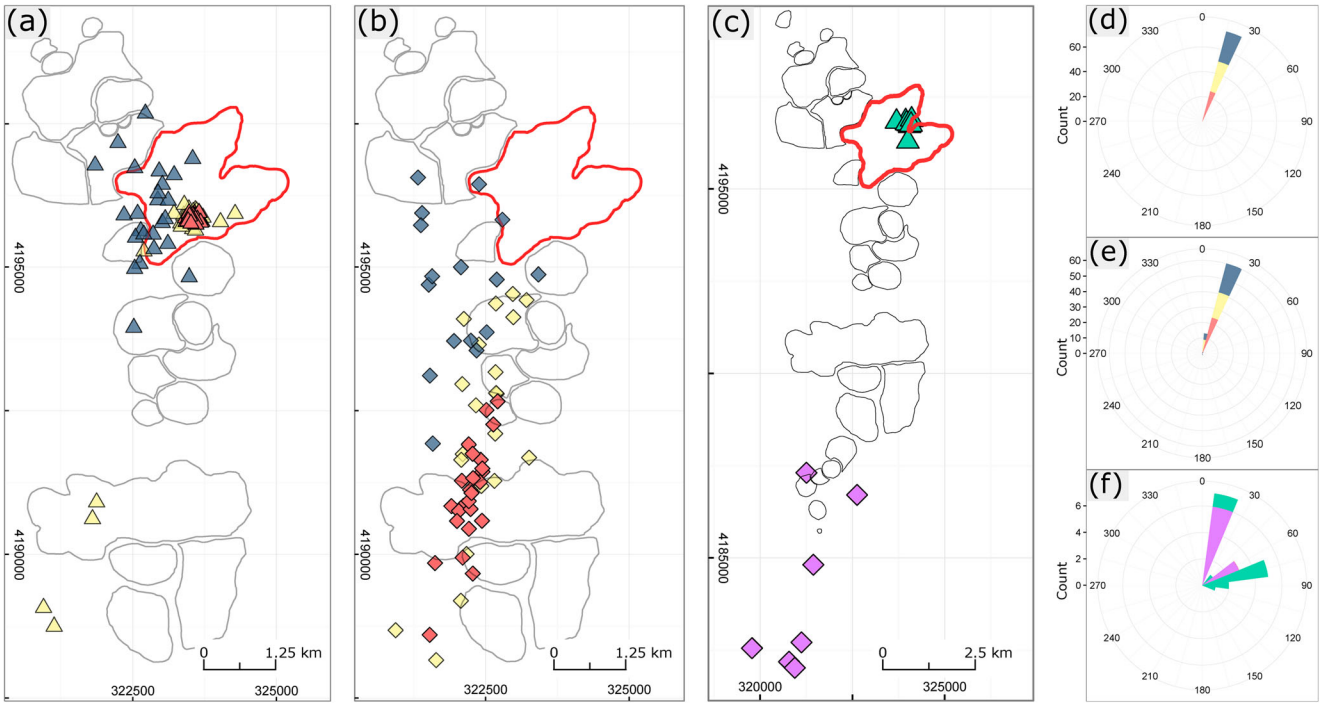


Fig. 3 **a**, **b** Estimated vent locations of NMB1 from using the power-law (triangle) and exponential (diamond) models, respectively, with 80 (red), 40 (yellow), and 20 (blue) input points randomly drawn from the original datasets. Shown are ensembles of 25 predictions for each dataset size. **c** Predicted vent positions of NMB1 from using the maximum clast size data with the power-law (light green triangle) and exponential (purple diamonds) models, respectively, with 30 input points randomly drawn for ten times from the original dataset. **d**–**f** Predicted dispersal directions corresponding to **a**–**c** with consistent color

dataset is used because more measurements were made at locations closer to the source vent.

Predicted vent locations for NMB2 (Fig. 4) using the power-law model are concentrated at the eastern half of the true vent. With 60, 40, and 20 input points, there are 25, 18, and 14 predictions within 1 km (the size of the vent is smaller compared with NMB1) from the center of the vent, respectively. For the exponential model, the corresponding counts

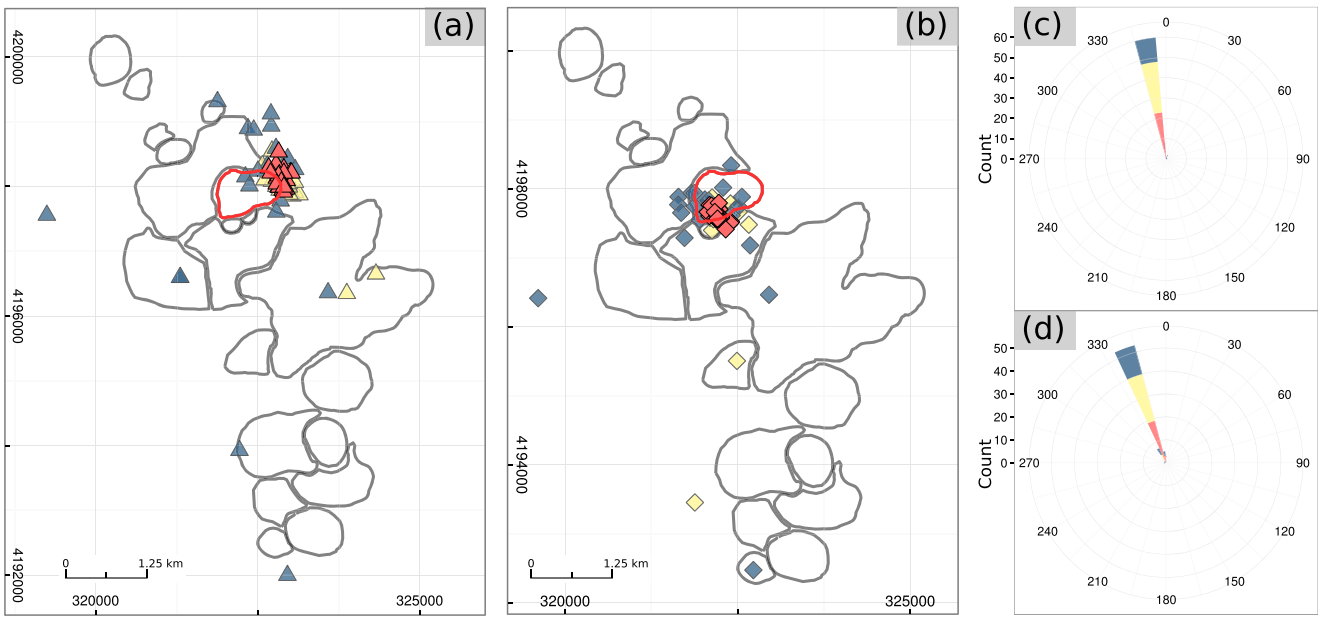


Fig. 4 **a**, **b** Estimated vent positions of NMB2 from using the power-law (triangle) and exponential (diamond) models, respectively, with 60 (red), 40 (yellow), and 20 (blue) input points randomly drawn from the original thickness dataset. Shown are ensembles of 25 predictions for each dataset size. **c**, **d** Predicted wind directions corresponding to **a** and **b** with consistent color

are 25, 23, and 18. Predictions from the exponential model are concentrated at the western half of the Upper Dome, in agreement with the published map of Sieh and Bursik (1986). Predicted wind directions (Fig. 4c, d) between the two models are consistent with the main dispersal direction of NMB2 despite a slight difference of 15°. The predicted wind directions calculated from the power-law and exponential models correspond to the dispersal patterns of the proximal and distal portions of the deposit (Table 1).

For the Fogo A deposit, using the power-law model, most of the predictions are at the southern rim of the Lagoa do Fogo caldera lake (Fig. 5a). This is partly due to digitization error. For example, Point a in Fig. 5a is a sample site located on the island north of its current marked position. With the power-law model, there are 25, 25, 18, and 14 predictions that are within 2 km from the center of the vent using 150, 100, 50, and 30 input points, respectively. With the exponential model, the corresponding counts are 24, 17, 16, and 12, less accurate compared with the power-law model. Predicted vent locations are more scattered in the latitudinal direction when the exponential model is used. In terms of predictions in the dispersal direction (Fig. 5c, d), the two semi-empirical models give similar results, pointing towards southeast, but the variance is greater when the exponential model is used. The wind directions shown in Fig. 5c, d are similar to the overall dispersal direction of the deposit inferred from previous studies (Walker and Croasdale 1971b; Engwell et al. 2015).

The above results show that the method can be used to estimate the vent location and dispersal axis of tephra fall deposits. The performance of the method varies with different datasets and is also dependent on the use of different semi-empirical models. For estimating the source vent location, the power-law model is more accurate for NMB1 and the Fogo A deposit, but the exponential model works better for NMB2.

Results from random subsetting

Knowing that the method is able to estimate the vent location of tephra deposits, we need to evaluate its performance given limited data. More experiments are performed in a similar way, but with fewer input data points and more runs. We obtain a thousand randomly selected subsets with 30, 20, 15, and 10 input points, respectively, for each dataset and apply them to the method. The predicted vent locations and dispersal axes are examined in this section. Previous experiments have shown that the power-law model is better than the exponential model at predicting the vent location for the NMB1 and Fogo deposit. Predicted vent locations from using the exponential model for these deposits are not shown, but the corresponding dispersal axes will be examined for NMB1.

Predicted vent locations

Predicted vent locations with different semi-empirical models for each dataset using 30 and 10 input points (as end members) are shown in Fig. 6. With 30 input points, results for the NMB1 (using the power-law model) and NMB2 (using exponential model) thickness datasets are clustered around the true

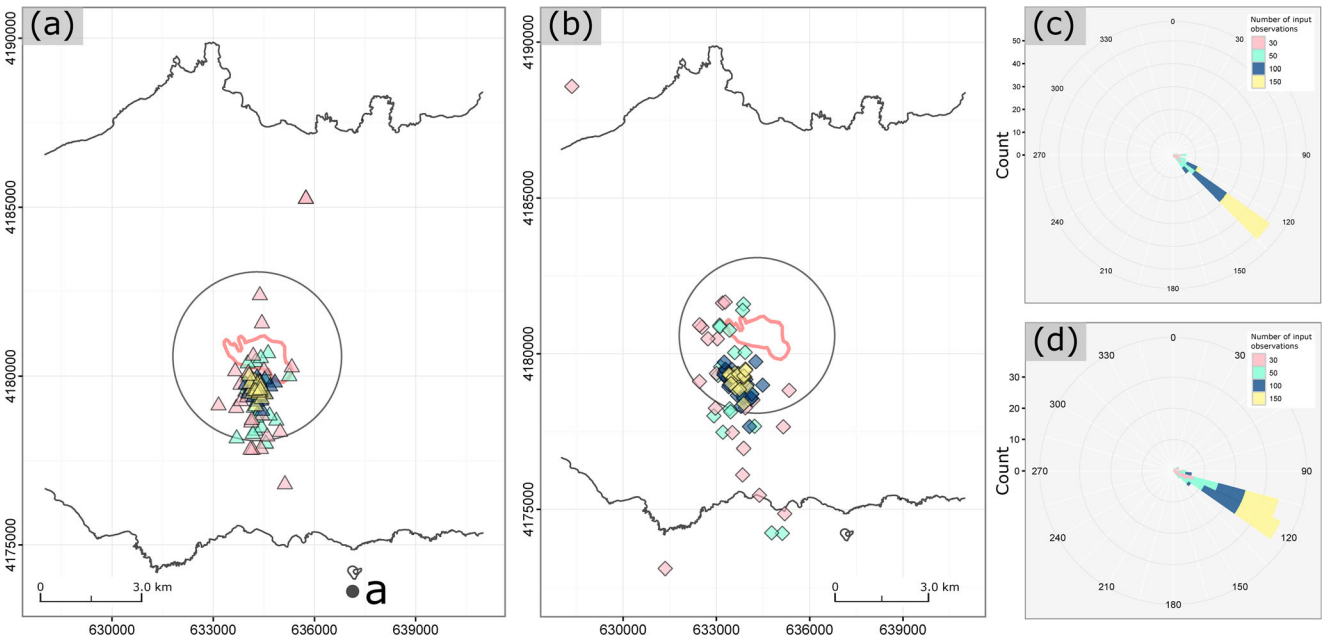


Fig. 5 **a**, **b** Estimated vent locations of the Fogo A deposit from using the power-law (triangle) and exponential (diamond) models, respectively, with 150 (yellow), 100 (blue), 50 (light green), and 30 (light pink) input points randomly drawn from the original dataset. Shown are ensembles of 25 predictions for each dataset size. **c**, **d** Predicted wind directions corresponding to **a** and **b** with consistent color. The southernmost sample site, Point a, is marked to denote digitization error. It is sampled on the small island in the north of it

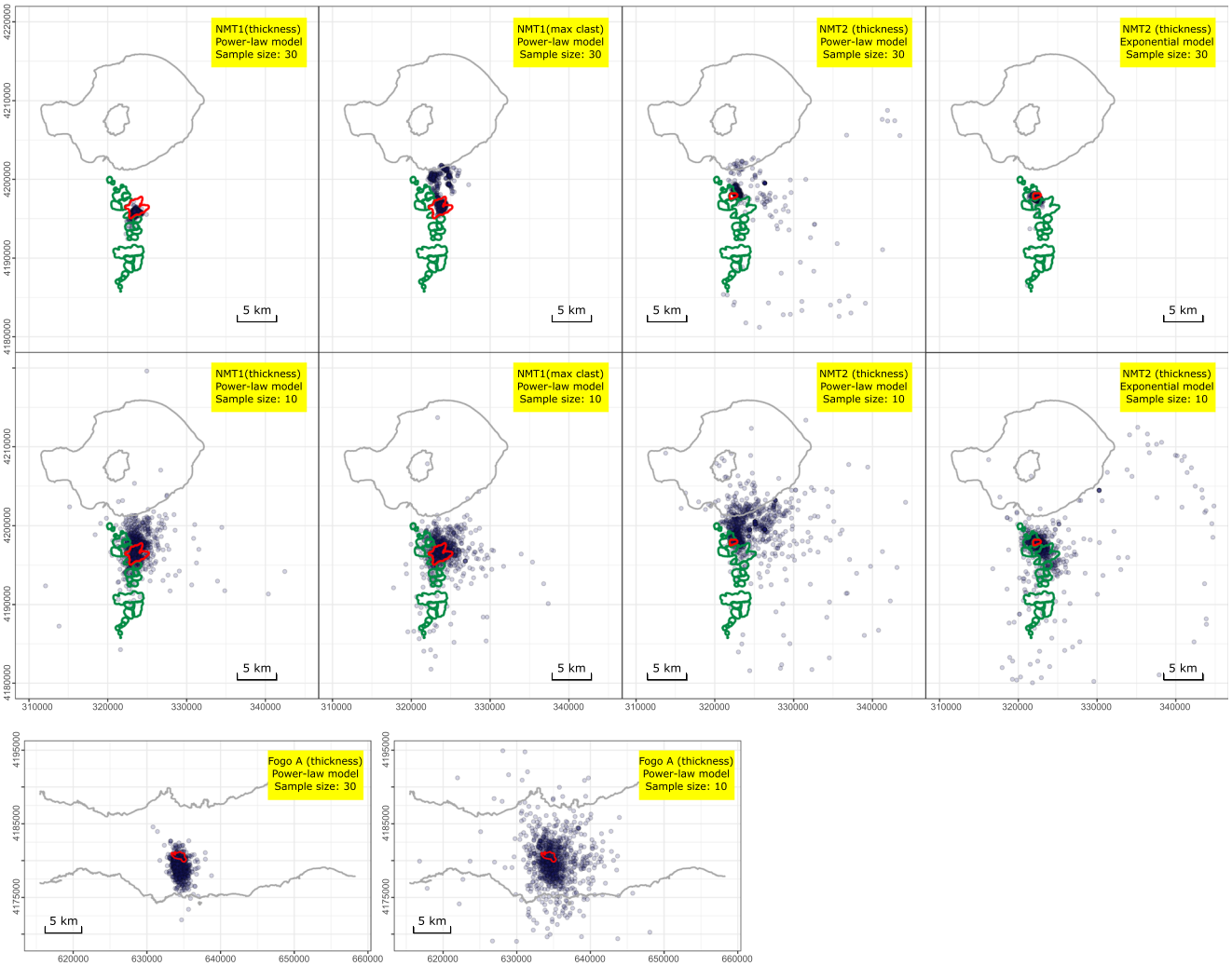


Fig. 6 Predicted vent locations (blue points) for NMB1 (thickness and maximum clast size), NMB2, and the Fogo A deposit. The results are derived from applying randomly selected subsets containing 30 and 10 sample points for 1000 times. The dataset and semi-empirical model being used and the size of input are highlighted in the upper-right corner of each figure

vent area. Results from the NMB1 maximum clast size and NMB2 thickness datasets using the power-law model are more scattered. For the Fogo A dataset, similar to the results shown in Fig. 5, the predictions are clustered at the southern part of the vent area with a greater variability in the *y*-direction. This is possibly due to digitization error and severe local thickness variation. With 10 input points, the variability of predicted vent coordinates increases for all cases. Distance from each predicted vent to the center of the true vent is calculated and plotted as histogram for each case in Fig. 7. Unphysical results (marked in gray) occur rarely, and the corresponding predictions are far from the center of the true vent.

We use the ratio of predicted vents that are within 2 km from the center of the true vent to denote accuracy of the method (unphysical predictions are excluded). Given sizes of the corresponding vents (Table 1), this length is considered to be sufficient. Results for each case are summarized in Fig. 9a–d and Table 2. The ratio of predictions that are within 1 and 3 km from the center of the true vent is also summarized in Fig. 9a–d (1, 2, and 3 km are also marked in Fig. 7).

With 30 input points, accuracies are 99.4% and 65.7% using the NMB1 thickness (power-law model) and maximum clast size (power-law model) data, 84.6% and 99.4% for NMB2 with the power-law and exponential models, and 88.3% for the Fogo A deposit (power-law model), respectively. Accuracies decrease with the input dataset size. This trend (Fig. 9a, d) is consistent and stable for the NMB1 (thickness) and the Fogo A datasets, in which cases the accuracies decrease to 84.1% (15 input points) and 56.6% (10 input points) for the former and 61.8% (15 input points) and 38.4% (10 input points) for the latter. The accuracy is not strongly affected by the input dataset size for the NMB1 maximum clast size dataset, which ranges from 52.4 to 65.4%. This is because NMB1 has different dispersal directions for the proximal and distal portions of the deposit. This is more distinct for the NMB1 maximum clast size data because more

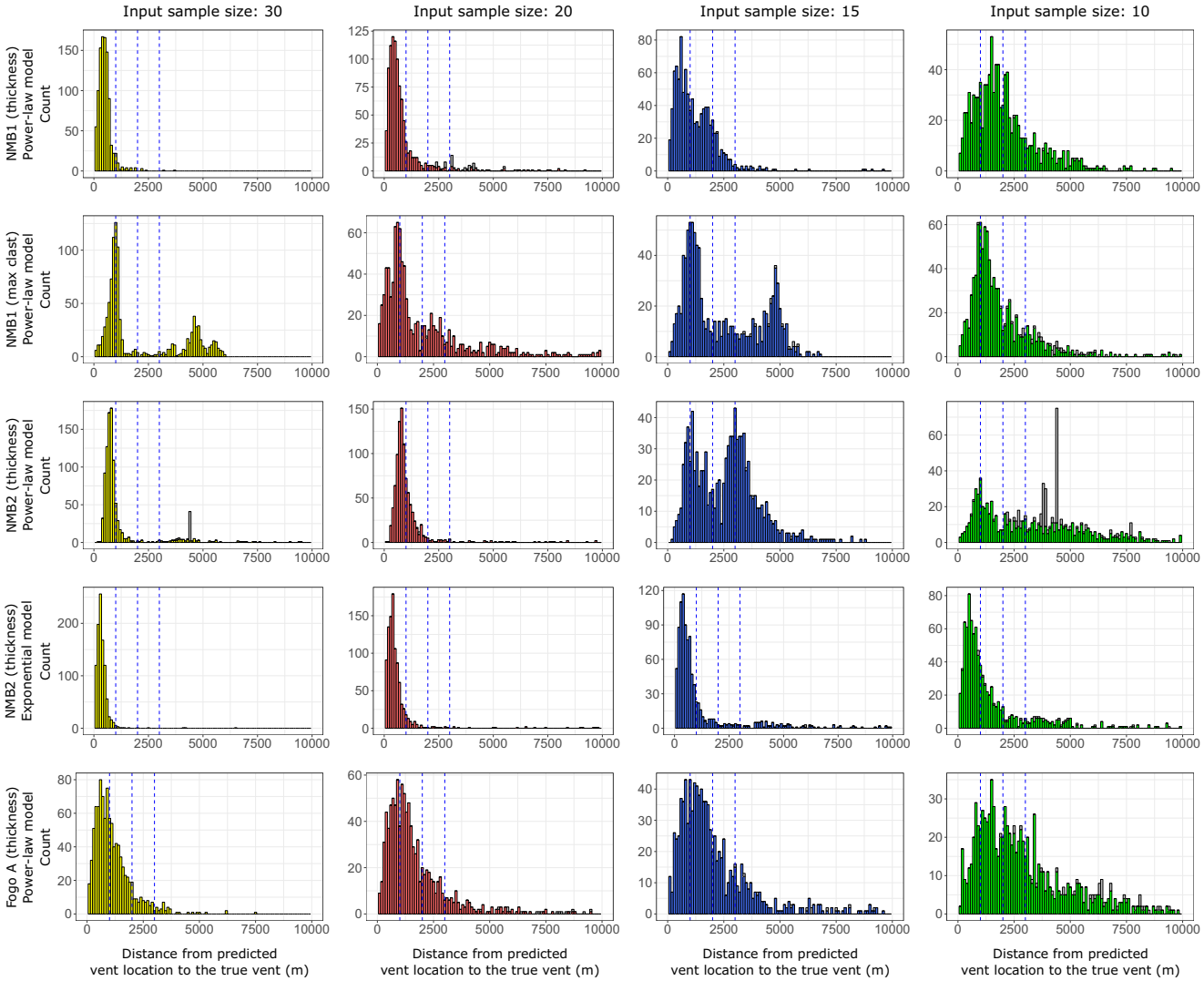


Fig. 7 Distribution of distances from predicted vent location to the center of the true vent location (power-law model is applied to all datasets, and the exponential model is only applied to NMB2). Predicted vents that are more than 10 km from the center of the true vent are rare and are thus not shown in this figure. The dataset and semi-empirical model being used and size of input data are labeled. The three dashed lines mark 1000, 2000, and 3000 m from the center of the vent. Gray bars correspond to unphysical predictions

measurements were made at sites proximal to the vent. The tradeoff between gaining useful information and incorporating observations that represent two dispersal patterns makes the accuracy less susceptible to the size of input dataset in this case. For NMB2, the exponential model outperforms the power-law model particularly given limited observations. The accuracies decrease to 37.9% (15 input points) and 30.8% (10 input points) with the power-law model and to 81.3% (15 input points) and 71.7% (10 input points) for the exponential model.

The above results show that the method can be used to estimate the source vent location of tephra deposits based on limited (≤ 30) thickness or maximum clast size measurements. The results also show that the performance of the method varies with different datasets and the semi-empirical model being used. Results for NMB2 indicate that the power-law model becomes less stable given limited observations (≤ 15).

Predicted dispersal axes

The dispersal axis can be defined by the vent location and dispersal direction. Our hypothesis (2) states that estimating the dispersal axis for tephra fall deposits affected by wind is less impacted (more robust) by different sources of uncertainty (e.g., data scarcity, measurement error, and epistemic uncertainty from the semi-empirical models) compared with estimating the vent location directly.

To test it, we examine the predicted dispersal axes for NMB1 (both thickness and maximum clast size data) and NMB2. The Fogo A dataset is excluded here because the

Table 2 Ratio of predicted vents and dispersal axes that are within 2 km from the center of the vent

Dataset	Input dataset size	Ratio (%) of predictions within 2 km from the true vent center (results from power-law/exponential models)	
		Vent location	Dispersal axis
NMB1 (thickness)	30	99.4/–	99.9/98.5
	20	88.8/–	93.3/72.5
	15	84.1/–	72.5/85.4
	10	56.6/–	46.3/54.8
NMB1 (maximum clast size)	30	65.7/–	27.7/88.1
	20	61.5/–	67.6/67.7
	15	52.4/–	22.9/42.7
	10	65.4/–	32.8/64.2
NMB2 (thickness)	30	84.6/99.4	83.9/98.3
	20	90.8/94.6	90.6/93.0
	15	37.9/81.3	37.9/85.5
	10	30.8/71.7	29.3/72.4
Fogo A (thickness)	30	88.3/–	–/–
	20	72.6/–	–/–
	15	61.8/–	–/–
	10	38.4/–	–/–

Note that results (predicted vent location) from the exponential model are not shown for the NMB1 and Fogo A datasets because it is already known that the exponential model does not work well with them. The Fogo A deposit was not strongly affected by wind, and therefore, the accuracy for predicted dispersal axis is not summarized

deposit was not strongly affected by wind and is composed of two sub-layers. Predicted dispersal axes from using both semi-empirical models are examined with the same input used for the previous section.

Distance from the predicted dispersal axis to the true vent location is summarized as histograms in Fig. 8. Predictions that are unphysical or have inconsistent dispersal direction with respect to the true pattern of the deposit are marked in gray. Here a dispersal direction is assumed to be inconsistent if it was outside the range of dispersal directions defined in Table 1 ± 10° (e.g., for NMB1, inconsistent dispersal direction needs to be outside the range of 8–28° and 30–50°). We use the ratio of dispersal axes that (1) are within 2 km from the center of the true vent, (2) correspond to physical prediction, and (3) have consistent dispersal direction to denote the accuracy (Fig. 9e–g and Table 2). Note that the third criterion may exclude some valid estimations, but such a criterion is necessary because a dispersal axis that is orthogonal to the true dispersal direction of the deposit is not a correct prediction. Here the range for consistent dispersal direction is chosen to be small such that the accuracies are characterized in a more rigorous way. When the exponential model is used, the accuracies are in general greater than the ones calculated from the power-law model. This is especially true given limited input points. With 30 samples, the accuracies are 99.9% and 98.5% for the NMB1 thickness dataset using the power-law and exponential models, respectively. The corresponding accuracies are 27.7% and 88.1% for the NMB1 maximum clast size data and 83.9% and 98.3% for the NMB2 thickness dataset. The sharp contrast in accuracy for the NMB1 maximum clast size data is due to our strict criterion to determine if a dispersal direction is consistent. Many results calculated from the power-law model using the maximum clast size dataset have dispersal directions slightly greater than 50°. They are still excluded for uniform and comparable criterion.

The accuracy decreases with the input dataset size as shown in Fig. 9e–g. For the NMB1 thickness dataset, the accuracies from using the power-law and exponential models are 46.3% and 54.8%, respectively, given 10 input points. For the NMB1 maximum clast size and NMB2 datasets, accuracies are 32.8% (power-law model) and 64.2% (exponential model) and 29.3% (power-law model) and 72.4% (exponential model), respectively.

The results confirm that the exponential model is more robust in predicting the dispersal axis given limited data. We examine the distribution of predicted dispersal axes for NMB1 (excluding results from maximum clast size data to avoid redundancy) and NMB2 in the *x*-*y* plane. The distributions, shown in Fig. 10, are plotted in the following procedure: all dispersal axes are first turned from vectors to grids. For one dispersal axis, its pixelated counterpart is a grid whose cell has the value one if it intersects the dispersal axis and is zero otherwise. As long as the grid for each dispersal axis has the same extent and cell size, the results can be plotted cumulatively by summing up the values for each pixel. The value in

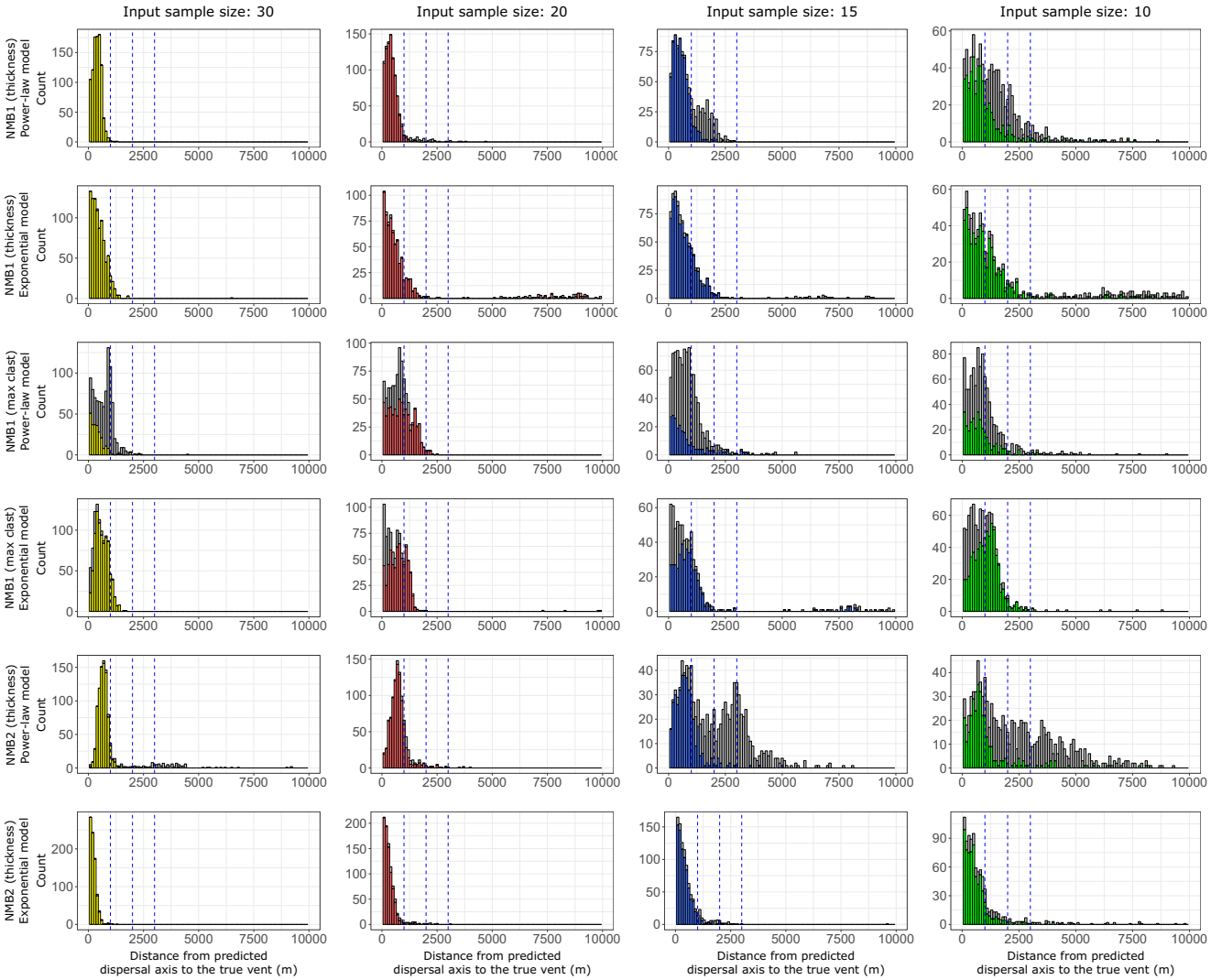


Fig. 8 Distribution of distances from predicted dispersal axis to the center of the true vent location (results from both power-law and exponential models; the Fogo A deposit is not used as an example here because it was not strongly affected by wind). The dataset and semi-empirical model being used and size of input data are labeled. The three dashed lines mark 1000, 2000, and 3000 m from the center of the vent. Predicted dispersal axes that are more than 10 km from the center of the true vent are rare and are thus not shown in this figure. Gray bars correspond to predictions that are unphysical or have inconsistent dispersal direction

each cell denotes the number of times that the dispersal axes cut across it. With 30 and 20 input points, for both datasets, distributions of the dispersal axes derived from the two semi-empirical models are well-constrained and consistent with the true dispersal patterns of the deposits. With 15 and 10 input points, the variability in the distribution increases. Notably, when the power-law model is used, more dispersal axes with incorrect dispersal pattern occur, while with the exponential model, most results are still consistent with the dispersal pattern of the deposits.

The results suggest that the predicted dispersal axes from the method can be used to constrain the source vent location of tephra deposits affected by wind. Given limited data, the exponential model is more stable and provides more accurate predictions compared with the power-law model.

Results from localized subsets

To analyze how the method reacts to tephra deposits whose transport was affected by change in wind direction, localized subsets of the NMB1 (both thickness and maximum clast size) and NMB2 datasets are applied to the method. The Fogo A deposit is not examined here due to severe local variation. Since sample site layout for the two deposits is oriented in the north-south direction, we take local subsets each containing 30 samples by a moving window along the latitudinal direction. As an example, for the first subset, we take the 30 southernmost sample points as input, and in the next subsetting, we exclude the southernmost and include the 31st southernmost measurement in the input dataset.

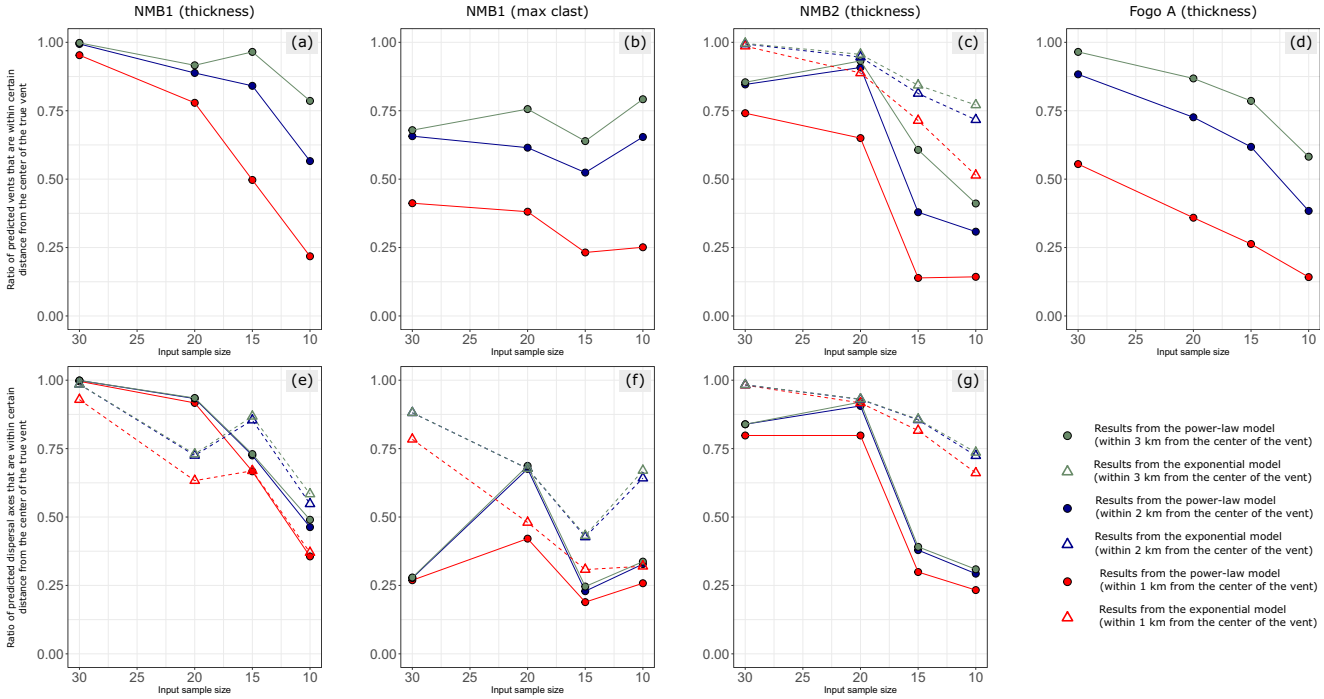


Fig. 9 **a**–**d** Ratio of predicted vent locations (*y*-axis; excluding unphysical predictions) that are within 3000 (green), 2000 (blue), and 1000 (red) m from the center of the true vent for different deposits with different sample sizes (*x*-axis). Results from the power-law (circle) model are shown for all datasets, and results from the exponential model (triangle) are shown only for the NMB2 thickness dataset. **e**–**g** Ratio of predicted dispersal axes (excluding predictions that are unphysical or have inconsistent predicted dispersal direction) that are within 3000 (green), 2000 (blue), and 1000 (red) m from the center of the true vent for NMB1 and NMB2. Results from using both semi-empirical models are shown (circle and triangle for results from the power-law and exponential models, respectively)

We apply both semi-empirical models to the thickness and maximum clast size datasets of NMB1 and NMB2. The results are shown in Figs. 11, 12, and 13, which display how the predicted vent location and dispersal direction change with the mean latitude of the subsets sampled from the moving window. Results for NMB1 derived from thickness data show that with input points proximal to the source vent, predictions are located at the true vent location regardless of which semi-empirical model is used (Fig. 11). As the moving window moves northwards, predicted vents move towards the north-northeast along the dispersal axis. There is a series of subsets with mean latitude below 4,200,000 that lead to unphysical results (red box in Fig. 11a, b), regardless of which semi-empirical model is used. This range of latitude corresponds to the area where the dispersal direction changes from northeast to north-northeast. Subsets with mean latitude greater than 4,200,000 lead to predictions that are in the north-northeast of the true vent along the dispersal axis. How the predicted dispersal direction changes with the mean latitude of each subset is better illustrated in Fig. 11c. For the southernmost subsets (represent proximal portion of the deposit), the predicted dispersal direction is mainly towards northeast (with a few outliers from the power-law model). For subsets further north, the predicted dispersal directions are correct and stable, pointing towards north-northeast. For subsets with mean latitude greater than 4,215,000 (near the northern shoreline of Mono Lake), unphysical predictions start to occur. This is because the deposit is thin (observed thickness range 1–24 mm) at sample sites farther north, and the ratio of useful information indicative of vent location or dispersal pattern to measurement error decreases.

Results (Fig. 12) derived from the maximum clast size data of NMB1 are similar to the ones shown in Fig. 11. Most predicted vents are close to the true vent location. The change in dispersal direction from north-northeast to northeast can be observed as the mean latitude of input dataset increases (Fig. 12c). Unphysical results are also obtained when sample sites of the input are close to the area where the dispersal pattern of the deposit changes (red boxes in Fig. 12a, b). It is noticed that when the power-law model is used, dispersal directions greater than 90° are obtained. This is probably related to the geometry of the vent that is oriented along this direction (Bursik 1993, Fig. 1a) and again implies the instability of the power-law model.

For NMB2, most predictions (Fig. 13a, b) using both models are located at or near the Upper Dome and are not strongly affected by the mean latitude of input sample sites. Overall, the exponential model outperforms the power-law model, since a few predictions from the latter

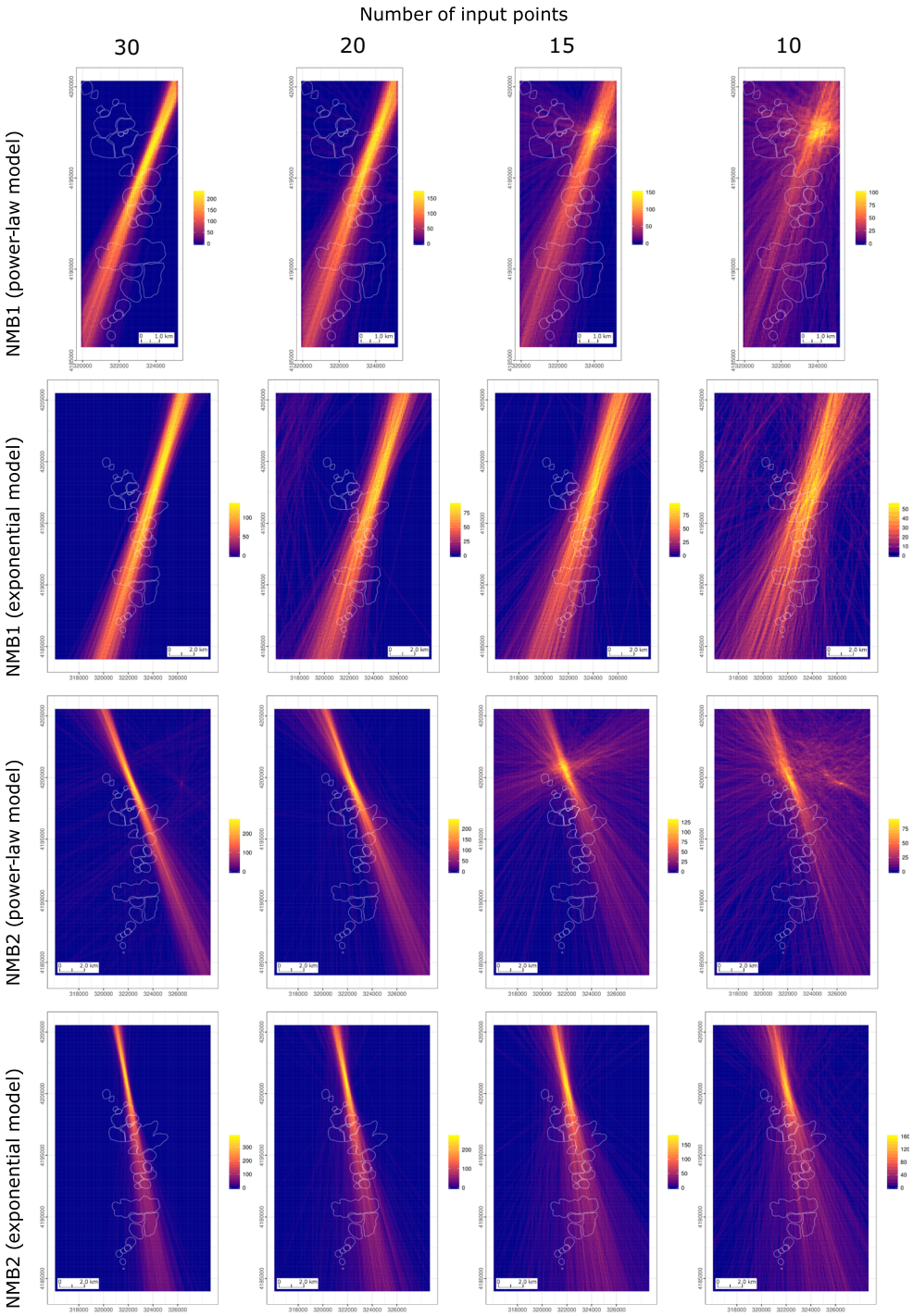


Fig. 10 Distribution of predicted dispersal axes from the results shown in Fig. 8 for NMB1 and NMB2 (using thickness data as input). The value of each cell represents the number of dispersal axes that intersect it. As the size of input dataset decreases, results from the power-law model become less stable

are located near the southern shoreline of Mono Lake, far from the source region. Predicted dispersal directions of NMB2 (Fig. 13c) are relatively stable with small fluctuations when southernmost subsets near the vent are used as input. A few subsets with mean latitude just above 4,200,000 lead to inconsistent dispersal directions. With the power-law model, the deviation from the true dispersal direction is greater. Dispersal directions from the power-law and exponential models are slightly different given subsets with mean latitude greater than ~4,205,000. Based on our moving-window sampling scheme, this difference is introduced by the inclusion of a single measurement and exclusion of another. This shows that one or two measurements could greatly affect the prediction from the power-law model, suggesting its instability. Results in Figs. 11, 12, and 13 show that by applying localized subsets to the method, the corresponding results reflect the local dispersal pattern of the deposit.

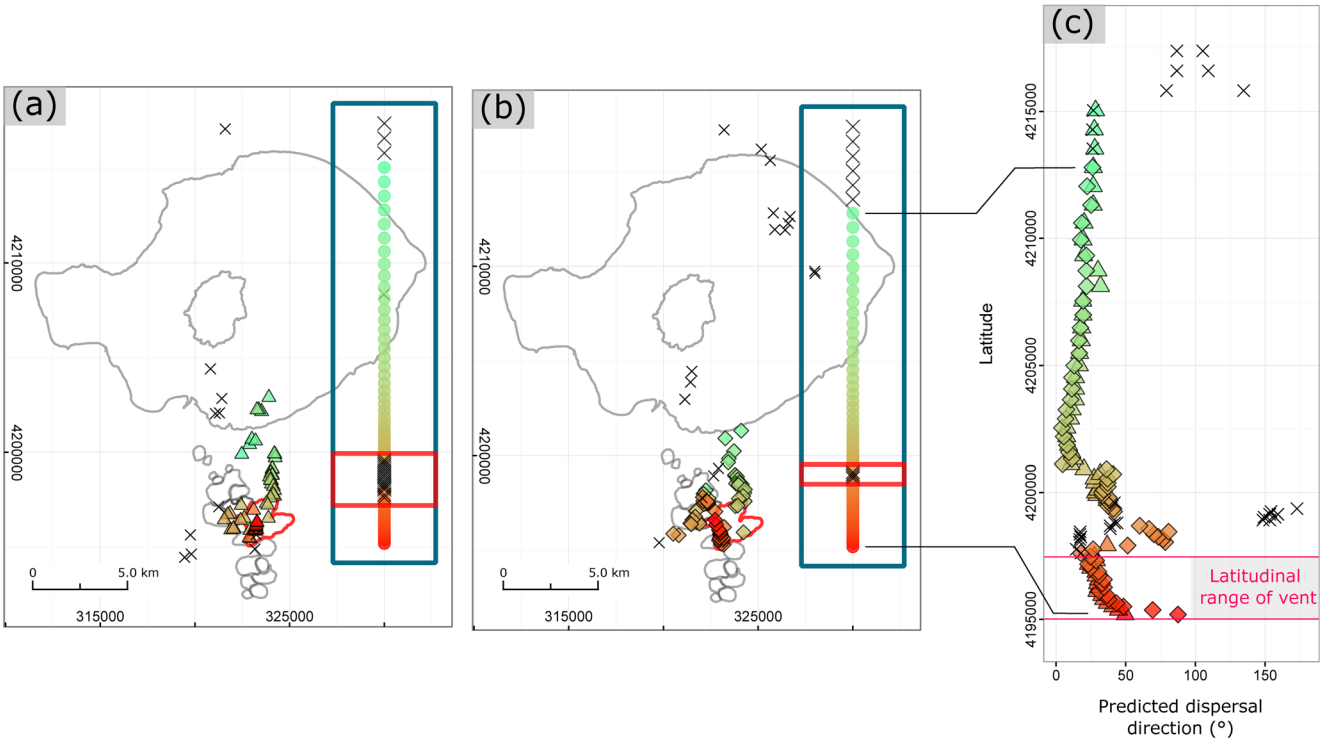


Fig. 11 Estimated vent locations of NMB1 derived from the power-law (**a**, triangle) and exponential (**b**, diamond) models using localized thickness observations (size 30). The subsetting is done using a moving window scheme along the latitudinal direction. The mean latitude of each subset is plotted in the blue box in **a** and **b**. Corresponding predicted vents are plotted in the map with consistent color. Crosses in the blue box and map represent unphysical predictions. **c** Mean latitude of each subset is plotted against the corresponding predicted dispersal direction (in degrees from north clockwise). The occurrence of crosses and color of the points in **c** are consistent with **a** and **b**

Results from statistical bagging

The Fogo A deposit provides a great example for us to examine the performance of the method when the deposit is characterized by severe local thickness variation and (or) composed of multiple lobes. Our experiments have shown that the accuracy is generally lower for the Fogo A deposit. We couple statistical bagging (bootstrap aggregating; Breiman 1996) with the method to test if observations that greatly affect the performance of the method can be detected in an objective way. The main idea of statistical bagging is to group or pick input data based on the corresponding outputs, which are derived from applying a certain model to the input data. In our case, the input and output data are thickness observations and predicted vent locations, respectively. Statistical bagging is implemented in the following way: subsets containing 30 samples are drawn randomly from a complete input dataset, 5000 times with redrawing, i.e., putting the data points back after each draw, and then, the method is applied. We pick out observations that could greatly affect the performance of the method by examining the 5000 predicted vent locations. In the case of the power-law model, which is used here since it works better with the Fogo A deposit, the resultant predictions are shown in Fig. 14a. Predictions outside the box in the figure are assumed to be outliers, as they are farther from the vent. The criterion for outliers can be defined in a flexible way. If the exact vent location is unknown, the outlier can be defined as predictions that are far from the center of the predicted vent "cloud." Here the vent location is assumed to be known, such that we could find out which observations contribute to incorrect predictions. The experiment gleans 848 outlier predictions, and the corresponding input subsets are gathered and combined to a single dataset. The occurrence of each thickness measurement in this combined dataset is summarized and plotted as a histogram in Fig. 14b. If each sample point has the same probability of leading to outlier, the histogram should resemble a uniform distribution, and their occurrence should all fluctuate slightly around $138 \approx 848 \times 30/184$ (the green horizontal line in Fig. 14b). However, it is clear that the occurrence of outlier prediction is dependent on the exclusion or inclusion of certain samples. We pick out these points by subjectively defined thresholds (red and yellow horizontal lines in Fig. 14b). Their thicknesses and those of nearby sites are marked in Fig. 14c.

The inclusion of yellow points in Fig. 14c in the input tends to have correct predictions. All but one of them are greater than 5 m thick. This is not surprising, as the source vent should be close to thicker sample sites. The inclusion of red points in

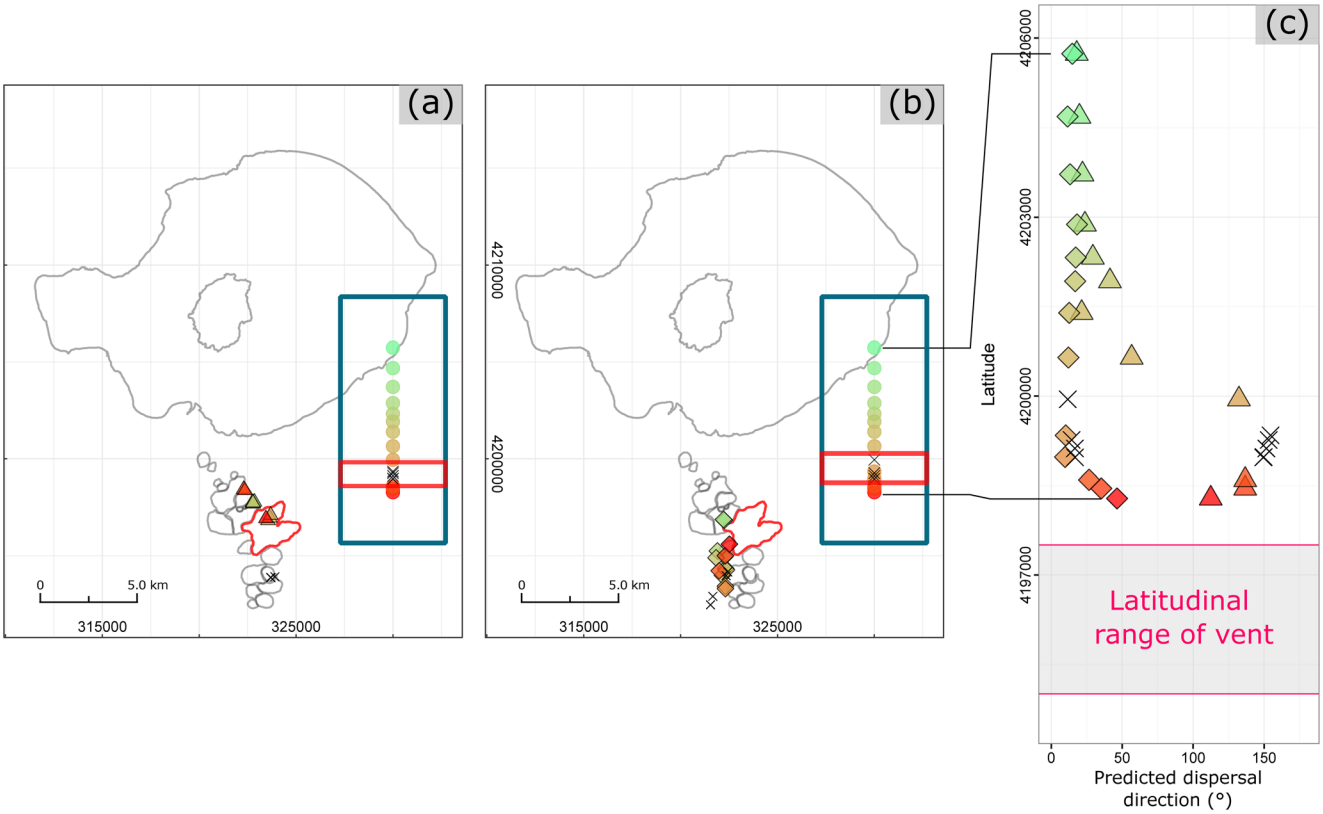


Fig. 12 Estimated vent locations of NMB1 derived from the power-law (**a**, triangle) and exponential (**b**, diamond) models using localized maximum clast size observations (size 30). The subsetting is done using a moving window scheme along the latitudinal direction. The mean latitude of each subset is plotted in the blue box. Corresponding predicted vents are plotted in the map with consistent color. Crosses in the blue box and map represent unphysical predictions. **c** Mean latitude of each subset is plotted against the corresponding predicted dispersal direction (in degrees from north clockwise). The occurrence of crosses and color of the points in **c** are consistent with **a** and **b**

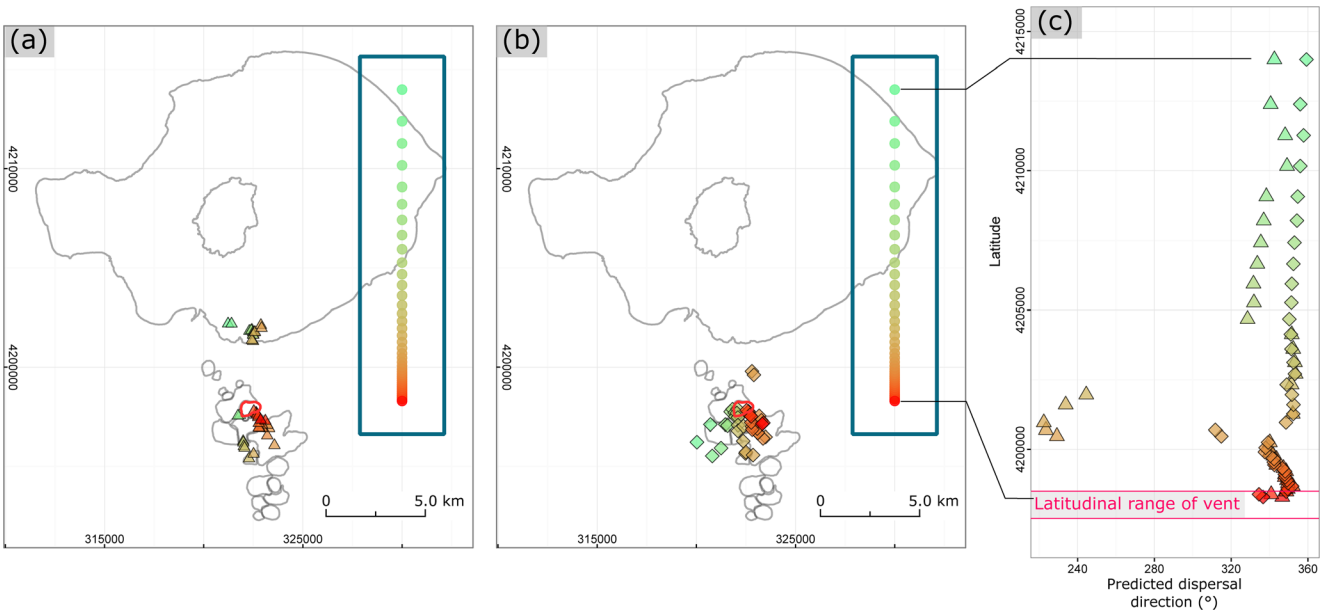


Fig. 13 Estimated vent positions of NMB2 derived from the power-law (**a**, triangle) and exponential (**b**, diamond) models using localized thickness observations (size 30). The subsetting is done using a moving window scheme along the latitudinal direction. The mean latitude of each subset is plotted in the blue box. Corresponding predicted vents are plotted in the map with consistent color. No unphysical predictions occur for the NMB2 dataset. **c** Mean latitude of each subset is plotted against the corresponding predicted dispersal direction (in degrees from north clockwise). Color of the points in **c** is consistent with **a** and **b**

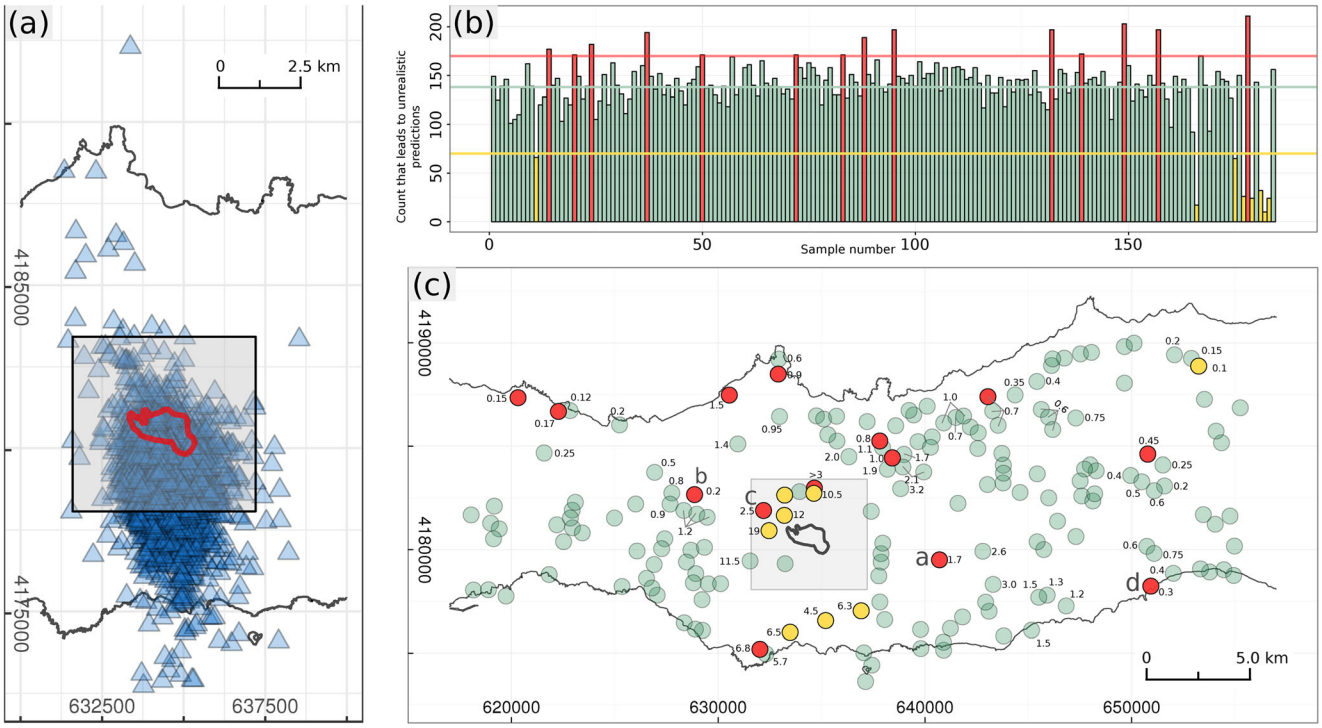


Fig. 14 Result from statistical bagging. **a** Predicted vent locations (blue triangles) of the Fogo A deposit from 5000 randomly selected subsets with 30 input points using the power-law model. Predictions outside the bounding box are assumed to be outliers. Input subsets leading to outlier predictions are collected and combined to a single dataset, and the occurrence (*y*-axis) of each observation within the combined dataset is plotted as a histogram in **b** (*x*-axis, index of each sample point). There are 848 outlier predictions, which suggests a total number of 25,440 = 848 × 30 points are drawn. If the likelihood of contributing to outlier is the same for each sample, the expected value in **b** should be 138 ≈ 848 × 30/184 (green horizontal line in **b**). Some points have a higher or lower chance of contributing to outlier predictions. Such points are selected based on two thresholds marked as yellow and red horizontal lines in **b**. Thicknesses (m) of these points and their neighboring sites are marked in **c** with corresponding color

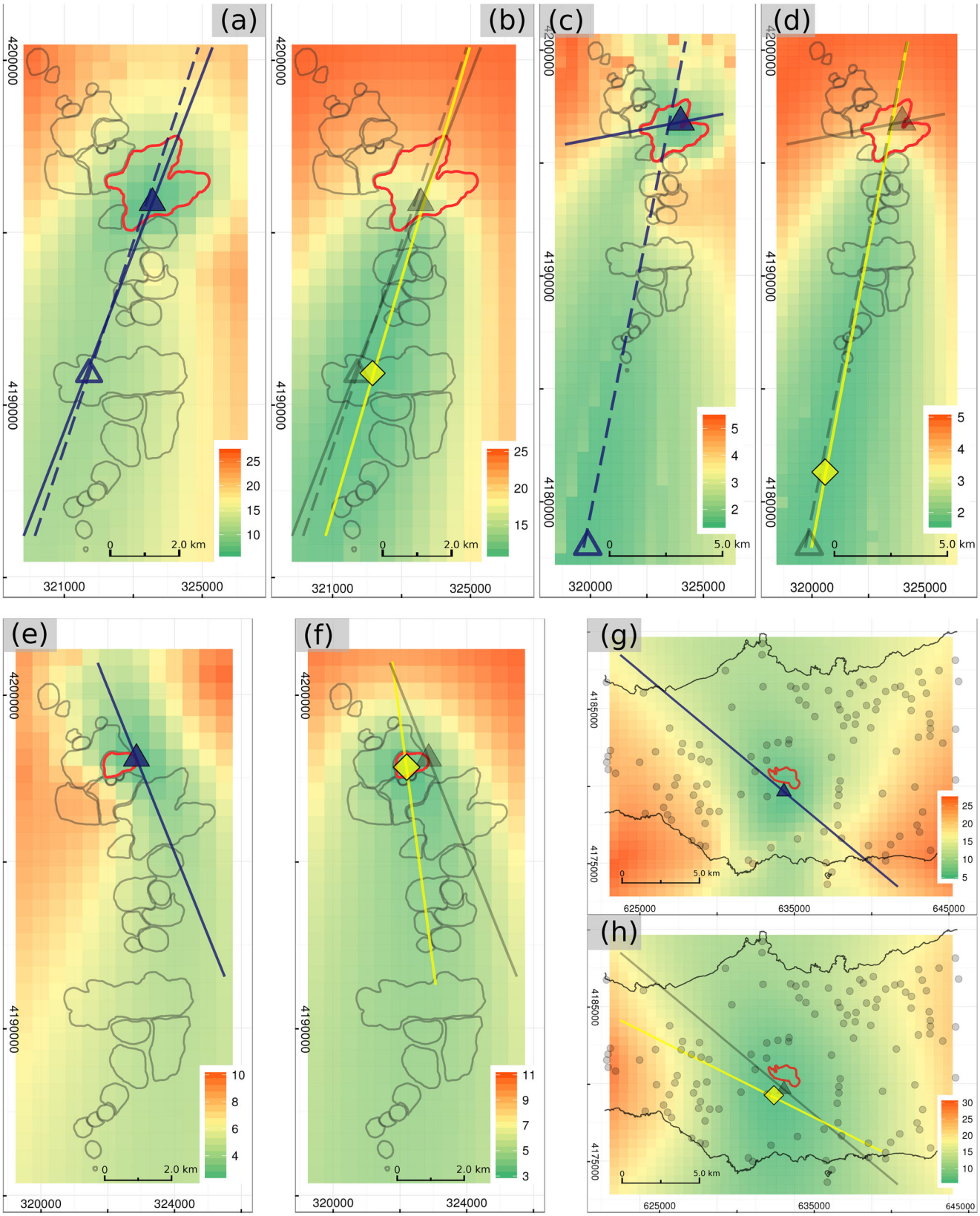
Fig. 14c in the input tends to have outlier predictions. Some of these points have distinct measured thickness compared with that of their neighboring sites. For example, the deposit at Points a, b, and c in Fig. 14c is much thinner compared with their neighboring sites. We also notice that thicknesses of Point d and its neighboring sites in Fig. 14c imply that multiple lobes exist for the deposit. Without the procedure of statistical bagging, we find it hard to detect this point subjectively. In addition, five of these red points were measured along the northern coast of the São Miguel Island, which is probably related to local erosion (Walker and Croasdale 1971b).

The above results show that the coupled use of statistical bagging and the method has the ability to detect observations that cannot be coherently fitted by the semi-empirical model. Some samples selected from this procedure that contribute to outliers are characterized by severe local variation or are indicative of multiple lobes. These results indicate that the coupled bagging and gradient descent approach has promise in quantitatively indicating the likelihood of a set of samples correctly predicting the source vent. The simplicity of the semi-empirical model suggests that epistemic uncertainty cannot be neglected. With the introduced procedure, it is expected that the detection of observations that are potentially critical to the interpretation of a tephra deposit can be done in an objective way.

Surface of the cost function

Visualizing the surface of the cost function $g(x_s, y_s)$ (Eq. 6) is another way of using the method. It is useful given limited observations. However, the complete dataset is used as input here, such that intrinsic and systematic factors affecting the performance of the method can be found. Surfaces of the cost function for each dataset using both semi-empirical models as

Fig. 15 Surface of the cost function (Eq. 6) and predicted vent location ▶ and dispersal axis using the complete dataset as input. **a**, **c**, **e**, **g** Results for NMB1 (thickness and maximum clast size data), NMB2, and the Fogo A deposit derived from the power-law model (solid triangle and solid line for global minimum and corresponding dispersal axis, respectively). Predictions at local minimum and corresponding dispersal axes are also shown (void triangle and dashed line) if present. **b**, **d**, **f**, **h** Results for NMB1 (thickness and maximum clast size), NMB2, and the Fogo A deposit derived from the exponential model (yellow diamond and solid line for global minimum and dispersal axis, respectively). Local minima do not occur when the exponential model is used. Results from using the power-law model are also displayed with transparent color in **b**, **d**, **f**, and **h** for easier comparison


(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
4200000
4190000
4180000
4185000
4175000
321000
325000
320000
324000
625000
635000
645000
0
2.0 km
5.0 km

well as the local and global minima and the corresponding dispersal axes are shown in Fig. 15. With the power-law model applied to the NMB1 thickness and maximum clast size datasets, local minimum occurs (Fig. 15a, c). When the exponential model is applied to the NMB1 datasets, no local minima occur (Fig. 15b, d). The local minimum with the power-law model for the NMB1 datasets is almost identical to the global minimum calculated from the exponential model. The corresponding dispersal axes are also highly consistent except for the global minimum using NMB1 maximum clast size dataset with the power-law model (Fig. 15c). The dispersal axis corresponding to the global minimum calculated from the NMB1 maximum clast size dataset is pointing towards northeast, reflecting the dispersal pattern of the proximal deposit. Local minima do not occur for the NMB2 and the Fogo A deposits regardless of which semi-empirical model is used (Figs. 16e–f). Surfaces of the cost function for the NMB1 datasets indicate that characteristics of the deposits in addition to data scarcity could also contribute to local minima.

Discussion

Our results indicate that the method is able to estimate the vent location of tephra deposits based on thickness or maximum clast size measurements. The dispersal axis calculated from the method can be used to constrain the vent location. Performance of the method varies with the deposit being analyzed as well as the use of different semi-empirical models. In particular, the power-law model outperforms the exponential model in cases of NMB1 and the Fogo A deposit, and the exponential model leads to more accurate and stable results

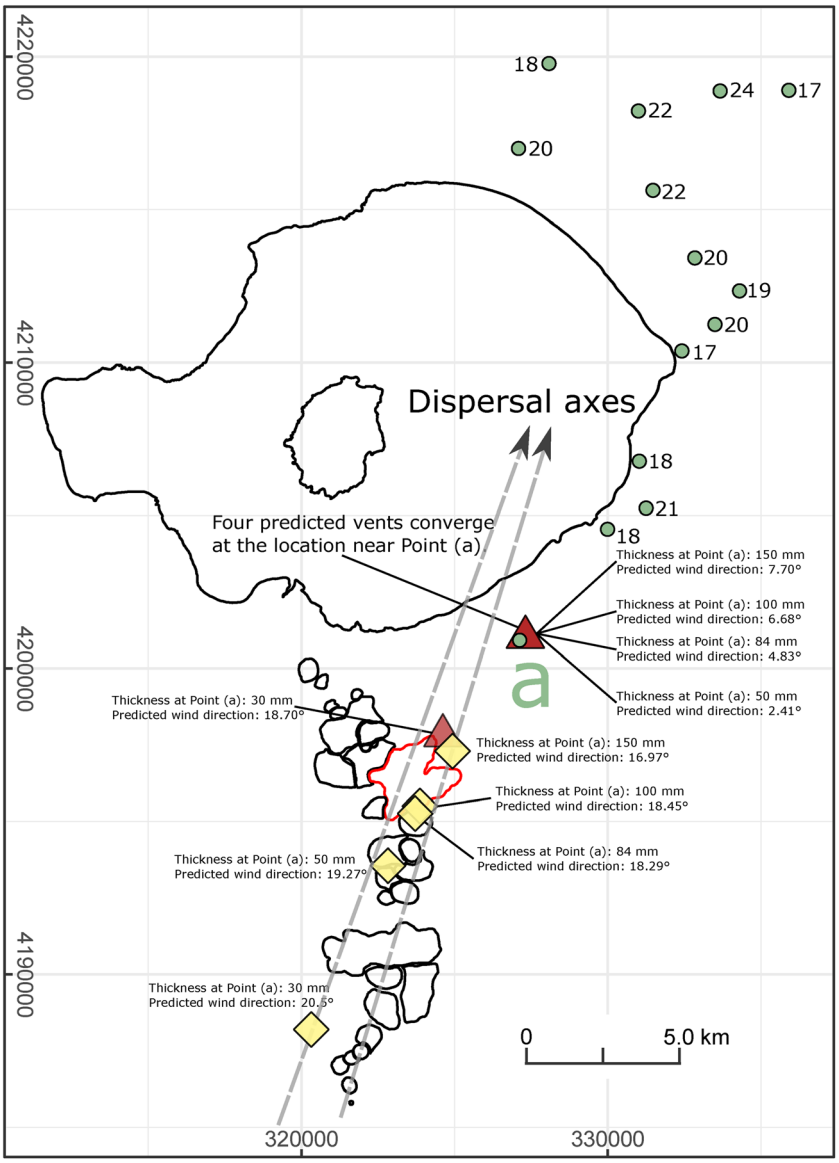


Fig. 16 Fourteen samples (thickness marked in mm) from the NMB1 thickness dataset are selected and applied to the method with both semi-empirical models. The observed thickness at Point a is 84 mm. Its thickness is reset to 50, 84, 100, and 150 mm and can be regarded as four datasets (their only difference is the thickness at Point a). The four datasets are applied to the method. Results for the power-law model are red triangles. Only one of them is located near the North Coulee (when Point a is set to be 30 mm thick), and the rest of them are located at the same position near Point a. Predicted vent locations (yellow diamond) from the exponential model are all located near the true dispersal axis, and the predicted dispersal directions are consistent (17.0–20.5° from north clockwise) with the dispersal pattern of the deposit. Two dispersal axes derived from the exponential model are shown that correspond to the results when Point a is set to be 30 and 150 mm thick, respectively, as end members. The different results highlight that the power-law model is less stable given sparse dataset for the method and is biased towards measurement(s) with greater magnitude

for NMB2. With limited observations, however, the exponential model is more stable and provides more accurate estimate on the dispersal axis. Experiments with the NMB1 and NMB2 datasets show that this method is able to detect local change in dispersal direction for tephra deposits given localized observations. The coupled use of statistical bagging with the method is able to pick out observations that may bear critical information about a deposit that cannot be well explained by the semi-empirical model. The simplicity of the forward model calls for attention to epistemic uncertainty. The following discussion focuses on factors that may affect the performance of the method viewing from different perspectives.

Features of tephra deposits affecting the performance of the method

Regardless of the eruption style as well as other physics governing the transport and sedimentation of tephra deposits, their net impact on tephra thickness or maximum clast size distribution can be described as thinning or fining along different directions with respect to the source vent. Therefore, we focus more on the physical properties of tephra deposits.

Decay pattern of thickness and maximum clast size

From the log(thickness)−$\sqrt{\text{isopach area}}$ plot (Pyle 1989), it was found that the thinning rate of tephra deposit could change with $\sqrt{\text{isopach area}}$. Segmented or non-segmented exponential (Pyle 1989), power-law (Bonadonna et al. 1998), and Weibull functions (Bonadonna and Costa 2012) have been proposed to address this concern. The different decay patterns may be related to the different dynamics near the vent (Bursik et al. 1992b), particle Reynolds number (Bonadonna et al. 1998), and many other factors, but the overall impact can be reflected from the log(thickness)−$\sqrt{\text{isopach area}}$ plot.

Similarly, the key difference between the power-law and exponential models is that the thinning or fining rate is not constant for the former under log scale. The flexibility of the power-law model makes it function well for the NMB1 deposit which is characterized by two distinct decay patterns in thickness and maximum clast size (Fig. 1e). The flexibility of the power-law model also makes it easier to be fitted to the Fogo A deposit which is composed of two lobes and characterized by local variation. For the exponential model, with the NMB1 dataset, as the result of constant decay rate under log scale, proximal observations exhibiting rapid decay in thickness and maximum clast size are treated as measurement errors by the method instead of systematic variation. This explains why predicted vent locations of NMB1 using the exponential model are often located along the upwind portion of the dispersal axis.

It is worth noting that local minimum from the power-law model is similar to the global minimum predicted from the exponential model for the NMB1 datasets (Fig. 15a, c). If fewer observations are made for the proximal portion of NMB1, predicted vents from the power-law and exponential models should both be located at the upwind portion of the dispersal axis (at the global minimum for the exponential model). In such circumstances, vent locations predicted from both semi-empirical models cannot be relied on, and we should instead use the dispersal axis to constrain the vent location. For NMB2, its decay rate does not change under log scale, which can be better modeled by the exponential model.

Wind directions

Thickness or maximum clast size distribution of tephra deposits can be greatly affected by wind. Under the assumption that the wind direction is constant, the axisymmetric (with respect to the dispersal axis) thickness or maximum clast size distribution of tephra deposits is supported by the physics of tephra transport and sedimentation.

Entrainment of horizontal momentum by a volcanic plume can lead to plume bending (Bursik 2001) or cause the development of a downwind propagating gravity current at the neutral buoyancy level (Bursik et al. 1992a). Ultimately, the horizontal transport of volcanic ash in the atmosphere is dominated by wind advection and turbulent diffusion (e.g., Suzuki 1983; Bonadonna et al. 2005; Schwaiger et al. 2012). The overall axisymmetric (with respect to the centerline of the eruption column) spreading of tephra when there is no wind, as a result of vertical plume rise, spreading as an axisymmetric gravity current, and turbulent diffusion, becomes less apparent in the presence of strong wind or weak plume. Ideal plume spread in the absence of wind leads to axisymmetric tephra thickness distributions on the ground. In the presence of wind, the solution to an advection-diffusion equation in 2D with a continuous point source is asymmetric up- and downwind with respect to the wind direction (Csanady 1973; Suzuki 1983), but has planar symmetry along the wind axis. Differences between this pattern and the output of tephra transport and sedimentation models (Suzuki 1983; Bursik et al. 1992a; Bonadonna et al. 2005; Schwaiger et al. 2012; Klawonn et al. 2012) are that the tephra is emitted from a vertical line source and then falls at terminal velocity in the z-direction. However, the resulting planar symmetry in the deposit is independent of eruption parameters such as column height and total volume and only relies on the assumption that the prevailing wind direction is constant. This justifies using dispersal axis to constrain the vent location given limited data: limited data may not be sufficient enough for the method (or any other methods) to capture features about how a deposit decays in thickness and maximum clast size with distance

along different directions, but the axisymmetric (with respect to the dispersal axis) thickness or maximum clast size distribution is easier to recognize.

When a tephra deposit was affected by change in wind direction during its deposition, epistemic uncertainty arises due to the simplicity of semi-empirical models. Our experiments with NMB1 and NMB2 with localized input data suggest that the method is able to identify the change in dispersal direction (Figs. 11, 12, and 13). Given sufficient data, use of the method could provide objective and more detailed constraints on the dispersal pattern of a tephra deposit. On the other hand, with limited observations, the outcome from the method might not be representative of the overall or local dispersal pattern. Under such circumstances, we recommend the use of leave-one-out validation or statistical bagging to separate or group observations that cannot be coherently explained by the semi-empirical model and analyze them separately.

Local variations and multiple lobes

Our experiments with the Fogo dataset show that the performance of the method is affected by local variation and presence of multiple lobes. Given sufficient observations, the coupled use of the method and statistical bagging is shown to be effective in selecting potentially critical observations that cannot be well explained by the semi-empirical model. Given sparse observations, statistical bagging can be replaced by leave-one-out or *k*-fold cross-validation with the same purpose.

Technical concerns

The instability of the power-law model

Viewing the method as a tool that solves least-square problem provides new perspectives about it and the semi-empirical models. In our results, it has been shown several times that the prediction from the power-law model becomes less stable given sparse observations. This is due to its greater degrees of freedom. Solving the least-square problem given fixed vent location and dispersal direction using pseudo-inverse for the power-law model gives more weight to measurements with greater magnitude. The distal observations, which usually preserve useful information about the dispersal pattern of the analyzed deposit, are given less weight. Results derived from using the power-law model highlight the tradeoff between flexibility and instability. With limited observations, the power-law model needs to be used with caution in predicting either the vent location or dispersal axis.

Log transformation

The observed thickness or maximum clast size measurements are log-transformed for the method. Theoretically, this should cause the prediction to be more influenced by small observations that are relatively far from the source vent, but the use of log transformation is justified here. Tephra deposits do not thin or fine linearly with distance (e.g., Carey and Sparks 1986; Pyle 1989; Bursik et al. 1992a, b; Fierstein and Nathenson 1992; Bonadonna et al. 1998; Kawabata et al. 2013; Yang and Bursik 2016). If raw observations are used as input, the prediction will be strongly affected by the measurements with large magnitude (higher values). This is fine, given sufficient observations. With limited and unevenly distributed samples, however, one or two measurements could be much greater than the others in magnitude. Under such conditions, if raw measurements are used as input, the predicted vent location would be close to the sample site with measurement with the greatest magnitude, and the method would be unable to capture the overall decay pattern of the studied deposit.

Distribution of sample sites

Although our results suggest that given localized subsets, the method is able to detect the local dispersal pattern of a tephra deposit, it is noted that with localized or clustered sample sites, the system would become "more overdetermined." As sample sites close to each other tend to have similar thickness and maximum clast size measurements, measurement error becomes relatively greater, which increases the uncertainty of the method.

Uncertainty propagation and scale dependence of accuracy

The accuracy of the prediction should be scaled with the area of interest or sampled area. It is natural that using more proximal observations could lead to more accurate predictions. On the other hand, however, for tephra deposits of continental scale, a rough but valid estimate on the vent location or dispersal axis could help a lot in constraining the potential vent location.

Given the simplicity of the semi-empirical models, our discussion mainly focuses on how and if epistemic uncertainty can be detected. To characterize measurement or aleatory uncertainty, one could simply apply datasets generated from Monte Carlo sampling based on original datasets to the method and examine the distribution of predictions. Our experiment that applies randomly selected subsets to the method ("Results from random subsetting") can be viewed as a similar procedure except that epistemic uncertainty is also taken into account.

Inclusion of zero observations

In the present work, locations with no tephra observed are not used as input for the method. A plus-one transformation, namely, adding one unit of thickness or maximum clast size to all observations, can effectively incorporate sample sites with no tephra observed into the input. The plus-one transformation has been adopted in Yang and Bursik (2016) to estimate the extent of tephra deposits. Here it is not used due to the concern that arbitrary number of zero-valued samples could be added to the input, which introduces additional uncertainty to the estimate. Unless the site with no tephra observed is believed to be close to the extent of the tephra deposit, they should not be used as input for the method with plus-one transformation.

Suggestions on using the method with datasets of different sizes and qualities and case study

Based on our experiments and discussion, suggestions on how to work with the method given dataset of different sizes and qualities are listed in Table 3. If 6–10 observations are available, users should couple the exponential model with the method given its stability. If observations suggest that the deposit was affected by wind, the dispersal axis should be used to constrain the vent location. Leave-one-out (or leave several observations out) validation or examining the surface of the cost function can be used to detect epistemic uncertainty and local minima. Given 11–30 observations, if a rapid change in decay rate can be confirmed, the power-law model is preferred to predict the vent location, and users should use the exponential model and its predicted vent location or dispersal axis (depending on prior knowledge about the deposit) to constrain the vent location otherwise. Statistical bagging or leave-one-out validation could be coupled with the method to detect the occurrence of multiple lobes, local thickness/maximum clast size variation, and whether the deposit was affected by change in dispersal direction. With more than 30 observations, if a rapid change in decay rate can be confirmed, the power-law model is preferred to predict the vent location, and users should use the exponential model to predict the vent location otherwise. Statistical bagging and applying localized subsets to the method can be used to detect the occurrence of multiple lobes, local thickness/maximum clast size variation, and whether the deposit was affected by change in wind direction. The following datasets with different sizes and features (e.g., thinning rate, sample site distribution, and sample site layout) are used to demonstrate the validity of the suggestions.

Case one: a subset of the NMB1 thickness dataset

We select 14 sample points from the NMB1 thickness dataset (green points in Fig. 16 with thickness measured), including one observation that is much thicker and closer to the source vent (Point a in Fig. 16; measured thickness: 84 mm). The selected sample sites are mostly located on the right side of the main dispersal axis of NMB1. They are clustered along the east and northeast shoreline of Mono Lake and are not the most proximal sample sites. Their thickness measurements vary within the range of 17–24 mm except for Point a and display local variations. We found it hard to subjectively construct isopachs or infer the dispersal axis given these observations.

We set the thickness at Point a to be 30, 50, 84, and 150 mm; apply the modified datasets to the power-law and

Table 3 Suggestions on how to work with the method given dataset of different sizes and qualities

Dataset size	Suggested semi-empirical model	Vent location or dispersal axis	Measures to detect epistemic uncertainty
6–10	Exponential model	Vent location/dispersal axis	Leave-one-out validation Examining the surface of the cost function
11–30	Change in decay rate recognizable: power-law model	Vent location	Statistical bagging or leave-one-out validation
	No change in decay rate: exponential model	Vent location/dispersal axis	Statistical bagging or leave-one-out validation
	Change in decay rate unrecognizable: exponential model	Dispersal axis	Statistical bagging or leave-one-out validation
>30	Change in decay rate recognizable: power-law model	Vent location	Statistical bagging Applying localized subsets to the method
	No change in decay rate: exponential model	Vent location	Statistical bagging Applying localized subsets to the method

exponential models, respectively; and examine the predicted vent locations and dispersal axes. Three predicted vent locations using the power-law model (predicted dispersal direction 2–7° from north clockwise) are concentrated at a location very close to Point a. This shows that with the power-law model, the method is strongly "biased" towards measurements with greater magnitude. The only exception occurs when thickness of Point a is set to be 150 mm, and the corresponding predicted vent is located near North Coulee (Fig. 16). For the exponential model, predicted vents are all located along the dispersal axis of the deposit within a narrow range (bounded by the two dispersal axes plotted), and predicted dispersal directions (16–21°) are consistent with the true transport pattern of NMB1. This example firmly confirms the instability of the power-law model for the method and supports our proposition that it is more robust to estimate the dispersal axis using the exponential model given limited data.

Cases two and three: THS and Rockland tephras

The THS and Rockland tephras are widely dispersed and significant Quaternary deposits, and their primary thickness has only been measured at eight sites each. Examining the surface of the cost function and leave-one-out validation have been applied to these datasets. Six sample sites of the THS tephra have been observed in northwestern Nevada, where the thickness shows great and systematic variation within a narrow swath. This suggests that transport was subject to a strong prevailing wind. Local minima tend to occur, given the paucity of sample sites. For the Rockland tephra, the sample sites span a broader area, and local minima do not occur in the surface of the cost function (not shown to avoid complexity). Since two thickness measurements might be reworked for the Rockland tephra, we include and exclude them to generate two input datasets. This procedure is comparable to cross-validation. The exponential model is used for both cases given the limited observations.

Surface of the cost function for the THS tephra in the *x*-*y* plane is plotted in Fig. 17a. Three local minima are obtained with one (green triangle in Fig. 17a) being unphysical; the others are close to Medicine Lake and Mount Shasta (see Fig. 17b for the location of potential vents). Medicine Lake and Mount Mazama are near the two, respective dispersal axes. The two predictions arise from the uncertainty associated with the two sample sites in California, which could be on the same or opposite side of the dispersal axis. Since the deposit was affected by strong wind, the predicted dispersal axis is preferred to constrain the vent location. Therefore, the method narrows the seven potential vents down to two, namely, Medicine Lake and Mount Mazama. The latter is the true vent for this deposit.

Two estimated vent sites near Lassen Peak (Fig. 17b) plus one unphysical prediction are obtained for the Rockland tephra. Thicknesses at the sample sites indicate that the tephra was widely dispersed downwind, but did not travel far in the upwind direction. Therefore, the source vent cannot be Mount Shasta, Medicine Lake, or the others farther north. The vent should be located near the two predicted locations shown in Fig. 17b. This inference is correct as the Rockland tephra was produced from the Rockland caldera complex, now underlying Brokeoff Volcano, a part of the Lassen Volcanic complex (Clynne and Muffler 2010; Pouget et al. 2014a).

We have estimated volumes of the two deposits based on the exponential model with the correct vent locations. We use volume estimation methods proposed by Pyle (1989), Fierstein and Nathenson (1992), and Nathenson and Fierstein (2014) to further process the isopach data. Deposit volumes of the THS and Rockland tephras are estimated to be 20.8 and 326.7km^3, respectively. The estimate for the THS tephra is consistent with the previous estimate (Pouget et al. 2014a), while the estimate for the Rockland tephra is beyond the range (50–248 km^3) found in previous studies (Sarna-Wojcicki et al. 1985; Pouget et al. 2014a). Given the methods used in previous studies, we believe that the current estimate of deposit volume represents a better, more objective, and unbiased value. These results will need to be evaluated carefully, especially for the Rockland tephra, as a volume of 326.7 km^3 suggests that it was an eruption of the same magnitude as that of the Bishop ash (Wilson and Hildreth 1997). Previous evaluation of the deposit suggested it was closer in volume to the Mazama ash (Clynne and Muffler 2010).

Application to Wilson Creek Formation ashes

We now apply the method, coupled with the exponential model, to the thickness dataset of the WCF Ash A4-d, to constrain its vent location. In another application, volumes of the WCF ashes B7-a, A4-d, and A3-f are estimated, using the exponential model with four to six observations. The measured thickness of Ash A4-d does not vary greatly from site to site, indicating that measurement uncertainty might affect the final estimate. Leave-one-out validation is used to examine whether the corresponding predictions are consistent. All but one of the predictions (unphysical prediction and the corresponding left-out observation marked in light green; dispersal axis not plotted) are consistent in both vent location and dispersal axis. The predicted dispersal axes suggest that the source vent for this sub-unit might be located near Dome 11. The inferred vent area for Ash A4-d is consistent with a previous, very poorly constrained and non-objective interpretation based on stratigraphic characteristics (Yang et al. 2019).

Volumes of ashes B7-a, A4-d, and A3-f are estimated following the same procedure done for the THS and Rockland tephras. Different potential vent locations (Marcaida et al. 2014 and this study) are assumed for these sub-units. Estimated volumes (Table 4) suggest that given the current

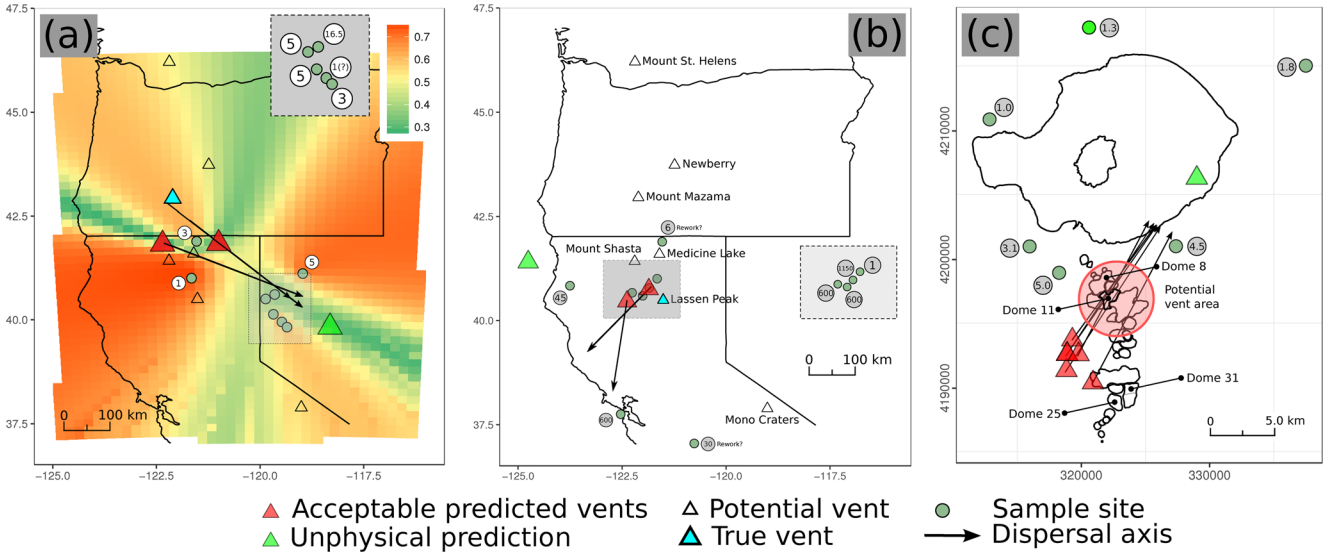


Fig. 17 Estimated vent locations (red triangle) and dispersal axes (black arrow lines) for the THS (**a**) and Rockland (**b**) tephras and the WCF Ash A4-d (**c**) using the exponential model. Thickness (cm) of the deposits at each sample site is labeled. Unphysical predictions are marked as green triangle. In **a** and **b**, potential vents (following Pouget et al. 2014a) for the two deposits are plotted as void triangle, and the true vents are marked in light blue. For the THS tephra, the surface of the cost function is plotted, and two predictions plus one unphysical prediction (light green triangle) are obtained. For the Rockland tephra, two estimates plus one unphysical prediction are derived from including and excluding two possibly reworked observations. For the WCF Ash A4-d (**c**), results from using the complete dataset (predicted vent location marked with darker red triangle) and subsets from leave-one-out sampling scheme as input are shown. There is one unphysical prediction that is derived from excluding the observation (marked in light green) in the north of Mono Lake. We draw a red circle to denote the inferred vent area for the deposit based on estimated dispersal axes. Locations of Domes 8, 11, 25, and 31 of the Mono Craters are labeled

Table 4 Estimated volumes of the THS and Rockland tephras and ashes B7-a, A4-d, and A3-f within the WCF

Tephra unit/sub-unit (dataset size)	Assumed vent location/eruption name	Estimated volume (dispersal direction)
THS (8)	Mt. Mazama	20.8
Rockland (8)	Rockland caldera complex	326.7
B7-a (4)	MC Dome 8	0.19 (243.2)
	MC Dome 31	0.12 (261.5)
	MC Dome 25	0.092 (289.6)
A4-d (6)	MC Dome 8	0.041 (48.1)
	MC Dome 11	0.041 (47.6)
	MC Dome 31	0.034 (35.3)
	MC Dome 25	0.031 (26.7)
A3-f (5)	MC Dome 8	0.017 (339.9)
	MC Dome 11	0.017 (339.3)
	MC Dome 31	0.014 (341.7)
	MC Dome 25	0.015 (349.1)
Orange-brown beds	South Mono	0.016
Basal beds		0.0054
North Mono Bed 1	North Mono	0.042
South Deadman 1	1350 A.D. Inyo eruption	0.01
South Deadman 2		0.04
Lower obsidian flow		< 0.01
Upper obsidian flow		0.01–0.02
Glass Creek		0.1

Volumes of selected tephra deposits from the most recent eruptions from the Mono-Inyo Craters are listed for reference (Miller 1985; Sieh and Bursik 1986; Nawotniak and Bursik 2010; Bursik et al. 2014). Units for estimated volume and dispersal direction are km^3 and degree from north clockwise, respectively. MC is short for Mono Craters

sample sites, volumes of these tephra layers are not significantly affected by the assumed vent location. The volume of Ash B7-a is ~ 0.092–0.19 km^3, while the others have estimated volumes below 0.05 km^3. Ash B7-a is thus the most voluminous single fall deposit yet known from the Mono Craters. Volume ranges for the other deposits are consistent with volumes of Holocene fall deposits from the Mono-Inyo Craters (Table 4, Miller 1985; Sieh and Bursik 1986; Nawotniak and Bursik 2010; Bursik et al. 2014). We believe that the crude estimates on vent location and volume of the listed WCF tephra sub-units can be used to improve our understanding on the volcanic history of the Mono Craters (Bailey 2004; Hildreth 2004; Marcaida 2015; Marcaida et al. 2019; Yang et al. 2019). These estimates provide new constraints for the volcanic hazard assessment of the Long Valley-Mono Craters region (Bevilacqua et al. 2017, 2018).

Conclusions

We have presented a new algorithm that can be used to estimate the source vent location and dispersal axis of tephra fall deposits based on thickness or maximum clast size measurements. The method is composed of a semi-empirical model that describes the thickness or maximum clast size distribution of tephra deposits, coupled with a gradient-descent method. There are two semi-empirical models (Gonzalez-Mellado and Cruz-Reyna 2010; Yang and Bursik 2016) that can be used. The method is validated on datasets of the North Mono Beds 1 and 2 and the Fogo A tephra deposit. The results show that the method can be used to estimate the vent location of tephra fall deposits or estimate the dispersal axis to constrain the vent location. With sufficient local input, the method can detect local changes in dispersal direction for tephra deposits. In addition, statistical bagging is shown to be another effective procedure in detecting epistemic uncertainty. Visualizing surface of the cost function is proposed to examine the occurrence of local minima given sparse dataset and is also another way of examining sources of uncertainty for the method.

The performance of the method is affected by the quality and size of the input dataset and the semi-empirical model being used. The flexibility of the power-law model makes the method more accurate given sufficient observations, but at the same time increases the instability of the method when working with sparse dataset. Our experiments show that, in the case of limited observations (≤ 10), estimating the dispersal axis first, instead of attempting to estimate the vent coordinates de novo, is a more robust way to constrain the vent location. This is because the planar symmetry of tephra thickness or maximum clast size distribution with respect to the dispersal axis, under the assumption of constant wind direction, is independent of many other eruptive and atmospheric parameters and can be easily detected and captured by the present method. Suggestions on how to use the method with dataset of varying sizes and qualities are given in Table 3.

The method assumes a constant wind direction, and the thickness and maximum clast size data are log-transformed prior to the fitting process. These assumptions are parsimonious, conforming with Occam's razor, which might introduce potential uncertainties to the final estimate; however, the method is shown to be useful and efficient and can be applied to tephra thickness or maximum clast size datasets of varying sizes and qualities. Due to the simplicity of the method, procedures (e.g., using localized subsets as input, statistical bagging, and cross-validation) proposed to characterize the uncertainty of the method focus on the detection of epistemic uncertainty. The flexibility of the method makes it easier to be incorporated into statistical systems and tephra databases to help identify the source vent location and correlate tephra deposits.

The method coupled with the exponential model is applied to sparse thickness datasets of the NMB1 (subset), Trego Hot Springs, and Rockland tephras, which provides useful constraints on their vent location. New estimates on their volume are given following the exponential model, which yields 20.8 and 326.7 km^3 (~ 8 and ~ 130 km^3 DRE), respectively. If these estimates are correct, then the Rockland tephra is among the most voluminous fall deposits of the late Quaternary in North America, formed 30–70 kyr after the caldera-forming event from the Yellowstone Volcano which produced the Lava Creek Tuff (Lanphere et al. 2002). Application of the method to tephra sub-units within the Wilson Creek Formation shows that Ash A4-d was erupted from a vent near Dome 11 of the Mono Craters and that at least one sub-unit of A4 was blown towards the northeast. From a transport pattern point of view, this suggests that the Lowder Creek ash could be Ash A4 (Madsen et al. 2002), as speculated by the original discoverers of that outcrop. The volume of Ash B7-a is estimated to be ~ 0.092–0.19 km^3, and estimated volumes of other analyzed sub-units are below 0.05 km^3. The use of the method provides new inferences about the eruptive history of the Mono Craters during the late Quaternary.

Acknowledgments S. Engwell is thanked for kindly sharing the data. We thank the reviewers for commenting on an earlier version of this manuscript.

Funding information This work was supported by National Science Foundation Hazard SEES grant number 1521855 to G. Valentine, M. Bursik, E.B. Pitman, and A.K. Patra and National Science Foundation Division of Mathematical Sciences grant number 1621853 to A.K. Patra, M. Bursik, and E.B. Pitman.

Compliance with ethical standards

Disclaimer The opinions expressed herein are those of the authors alone and do not reflect the opinion of the NSF.

References

Adams KD (2010) Lake levels and sedimentary environments during deposition of the Trego Hot Springs and Wono tephras in the lake Lahontan basin, Nevada, USA. Quat Res 73(1):118–129

Alloway B, Westgate J, Sandhu A, Bright R (1992) Isothermal plateau fission-track age and revised distribution of the widespread mid-Pleistocene Rockland tephra in west-central United States. Geophys Res Lett 19(6):569–572

Bailey RA (2004) Eruptive history and chemical evolution of the precaldera and postcaldera basalt-dacite sequences, Long Valley, California: implications for magma sources, current seismic unrest, and future volcanism. Number 1692. US Department of the Interior, US Geological Survey

Benson LV, Smoot JP, Kashgarian M, Sarna-Wojcicki A, Burdett JW (1997) Radiocarbon ages and environments of deposition of the Wono and Trego Hot Springs tephra layers in the Pyramid Lake subbasin, Nevada. Quat Res 47(3):251–260

Benson L, Liddicoat J, Smoot J, Sarna-Wojcicki A, Negrini R, Lund S (2003) Age of the Mono Lake excursion and associated tephra. Quat Sci Rev 22(2):135–140

Benson LV, Smoot JP, Lund SP, Mensing SA, Foit F Jr, Rye RO (2013) Insights from a synthesis of old and new climate-proxy data from the Pyramid and Winnemucca lake basins for the period 48 to 11.5 cal ka. Quat Int 310:62–82

Bevilacqua A, Bursik M, Patra A, Pitman EB, Till R (2017) Bayesian construction of a long-term vent opening probability map in the Long Valley volcanic region (CA, USA). Stat Volcanol 3(1):1–36

Bevilacqua A, Bursik M, Patra A, Bruce Pitman E, Yang Q, Sangani R, Kobs-Nawotniak S (2018) Late Quaternary eruption record and probability of future volcanic eruptions in the Long Valley volcanic region (CA, USA). J Geophys Res Solid Earth 123(7):5466–5494

Biass S, Bonadonna C (2011) A quantitative uncertainty assessment of eruptive parameters derived from tephra deposits: the example of two large eruptions of Cotopaxi volcano, Ecuador. Bull Volcanol 73(1):73–90

Biasse S, Bagheri G, Aeberhard W, Bonadonna C (2014) TError: towards a better quantification of the uncertainty propagated during the characterization of tephra deposits. Stat Volcanol 1(2):1–27

Bonadonna C, Costa A (2012) Estimating the volume of tephra deposits: a new simple strategy. Geology, pages G32769–1

Bonadonna C, Phillips JC (2003) Sedimentation from strong volcanic plumes. J Geophys Res Solid Earth 108(B7)

Bonadonna C, Ernst G, Sparks R (1998) Thickness variations and volume estimates of tephra fall deposits: the importance of particle Reynolds number. J Volcanol Geotherm Res 81(3):173–187

Bonadonna C, Mayberry G, Calder E, Sparks R, Choux C, Jackson P, Lejeune A, Loughlin S, Norton G, Rose W et al (2002) Tephra fallout in the eruption of Soufrière Hills Volcano, Montserrat. Geol Soc Lond Mem 21(1):483–516

Bonadonna C, Connor CB, Houghton B, Connor L, Byrne M, Laing A, Hincks T (2005) Probabilistic modeling of tephra dispersal: hazard assessment of a multiphase rhyolitic eruption at Tarawera, New Zealand. J Geophys Res Solid Earth 110(B3)

Bonadonna C, Biass S, Costa A (2015) Physical characterization of explosive volcanic eruptions based on tephra deposits: propagation of uncertainties and sensitivity analysis. J Volcanol Geotherm Res 296: 80–100

Bonasia R, Macedonio G, Costa A, Mele D, Sulpizio R (2010) Numerical inversion and analysis of tephra fallout deposits from the 472 AD sub-Plinian eruption at Vesuvius (Italy) through a new best-fit procedure. J Volcanol Geotherm Res 189(3):238–246

Bowers R (2009) Quaternary stratigraphy, geomorphology, and hydrologic history of pluvial Lake Madeline, Lassen County, northeastern California. PhD thesis, Humboldt State University

Breiman L (1996) Bagging predictors. Mach Learn 24(2):123–140

Burden R, Phillips J, Hincks T (2011) Estimating volcanic plume heights from depositional clast size. J Geophys Res Solid Earth 116(B11)

Burden R, Chen L, Phillips J (2013) A statistical method for determining the volume of volcanic fall deposits. Bull Volcanol 75(6):707

Bursik M (1993) Subplinian eruption mechanisms inferred from volatile and clast dispersal data. J Volcanol Geotherm Res 57(1):57–70

Bursik M (2001) Effect of wind on the rise height of volcanic plumes. Geophys Res Lett 28(18):3621–3624

Bursik M, Rogova G (2006) Use of neural networks and decision fusion for lithostratigraphic correlation with sparse data, Mono-Inyo Craters, California. Comput Geosci 32(10):1564–1572

Bursik M, Sieh K (1989) Range front faulting and volcanism in the Mono Basin, eastern California. J Geophys Res Solid Earth 94(B11): 15587–15609

Bursik M, Sieh K (2013) Digital database of the Holocene tephras of the Mono-Inyo Craters, California. Technical report, US Geological Survey

Bursik M, Carey S, Sparks R (1992a) A gravity current model for the May 18, 1980 Mount St. Helens plume. Geophys Res Lett 19(16): 1663–1666

Bursik M, Sparks R, Gilbert J, Carey S (1992b) Sedimentation of tephra by volcanic plumes: I. theory and its comparison with a study of the Fogo A plinian deposit, São Miguel (Azores). Bull Volcanol 54(4): 329–344

Bursik M, Sieh K, Meltzner A (2014) Deposits of the most recent eruption in the Southern Mono Craters, California: description, interpretation and implications for regional marker tephras. J Volcanol Geotherm Res 275:114–131

Carey S, Sparks R (1986) Quantitative models of the fallout and dispersal of tephra from volcanic eruption columns. Bull Volcanol 48(2–3): 109–125

Clynne M (1984) Stratigraphy and major element geochemistry of the Lassen volcanic center, California. Technical report, Geological Survey, Menlo Park, CA (USA)

Clynne MA, Muffler LJP (2010) Geologic map of Lassen Volcanic National Park and vicinity, California. US Department of the Interior, US Geological Survey

Connor L, Connor CB (2006) Inversion is the key to dispersion: understanding eruption dynamics by inverting tephra fallout. Statistics in volcanology. Geological Society, London, pp 231–242

Costa A, Folch A, Macedonio G (2013) Density-driven transport in the umbrella region of volcanic clouds: implications for tephra dispersion models. Geophys Res Lett 40(18):4823–4827

Csanady GT (1973) The fluctuation problem in turbulent diffusion. In: Turbulent diffusion in the environment. Springer, pp. 10–26

Davis JO (1978) Quaternary tephrochronology of the Lake Lahontan area, Nevada and California. Doctoral dissertation, University of Nevada

Engwell S, Sparks R, Aspinall W (2013) Quantifying uncertainties in the measurement of tephra fall thickness. J Appl Volcanol 2(1):5

Engwell S, Aspinall W, Sparks R (2015) An objective method for the production of isopach maps and implications for the estimation of tephra deposit volumes and their uncertainties. Bull Volcanol 77(7): 1–18

Fierstein J, Nathenson M (1992) Another look at the calculation of fallout tephra volumes. Bull Volcanol 54(2):156–167

Gonzalez-Mellado A, Cruz-Reyna S (2010) A simple semi-empirical approach to model thickness of ash-deposits for different eruption scenarios. Nat Hazards Earth Syst Sci 10(11):2241–2257

Green RM, Bebbington MS, Cronin SJ, Jones G (2014) Automated statistical matching of multiple tephra records exemplified using five long maar sequences younger than 75ka, Auckland, New Zealand. Quat Res 82(2):405–419

Green RM, Bebbington MS, Jones G, Cronin SJ, Turner MB (2016) Estimation of tephra volumes from sparse and incompletely observed deposit thicknesses. Bull Volcanol 78(4):1–18

Hildreth W (2004) Volcanological perspectives on Long Valley, Mammoth Mountain, and Mono Craters: several contiguous but discrete systems. J Volcanol Geotherm Res 136(3):169–198

Kawabata E, Bebbington MS, Cronin SJ, Wang T (2013) Modeling thickness variability in tephra deposition. Bull Volcanol 75(8):1–14

Kawabata E, Cronin S, Bebbington M, Moufti M, El-Masry N, Wang T (2015) Identifying multiple eruption phases from a compound tephra blanket: an example of the AD 1256 AI-Madinah eruption, Saudi Arabia. Bull Volcanol 77(1):1–13

Kawabata E, Bebbington MS, Cronin SJ, Wang T (2016) Optimal likelihood-based matching of volcanic sources and deposits in the Auckland volcanic field. J Volcanol Geotherm Res 323:194–208

King M, Busacca AJ, Foit FF, Kemp RA (2001) Identification of disseminated Trego Hot Springs tephra in the Palouse, Washington State. Quat Res 56(2):165–169

Klawonn M, Wolfe CJ, Frazer LN, Houghton BF (2012) Novel inversion approach to constrain plume sedimentation from tephra deposit data: application to the 17 June 1996 eruption of Ruapehu volcano, New Zealand. J Geophys Res Solid Earth 117(B5)

Klawonn M, Houghton BF, Swanson DA, Fagents SA, Wessel P, Wolfe CJ (2014) Constraining explosive volcanism: subjective choices during estimates of eruption magnitude. Bull Volcanol 76(2):793

Koyaguchi T (1994) Grain-size variation of tephra derived from volcanic umbrella clouds. Bull Volcanol 56(1):1–9

Lajoie KR (1968) Late Quaternary stratigraphy and geologic history of Mono Basin, eastern California. University of California

Lanphere MA, Champion DE, Clynne MA, Patrick Muffler L (1999) Revised age of the Rockland tephra, northern California: implications for climate and stratigraphic reconstructions in the western United States. Geology 27(2):135–138

Lanphere MA, Champion DE, Christiansen RL, Izett GA, Obradovich JD (2002) Revised ages for tuffs of the Yellowstone Plateau volcanic field: assignment of the Huckleberry Ridge Tuff to a new geomagnetic polarity event. Geol Soc Am Bull 114(5):559–568

Lanphere M, Champion D, Clynne M, Lowenstern J, Sarna-Wojcicki A, Wooden J (2004) Age of the Rockland tephra, western USA. Quat Res 62(1):94–104

Madsen DB, Sarna-Wojcicki AM, Thompson RS (2002) A late Pleistocene tephra layer in the southern Great Basin and Colorado Plateau derived from Mono Craters, California. Quat Res 57(3): 382–390

Marcaida M (2015) Resolving the timing of late Pleistocene dome emplacement at Mono Craters, California, from ^{238}U–^{230}Th and $^{40}Ar/^{39}Ar$ dating. San Jose State University

Marcaida M, Mangan MT, Vazquez JA, Bursik M, Lidzbarski MI (2014) Geochemical fingerprinting of Wilson Creek formation tephra layers (Mono Basin, California) using titanomagnetite compositions. J Volcanol Geotherm Res 273:1–14

Marcaida M, Vazquez JA, Stelten ME, Miller JS (2019) Constraining the Early Eruptive History of the Mono Craters Rhyolites, California, Based on U- Th Isochron Dating of Their Explosive and Effusive Products. Geochemistry, Geophysics, Geosystems 20 (3):1539–1556

Miller CD (1985) Holocene eruptions at the Inyo volcanic chain, California: implications for possible eruptions in Long Valley caldera. Geology 13(1):14–17

Moore RB (1990) Volcanic geology and eruption frequency, São Miguel, Azores. Bull Volcanol 52(8):602–614

Nathenson M, Fierstein J (2014) Spread sheet to calculate tephra volume for exponential thinning. https://vhub.org/resources/3716

Nawotniak SK, Bursik M (2010) Subplinian fall deposits of Inyo Craters, CA. J Volcanol Geotherm Res 198(3):433–446

Negrini RM, Verosub KL, Davis JO (1988) The middle to late Pleistocene geomagnetic field recorded in fine-grained sediments from Summer Lake, Oregon, and Double Hot Springs, Nevada, USA. Earth Planet Sci Lett 87(1–2):173–192

Pouget S, Bursik M, Cortes JA, Hayward C (2014a) Use of principal component analysis for identification of Rockland and Trego Hot Springs tephras in the Hat Creek Graben, northeastern California, USA. Quat Res 81(1):125–137

Pouget S, Bursik M, Rogova G (2014b) Tephra redeposition and mixing in a late-glacial hillside basin determined by fusion of clustering analyses of glass-shard geochemistry. J Quat Sci 29(8):789–802

Pouget S, Bursik M, Singla P, Singh T (2016) Sensitivity analysis of a one-dimensional model of a volcanic plume with particle fallout and collapse behavior. J Volcanol Geotherm Res 326:43–53

Pyle DM (1989) The thickness, volume and grainsize of tephra fall deposits. Bull Volcanol 51(1):1–15

R Core Team (2017) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna

Rhoades DA, Dowrick D, Wilson C (2002) Volcanic hazard in New Zealand: scaling and attenuation relations for tephra fall deposits from Taupo Volcano. Nat Hazards 26(2):147–174

Rieck HJ, Sarna-Wojcicki AM, Meyer CE, Adam DP (1992) Magnetostratigraphy and tephrochronology of an upper Pliocene to Holocene record in lake sediments at Tulelake, northern California. Geol Soc Am Bull 104(4):409–428

Rogova GL, Bursik MI, Hanson-Hedgecock S (2007) Interpreting the pattern of volcano eruptions: intelligent system for tephra layer correlation. In: Information fusion, 2007 10th international conference on, pages 1–7. IEEE

Rogova GL, Bursik MI, Pouget S (2015) Learning under uncertainty for interpreting the pattern of volcanic eruptions. In Information Fusion (Fusion), 2015 18th International Conference on, pages 375–382. IEEE

Sarna-Wojcicki A (2000) Revised age of the Rockland tephra, northern California: implications for climate and stratigraphic reconstructions in the western United States: comment: COMMENT. Geology 28(3):286–286

Sarna-Wojcicki AM, Meyer CE, Bowman HR, Hall NT, Russell PC, Woodward MJ, Slate JL (1985) Correlation of the Rockland ash bed, a 400,000-year-old stratigraphic marker in northern California and western Nevada, and implications for middle Pleistocene paleogeography of central California. Quat Res 23(2):236–257

Sarna-Wojcicki A, Morrison SD, Meyer C, Hillhouse J (1987) Correlation of upper Cenozoic tephra layers between sediments of the western United States and eastern Pacific Ocean and comparison with biostratigraphic and magnetostratigraphic age data. Geol Soc Am Bull 98(2):207–223

Sarna-Wojcicki AM, Davis JO, Morrison R (1991) Quaternary tephrochronology, Quaternary nonglacial geology: Conterminous United States: Boulder, Colorado, Geological Society of America. Geol North Am 2:93–116

Schwaiger HF, Denlinger RP, Mastin LG (2012) Ash3d: a finite-volume, conservative numerical model for ash transport and tephra deposition. J Geophys Res Solid Earth 117(B4)

Scollo S, Tarantola S, Bonadonna C, Coltelli M, Saltelli A (2008) Sensitivity analysis and uncertainty estimation for tephra dispersal models. J Geophys Res Solid Earth 113(B6)

Self S (1983) Large-scale phreatomagmatic silicic volcanism: a case study from New Zealand. J Volcanol Geotherm Res 17(1–4):433–469

Sieh K, Bursik M (1986) Most recent eruption of the Mono Craters, eastern central California. J Geophys Res Solid Earth 91(B12): 12539–12571

Sigl M, Winstrup M, McConnell J, Welten K, Plunkett G, Ludlow F, Buntgen U, Caffee M, Chellman N, Dahl-Jensen D et al (2015)

Timing and climate forcing of volcanic eruptions for the past 2,500 years. Nature 523(7562):543–549

Sparks R, Carey S, Sigurdsson H (1991) Sedimentation from gravity currents generated by turbulent plumes. Sedimentology 38(5):839–856

Sparks R, Bursik M, Ablay G, Thomas R, Carey S (1992) Sedimentation of tephra by volcanic plumes. Part 2: controls on thickness and grain-size variations of tephra fall deposits. Bull Volcanol 54(8): 685–695

Suzuki T (1983) A theoretical model for dispersion of tephra. Arc Volcanism: Physics and Tectonics 95:113

Turner MB, Bebbington MS, Cronin SJ, Stewart RB (2009) Merging eruption datasets: building an integrated Holocene eruptive record for Mt Taranaki, New Zealand. Bull Volcanol 71(8):903–918

Volentik AC, Bonadonna C, Connor CB, Connor LJ, Rosi M (2010) Modeling tephra dispersal in absence of wind: insights from the climactic phase of the 2450 BP Plinian eruption of Pululagua volcano (Ecuador). J Volcanol Geotherm Res 193(1):117–136

Walker G, Croasdale R (1971a) Characteristics of some basaltic pyroclastics. Bull Volcanol 35(2):303–317

Walker G, Croasdale R (1971b) Two Plinian-type eruptions in the Azores. J Geol Soc 127(1):17–55

White J, Connor C, Connor L, Hasenaka T (2017) Efficient inversion and uncertainty quantification of a tephra fallout model. J Geophys Res Solid Earth 122(1):281–294

Wilson CJ, Hildreth W (1997) The Bishop Tuff: new insights from eruptive stratigraphy. J Geol 105(4):407–440

Wood SH (1977) Distribution, correlation, and radiocarbon dating of late Holocene tephra, Mono and Inyo craters, eastern California. Geol Soc Am Bull 88(1):89–95

Yang Q, Bursik M (2016) A new interpolation method to model thickness, isopachs, extent, and volume of tephra fall deposits. Bull Volcanol 78(10):68

Yang Q, Bursik MI, Pitman EB (2018) A new method to identify the vent location of tephra fall deposits based on thickness or maximum clast size measurements (SVL). https://vhub.org/resources/4377

Yang Q, Bursik M, Pouget S (2019) Stratigraphic and sedimentologic framework for tephras in the Wilson Creek Formation, Mono Basin, California, USA. J Volcanol Geotherm Res