Journal of Applied Volcanology
a SpringerOpen Journal

RESEARCH   **Open Access**

AshCalc–a new tool for the comparison of the exponential, power-law and Weibull models of tephra deposition

Matthew L Daggitt[1*], Tamsin A Mather[2*], David M Pyle[2] and Stephen Page[3]

Abstract

This paper presents a new tool, AshCalc, for the comparison of the three most commonly used models for the calculation of the bulk volume of volcanic tephra fall deposits: the exponential model, the power law model and the Weibull model. AshCalc provides a simple and intuitive tool to speed up the analysis of tephra deposits and compare and contrast the fits for each model. Two improvements in terms of computational performance are implemented in AshCalc for the estimation of the parameters for the Weibull model. The first is an analytic method for reducing the number of free parameters, whilst the second exaggerates the minima in parameter space, leading to a more robust solution. We show that AshCalc provides volume estimates in line with other previously published estimates and hence can be used with a high degree of confidence. We include the open source python code for Ashcalc with the intention that it can be used both as a stand-alone program and integrated into other python projects.

Keywords: AshCalc; Tephra volumes; Exponential; Power-law; Weibull

Introduction

Studying the variation in thickness of tephra fall deposits has long been used as a key tool for estimating the volumes of explosive volcanic eruptions (Pyle, 1989). However, as yet, there is no agreement on the most appropriate model that can be integrated to yield a total bulk volume on the basis of sparse field measurements. Most approaches currently fit isopachs (contours of equal thickness) to field measurements of deposit thickness, and then use the thickness–area relationships to derive a volume. Using a symmetrical distribution assumption, the relationship between the volume of ejecta eruption, V, and the square root of isopach area, x, is as follows:

$$V = \int_0^\infty T(x)dA = 2\int_0^\infty xT(x)dx$$

Various models have been proposed for thickness as a function of isopach area, T(x). The most widely used to date are the multi-segment exponential model (Pyle, 1989;

1990; 1995), the power-law model (Bonadonna et al. 1998; Bonadonna and Houghton, 2005) and more recently a Weibull distribution has been proposed as an alternative model (Bonadonna & Costa, 2012; 2013). Each model has its own advantages and drawbacks and their suitability may depend on the characteristics of the deposit. Despite the multiple models, there is, as yet, no routine way to facilitate comparisons between the different approaches. For this reason, we have developed AshCalc, which is designed to do just this.

An alternative approach to volume estimation has recently been proposed by Burden et al. (2013), which uses a purely statistical method to determine the volume of fall deposits without relying on the construction of isopach maps. This is certainly a viable and complementary approach, but it is not one that we shall explore further. As Burden et al. (2013) showed, the results of their approach are consistent with those derived from isopach-based models.

Introduction to ash volume estimate methodologies
Exponential model
The exponential model (Pyle, 1989) states that the relationship between that thickness, T, is a function of root isopach area, x, as follows:

$$T(x) = ce^{-mx}$$

* Correspondence: mld46@cam.ac.uk; Tamsin.Mather@earth.ox.ac.uk
[1]Jesus College, University of Cambridge, Cambridge CB5 8BL, UK
[2]Department of Earth Sciences, University of Oxford, South Parks Road, Oxford OX1 3AN, UK
Full list of author information is available at the end of the article

Springer

where c represents the theoretical maximum thickness at the vent and m characterises the rate of decrease in tephra thickness. Taking the logarithm of both sides linearises the equation and the parameters, c and m, are found by applying least squares regression.

By fitting multiple exponential segments to the data, the exponential law is capable of modelling deposits which have a variable rate of thinning with distance away from the vent–a feature which is common to many well constrained deposits (e.g. Pyle, 1989; 1990; Fierstein and Nathanson, 1992; Bonadonna and Houghton 2005; Watt et al. 2009). This approach lends itself to a variety of simple approximations, for example to determine the minimum volume of a deposit from sparse information (e.g. Pyle, 1995; 1999; Legros, 2000; Sulpizio, 2005).

Power law model

The power law model (Bonadonna et al. 1998; Bonadonna & Houghton, 2005) states that the relationship between thickness, T, is a function of root isopach area, x, as follows:

$$T(x) = cx^{-m}$$

Where c is a linear scaling factor and m characterises the rate of decrease in tephra thickness. Taking the log of both sides linearises the equation and the parameters, c and m, are found by applying least squares regression.

The main disadvantage of the power law model is that T(x) is not integrable between 0 and ∞ and therefore proximal and distal limits of integration have to be chosen (e.g., Bonadonna and Costa 2012).

Weibull model

The Weibull model was proposed more recently than the exponential and power law models with the aim of combining the advantage of the exponential model (being integrable between 0 and ∞) and the power law model (variable rate of decrease) (Bonadonna & Costa, 2012). In the model thickness is related to x as follows:

$$T(x) = \theta\left(\frac{x}{\lambda}\right)^{k-2} e^{\left(\frac{x}{\lambda}\right)^k}$$

The additional parameter, λ, allows the Weibull model to capture the variation in thinning rate that necessitates multiple segments in the exponential model.

Methods

The aim of AshCalc is to be an easy to use tool distributed under an open source license for volcanologists to compare the suitability of the three models for their data. AshCalc consists of a graphical user interface (Figure 1), as well as the python source code for the calculations which is easily extensible and able to be integrated into other projects. AshCalc's source code is provided in Additional file 1

(Additional file 2 is the Ashcalc User Manual and source code bundled together in case more efficient for distribution).

General inputs and outputs

As input, AshCalc takes a list of thickness(m), area(km^2) pairs for each isopach as well as additional model specific parameters (see Sections 2.2-2.4 below). For all models, AshCalc displays a plot of thickness against $\sqrt{\text{Area}}$, the total estimated tephra deposit, the estimates for the model parameters, the mean relative squared error in the fit of the model to the data and the equation for T(x). In addition to generating optimal parameters it allows the user to experiment with their own values for the parameters, updating the graphs and volume and error estimates accordingly. The user manual, containing installation instructions, is provided as Additional file 3. Specific details relating to running AshCalc for each model are highlighted below.

Exponential model

AshCalc uses the scientific python library, SciPy, to perform multiple-segment least squares regression to determine the segment parameters. When more than a single segment is used, AshCalc chooses the segment boundaries as to ensure the continuity of the thickness function where possible.

Model specific inputs: The number of exponential segments

Model specific outputs: A graph of the least squares regression fit, as well as individual segment volumes, equations and estimates.

Power law model

AshCalc performs simple least squares regression to determine the power law parameters.

Model specific inputs: Proximal and distal limits of integration

Model specific outputs: Graph of the least squares regression fit, graph of the error surface for parameters c and m, an estimate of the proximal bound as described in Bonadonna and Houghton (2005).

Weibull model

The Weibull model is a three parameter model and hence, unlike the exponential and power law distributions, optimal values for λ, k and θ cannot be determined analytically. Instead an optimisation algorithm must explore the parameter space and come up with a suitable estimation of their true values. Optimisation algorithms work by trying to minimise a function of the parameters. The function should be chosen such that the minimal parameters result in the distribution with the best fit to the data.

Decreasing the number of parameters in an optimisation problem greatly reduces the size of the parameter space needing to be searched. This significantly increases
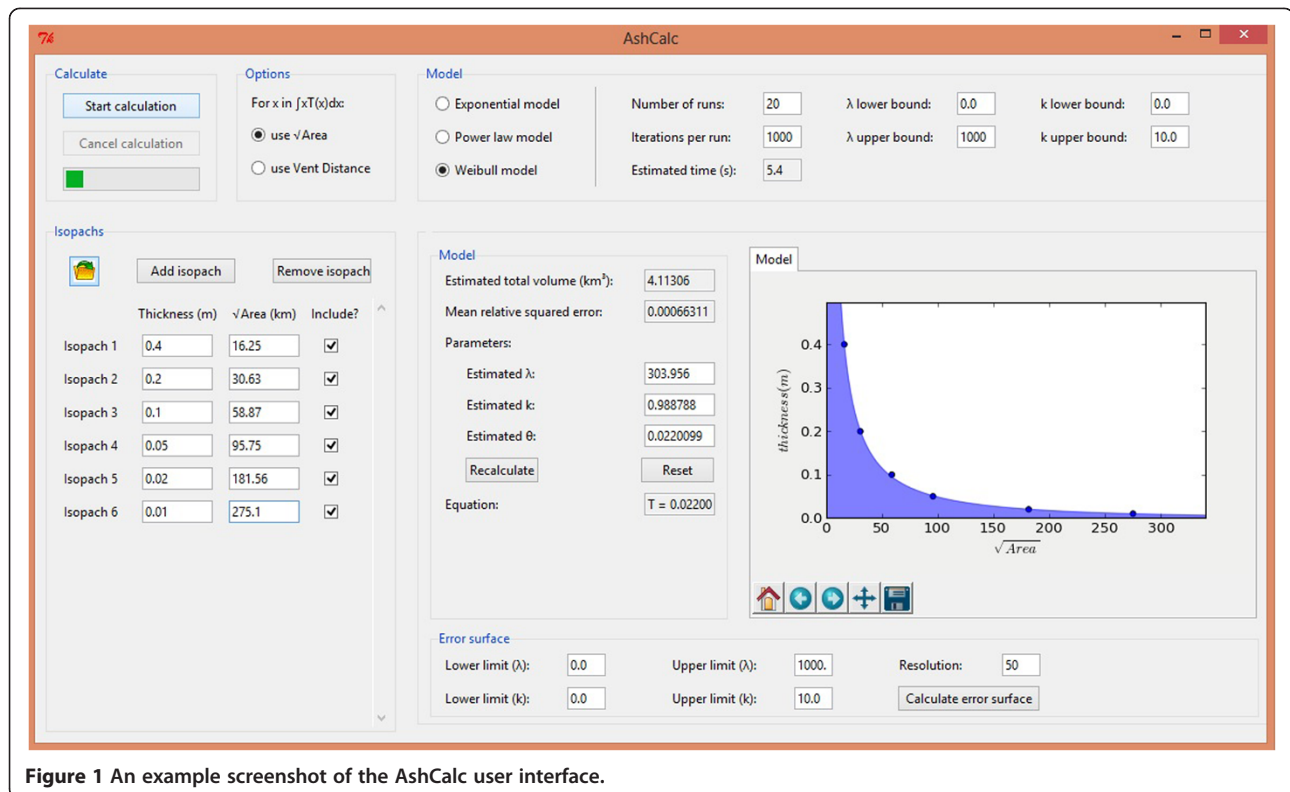
Figure 1 An example screenshot of the AshCalc user interface.

the performance of the optimisation algorithm when finding solutions. The magnitude of the derivatives of the function (gradient) around the minima also affects the performance of algorithm in finding minima.

Bonadonna and Costa (2013) chose the relative squared error (RSE) as a function of λ, k and θ as the function to minimise.

$$RSE(\lambda, k, \theta) = \sum_{x \in data} \left(\left(\frac{T_{observed}(x) - T(x)}{T_{observed}(x)} \right)^2 \right) \quad (1)$$

where

$$T(x) = \theta \left(\frac{x}{\lambda} \right)^{k-2} e^{-\left(\frac{x}{\lambda} \right)^k}$$

The RSE is greater than or equal to zero for all parameter values, with zero representing a fit that goes through every data point perfectly.

Two possible improvements to using RSE(λ,k,θ) as the minimising function, with the aim of increasing the performance of the optimisation algorithms, are proposed here:

1) Eliminating theta by deriving an optimal value for it in terms of λ and k and the given isopach data. This turns the three parameter function RSE(λ,k,θ) into the two parameter function RSE(λ,k), resulting in increased performance of the optimisation algorithm.

2) Minimising a new quantity Error(λ, k) = RSE(λ, k)+ ln(RSE(λ, k)) as alternative to *RSE*(λ,k). The new function has much larger gradients around the minimum of *RSE*(λ,k), improving the performance of the optimisation algorithms.

In this instance the performance of an algorithm is measured by how many computation steps it has to perform to get to within some pre-defined range around the optimal parameters. As it is measured in computation steps rather than computation time this measure is independent of the hardware the algorithm is run on.

Eliminating theta
Eq 1 is a quadratic in θ and hence has a single analytically derivable minimum with respect to θ.

Differentiating Eq 1 with respect to θ yields an equation linear in θ. Setting the derivative to zero and solving for θ arrives at an expression for the value that minimises Eq 1 for any given data set and values of λ and k:

$$\theta(\lambda, k) = \lambda^{k-2} \sum_{x \in data} \frac{x^{k-2}}{T_{observed}(x)} e^{-\left(\frac{x}{\lambda} \right)^k}$$

$$\times \left(\sum_{x \in data} \left(\frac{x^{k-2}}{T_{observed}(x)} e^{-\left(\frac{x}{\lambda} \right)^k} \right)^2 \right)^{-1}$$
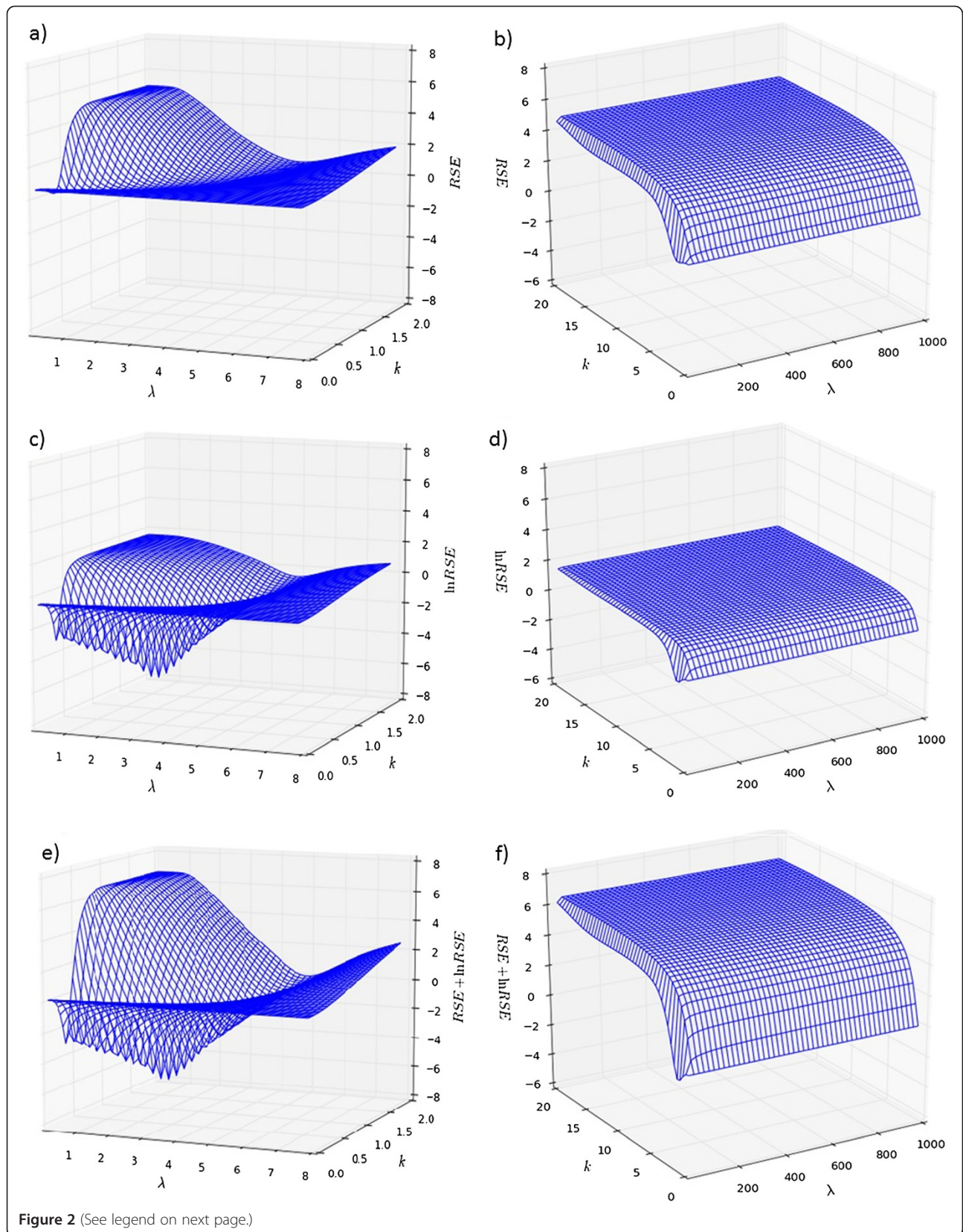
(2)

Figure 2 (See legend on next page.)

(See figure on previous page.)
Figure 2 Weibull model error surfaces produced by AshCalc for the 1992 Cerro Negro eruption. Graphs show different error functions (Relative Squared Error (RSE), lnRSE and RSE + lnRSE) plotted against λ and k over large and small ranges of the values. The minimum is in the region of $4.0 \leq \lambda \leq 4.4, 0.75 \leq k \leq 0.85$. **a)** RSE error function–close to the minimum. **b)** RSE error function–larger view of parameter space. **c)** ln (RSE) error function–close to the minimum. **d)** ln(RSE) error function–larger view of parameter space. **e)** ln(RSE) + RSE error function–close to the minimum. **f)** ln(RSE) + RSE error function–larger view of parameter space.

The full derivation of Eq 2 can be found in Additional file 4.

The elimination of θ as a free parameter from the model reduces it to a two parameter model. As well as increasing the performance of the optimisation algorithm this also allows the parameter space to be plotted in three dimensions. This allows a visual inspection of the parameter space by users, aiding in the process of finding appropriate bounds to be set for the parameters in the optimisation algorithm.

Emphasising the minima

The function RSE has the drawback that the denominator of each term, the observed thickness $T_{observed}(x)$ maps the squared errors smaller than the observed thickness into the range [0,1]. This mapping results in large flat basins around the minimum of the RSE function, decreasing the performance of optimisation algorithms.

When the 2-dimensional error function is plotted as a function of λ and k, empirical observations confirm that resulting surfaces tend to be flat around the minima (Figure 2a). In such flat basins many optimisation algorithms' convergence to the minima is slow.

Taking the natural logarithm of the RSE maps values in [0,1] to [-∞, 1], eliminates these flat basins (Figure 2c).

The logarithm function is a monotonically increasing function and therefore:

$$\text{RSE}(\lambda_1, k_1) \leq \text{RSE}(\lambda_2, k_2) \leftrightarrow \ln(\text{RSE}(\lambda_1, k_2)) \leq \ln(\text{RSE}(\lambda_1, k_2))$$

This means that if (λ_1, k_1) is a minimum for $\ln(\text{RSE}(\lambda_1, k_2))$ then it is also a minimum for $\text{RSE}(\lambda_1, k_1)$. Hence minimising $\ln(\text{RSE}(\lambda_1, k_2))$ is equivalent to minimising $\text{RSE}(\lambda_1, k_1)$.

An undesirable property of $\ln(\text{RSE}(\lambda_1, k_2))$ is that it flattens the error surface in the range $[1, \infty]$ (Figure 2d). Ideally, for $\text{RSE}(\lambda_1, k_1) \geq 1$ the properties of $\text{RSE}(\lambda_1, k_1)$ are desirable, whilst for $\text{RSE}(\lambda_1, k_1) \leq 1$ the properties of $\ln(\text{RSE}(\lambda_1, k_2))$ are desired.

Therefore it is proposed that a new error function is used for optimisation of parameters λ and k:

$$\text{Error}(\lambda, k) = \text{RSE}(\lambda, k) + \ln(\text{RSE}(\lambda, k)) \qquad (3)$$

This function preserves the ordering of all points on the original error surface as the logarithm is a monotonic function. Therefore minimising Eq 3 is equivalent to minimising Eq 1.

When RSE is in the range [0,1], the logarithmic term dominates Eq 3 mapping it to the range [-∞, 1] (Figure 2e) and when the RSE is in the range $[1, \infty]$ the non-
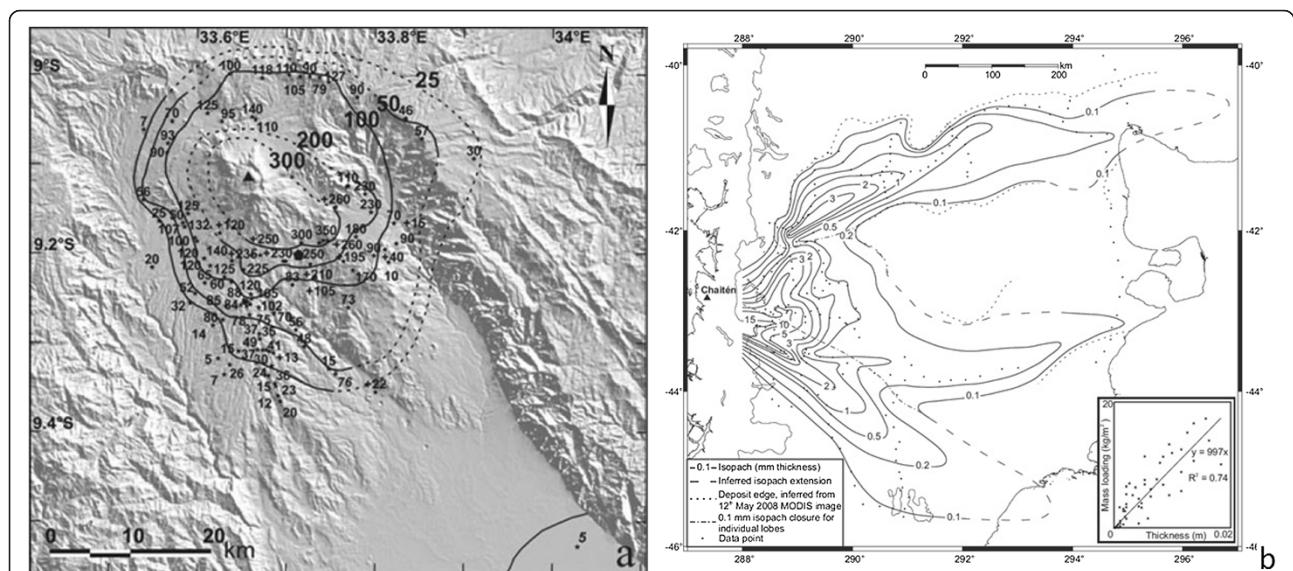

Figure 3 Isopach maps for a) Rungwe pumice eruption (figure adapted from Fontijn et al. 2011) and b) 2008 Chaitén eruption (figure adapted from Watt et al. 2009; Alfano et al. 2011).

Table 1 A comparison between values produced by AshCalc and previously published values for three eruptions of varying magnitude

Eruption	Source	Exponential (single segment) eruption volume (km^3)	Exponential (two segments) eruption volume (km^3)	Power law eruption volume (km^3)	Weibull eruption volume[a] (km^3)
Chaitén 2008–Whole deposit (isopachs closed)	Watt et al. 2009	-	0.17	-	-
	AshCalc	-	0.1713	0.2158[b]	0.2033
Chaitén 2008–Whole deposit (mapped area)	Watt et al. 2009	-	0.15	-	-
	AshCalc	-	0.1518	0.1548[b]	0.1566
Chaitén 2008–3rd of May lobe	Watt et al. 2009	-	0.05	-	-
	AshCalc	-	0.0478	0.0506[b]	0.0535
Chaitén 2008–6th of May lobe	Watt et al. 2009	-	0.03	0.04[b]	-
	AshCalc	-	0.0347	0.0420[b]	0.0361
Rungwe pumice eruption–Malawi (1 cm)	Fontijn et al. 2011	2.2[g]	3.2	3.3[c]	-
	AshCalc	2.145	3.236	3.251[c]	-
Rungwe pumice eruption–Malawi (3 cm)	Fontijn et al. 2011	-	-	5.8[d]	-
	AshCalc	-	4.301	5.818[d]	-
Rungwe pumice eruption–Malawi (5 cm)	Fontijn et al. 2011	-	5.6	8.5[e]	-
	AshCalc	-	5.559	8.493[e]	-
Minoan eruption	Pyle, 1990	-	38.5	-	-
	AshCalc	-	38.83	39.01[f]	42.34

[a]Due to the stochastic nature of the fit the values are taken from the medial volume of five successive calculations by AshCalc.
[b]Proximal limit of integration–5.0 km, distal limit of integration–250 km.
[c]Proximal limit of integration–8.6 km, distal limit of integration–1000 km.
[d]Proximal limit of integration–7.0 km, distal limit of integration–1000 km.
[e]Proximal limit of integration–6.1 km, distal limit of integration–1000 km.
[f]Proximal limit of integration–1.0 km, distal limit of integration–600 km.
[g]Mistake in published value, acknowledged by K. Fontijn in personal communication. Value should be 2.1 km^3.

logarithmic term dominates, preserving the greater magnitude of the derivatives (Figure 2f). This transformation reduces the size and flatness of any basins, hence increasing the probability that the optimisation algorithm will find better values for λ and k for a given number of iterations.

AshCalc and the Weibull model
AshCalc takes advantage of the two improvements above and uses a variant of random-restart stochastic hill-climbing for its optimisation algorithm. Within each run, the maximum amount each parameter can change decreases after each iteration.

Model specific inputs: Number of hill-climbing runs, number of iterations per hill-climbing run and upper and lower bounds for λ and k. Guidelines on what to choose for these inputs can be found in the user manual in Additional file 2.

Model specific outputs: Graph of the error surface for parameters λ and k.

Results
In order to assess the accuracy of AshCalc, the program's outputs were compared to published values for four eruptions. The four were chosen with the aim of being spread over a range of eruptive magnitudes and deposit shapes and, in order of eruptive magnitude, they are the 1992 Cerro Negro eruption (Bonadonna & Costa, 2012), the 2008 Chaitén eruption (Watt et al. 2009), the Rungwe pumice eruption (Fontijn et al. 2011) and the Minoan eruption (Pyle, 1990).

The Rungwe pumice eruption was chosen as an example of an elliptical deposit while the 2008 Chaitén eruption is an example of an irregularly shaped deposit heavily affected by wind direction. The isopach maps for the two eruptions are shown in Figure 3 with the uncertain Lake Malawi data point for the Rungwe eruption visible in the bottom right hand corner of Figure 3a.

Table 2 A comparison between values produced by AshCalc and previously published values for the 1992 Cerro Negro eruption (Bonadonna & Costa, 2012)

Resulting Weibull fit	Bonadonna & Costa., 2008	AshCalc[a]
Volume (km^3)	0.045	0.04556
Λ	4.2	4.131
k	0.78	0.7667
Θ	1.0004	1.02325
Relative mean squared error	0.000541037	0.00050026

[a]Due to the stochastic nature of the fit the values are taken from the medial volume of five successive calculations by AshCalc.

The results for the comparisons with the 2008 Chaitén eruption, the Rungwe pumice eruption and the Minoan eruption are presented in Table 1.

AshCalc provides values equal, within the number of significant figures provided, to all values calculated using the exponential model by Watt et al. (2009) for Chaitén. Additionally AshCalc's Weibull and power law fits provide reasonable volumes compared to the exponential estimates by Watt et al. (2009).

For the Rungwe pumice eruption, AshCalc again provides volumes equal to the published values of all fits, within the number of published significant figures. The true value of the single segment exponential fit to 2 s.f. has been acknowledged by K. Fontijn to be 2.1 km^3 rather than the 2.2 km^3 published (pers. comm.).

For the Minoan eruption, AshCalc calculates a two segment exponential volume within 1% of the original value.

Table 2 shows the comparison between AshCalc's Weibull model and the published Weibull model (Bonadonna & Costa, 2012) for the 1992 Cerro Negro eruption. As the process of determining the Weibull parameters is inherently random it is unlikely to generate an exact match for the published model. None-the-less all three parameters are within 3% of those previously published and the relative MSE (mean squared error, the measure used by Bonadonna & Costa, 2012) of the fit provided by AshCalc is smaller than that of the fit of the published model. This means that, as measured by the relative MSE, in some sense AshCalc has provided a "better" fit for the data.

In this paper, we have focussed on the analysis of deposit thickness as a tool for determining erupted volume as deposit thicknesses are most readily and commonly measured in the field. An exactly similar approach applies for the analysis of deposit mass per unit area as a way of determining total erupted mass, and the programme can readily be adapted for this purpose by multiplying the final calculated volume by the estimated mean density.

Conclusions

AshCalc has the potential to be a powerful tool for volcanologists to calculate and compare the various models (exponential, power law and Weibull) for tephra deposits as matches for their data and calculation of volumes. It has been shown that AshCalc provides volume estimates in line with other previously published estimates and hence can be used with a high degree of confidence. We implement two improvements in terms of computational performance for the estimation of the parameters for the Weibull model in AshCalc. The first is an analytic method for reducing the number of free parameters and the second exaggerates the minima in parameter space, leading to a more robust solution. As both a stand-alone program and a series of python modules that can be integrated into other python projects, it is hoped that AshCalc will become an important tool in the volcanology community.

Additional files

Additional file 1: AshCalc source code.

Additional file 2: AshCalc user manual and source code bundled together for efficient distribution.

Additional file 3: AshCalc: User manual.

Additional file 4: Calculation for the elimination of θ.

Authors' contributions
MD performed the coding associated with the project with additional mathematical inputs from SP. DP and TM formulated the project brief and assisted MD with the project's context and the drafting of the manuscript. All authors read and approved the final manuscript.

Author details
[1]Jesus College, University of Cambridge, Cambridge CB5 8BL, UK.
[2]Department of Earth Sciences, University of Oxford, South Parks Road, Oxford OX1 3AN, UK. [3]St Catharine's College, University of Cambridge, Cambridge CB2 1RL, UK.

References
Alfano F, Bonadonna C, Volentik ACM, Connor CB, Watt SFL, Pyle DM, Connor LJ (2011) Tephra stratigraphy and eruptive volume of the May, 2008, Chaitén eruption, Chile. Bull Volcanol 73:613–630

Bonadonna C, Costa A (2012) Estimating the volume of tephra deposits: a new simple strategy. Geology 40:415–418

Bonadonna C, Costa A (2013) Plume height, volume, and classification of explosive volcanic eruptions based on the Weibull function. Bull Volcanol 75:742

Bonadonna C, Houghton BF (2005) Total grain-size distribution and volume of tephra-fall deposits. Bull Volcanol 67:441–456

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

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

Fierstein JN (1992) Another look at the calculation of fallout tephra volume. Bull Volcanol 54:156–167

Fontijn K, Ernst GG, Bonadonna C, Elburg MA, Mbede E, Jacobs P (2011) The ~4-ka Rungwe Pumice (South-Western Tanzania): a wind-still Plinian eruption. Bull Volcanol, 73:1353-1368

Legros F (2000) Minimum volume of a tephra fallout deposit estimated from a single isopach. J Volcanol Geotherm Res 96:25–32

Pyle D (1999) Quaternary tephra in Africa. Glob Planet Chang 21:95–112

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

Pyle DM (1990) New volume estimates for the Minoan eruption. Thera Aegean World III 2:113–121

Pyle DM (1995) Assessment of the minimum volume of tephra fall deposits. J
    Volcanol Geotherm Res 69:379–382
Sulpizio R (2005) Three empirical methods for the calculation of distal volume of
    tephra-fall deposits. J Volcanol Geotherm Res 145:315–336
Watt SFL, Pyle DM, Mather TA, Martin RS, Matthews NE (2009) Fallout and
    distribution of volcanic ash over Argentina following the May 2008 explosive
    eruption of Chaitén, Chile. J Geophys Res 114:B04207