**DATABASE**       **Open Access**

# Global database on large magnitude explosive volcanic eruptions (LaMEVE)

Helen Sian Crosweller[1*], Baneet Arora[2], Sarah Krystyna Brown[1], Elizabeth Cottrell[3], Natalia Irma Deligne[1,4], Natalie Ortiz Guerrero[1,5,6], Laura Hobbs[1,7], Koji Kiyosugi[8], Susan Clare Loughlin[9], Jonathan Lowndes[9], Martin Nayembil[2], Lee Siebert[3], Robert Stephen John Sparks[1], Shinji Takarada[10] and Edward Venzke[3]

## Abstract

To facilitate the assessment of hazards and risk from volcanoes, we have created a comprehensive global database of Quaternary Large Magnitude Explosive Volcanic Eruptions (LaMEVE). This forms part of the larger Volcanic Global Risk Identification and Analysis Project (VOGRIPA), and also forms part of the Global Volcano Model (GVM) initiative (www.globalvolcanomodel.org). A flexible search tool allows users to select data on a global, regional or local scale; the selected data can be downloaded into a spreadsheet. The database is publically available online at www.bgs.ac.uk/vogripa and currently contains information on nearly 3,000 volcanoes and over 1,800 Quaternary eruption records. Not all volcanoes currently have eruptions associated with them but have been included to allow for easy expansion of the database as more data are found. Data fields include: magnitude, Volcanic Explosivity Index (VEI), deposit volumes, eruption dates, and rock type. The scientific community is invited to contribute new data and also alert the database manager to potentially incorrect data. Whilst the database currently focuses only on large magnitude eruptions, it will be expanded to include data specifically relating to the principal volcanic hazards (e.g. pyroclastic flows, tephra fall, lahars, debris avalanches, ballistics), as well as vulnerability (e.g. population figures, building type) to facilitate risk assessments of future eruptions.

**Keywords:** Volcanic hazard, Explosive eruptions, Database, Global volcanism, Risk

## Background

Explosive volcanism represents a key Earth process that transfers heat and mass from the Earth's interior to the surface. Its products contribute constituents to global geochemical cycles, affecting the near surface geosphere, hydrosphere and atmosphere. Volcanic eruptions have the potential to cause loss of life, disrupt air traffic, impact the climate and significantly alter the surrounding landscape. Knowledge of magnitude-frequency relationships is poorly constrained at a majority of volcanoes but is key for assessing environmental and societal impacts of volcanism on global, regional and local scales. We have created a database of Large Magnitude Explosive Volcanic Eruptions (LaMEVE) to support these assessments and provide basic information on global explosive volcanism.

The LaMEVE database contains the nearly 3,000 Quaternary volcanoes catalogued by the Smithsonian's Global Volcanism Program and over 1,800 explosive eruption records spanning the last 1.8 My. Not all volcanoes currently have eruptions associated with them but have been included to allow for easy expansion of the database as more data are found. The ultimate objective is to go back to the start of the Quaternary at its current definition of 2.58 Ma, but at present the database is limited to the previous definition of 1.8 Ma.

The database, accessible online at www.bgs.ac.uk/vogripa, primarily consists of data on eruption magnitude, age and source volcano location. The data were collected from published literature; principally Volcanoes of the World by Siebert et al. (2010) and journal articles, but also other online databases, books, public reports, conference proceedings, etc. This database should be of particular interest to volcanology researchers, hazard modellers and civil authorities responsible for crisis management.

* Correspondence: Sian.Crosweller@bristol.ac.uk
[1]School of Earth Sciences, University of Bristol, Queens Road, Bristol BS8 1RJ, UK
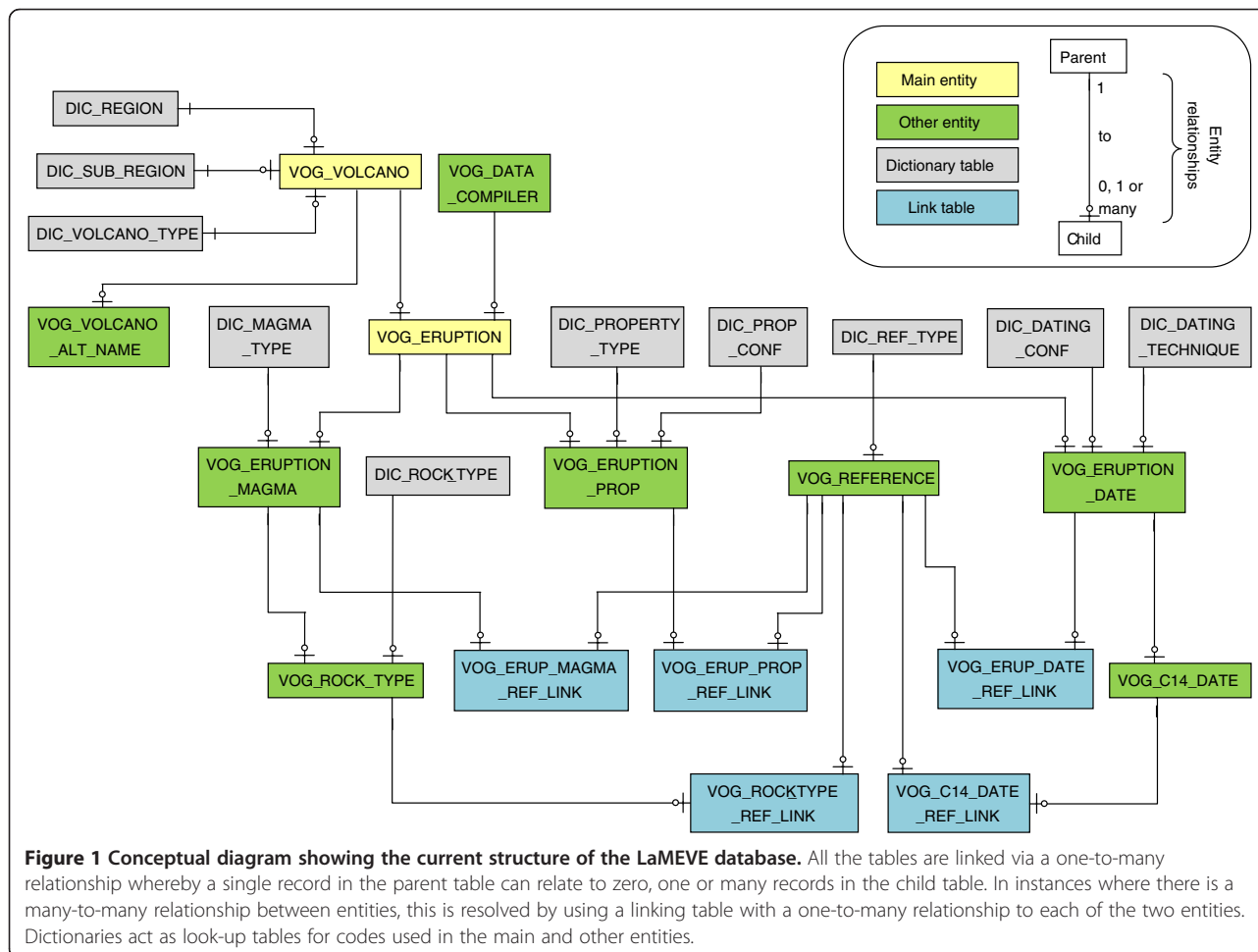Full list of author information is available at the end of the article

There are 3 criteria an eruption must satisfy in order to be included in the database:

1. It must be magnitude (M) (Pyle, 2000) or VEI (Volcanic Explosivity Index) (Newhall and Self, 1982) of 4 or above.
2. It must be dated.
3. It must be from a known source volcano.

The database also contains information on deposit type, deposit volumes, eruption intensity, rock type classification, data errors and uncertainties with indices of data reliability. The database is currently restricted to larger magnitude eruptions because they have large hazard footprints that can threaten large populations. They are difficult to predict and occur less frequently than smaller eruptions. Some of the largest historic eruptions occur at volcanoes presumed dormant; as such, preparation and response are often poor. These larger eruptions are typically responsible for the greatest loss of life in the historical period (Siebert et al., 2010, pp. 41).

There exist two other databases similar in scope to LaMEVE: that of the Smithsonian Institution's Global Volcanism Program (GVP, www.volcano.si.edu; Siebert et al. 2010) and the database by Mason et al. (2004) with eruptions of M≥8. LaMEVE, however, is distinctly different from these two. The database maintained by the Smithsonian's Global Volcanism Program is the leading source of global information on volcanic eruptions at all magnitude scales; however, it currently focuses on Holocene eruptions (the last 10,000 years). Analysis of large magnitude eruptions over this timescale has shown that this is an insufficient amount of time for adequate sampling of the largest eruptions (M>6.5) (Deligne et al., 2010). The database of Mason et al. (2004) concentrated on the very largest magnitude eruptions, which are considered to have the greatest potential for global repercussions. LaMEVE covers all eruptions of M≥4, as these eruptions still have the capacity to cause considerable regional and local disruption.

LaMEVE is the first component of the VOGRIPA (Volcanic Global Risk Identification and Analysis



**Figure 1 Conceptual diagram showing the current structure of the LaMEVE database.** All the tables are linked via a one-to-many relationship whereby a single record in the parent table can relate to zero, one or many records in the child table. In instances where there is a many-to-many relationship between entities, this is resolved by using a linking table with a one-to-many relationship to each of the two entities. Dictionaries act as look-up tables for codes used in the main and other entities.

Project) database of volcanic hazards, which is being developed as part of the Global Volcano Model (GVM, www.globalvolcanomodel.org). The long-term aim of GVM is to have a global source of freely available information on volcanic hazards and volcanism produced to agreed standards and protocols.

This article discusses the structure and fields of the LaMEVE database, including its links with other databases, some of the difficulties encountered whilst constructing the database, how these issues were dealt with, perceived uses, and planned future developments. The paper also outlines planned steps to make the database accessible as well as further development ideas, including the ability for the scientific community to add new data.

## Database Design

The LaMEVE database uses Oracle 10g R2 as a platform and contains 2 main tables and 8 secondary tables with additional look-up tables for codes used. The data format closely follows that of the Smithsonian's GVP database with many of the same data fields and codes, although the underlying structure of LaMEVE is not identical. The Global Volcanism Program's database is currently undergoing restructuring (Cottrell et al., in prep.) to allow public access and to better coordinate with the needs of VOGRIPA and other international databases. The LaMEVE database structure is outlined in Figure 1 with a full schema included in a supplementary file (see Additional file 1).

The following section is split into sub-sections based on the tables within the database. The two main tables are 'Volcano' and 'Eruption,' with the other tables linking to these; therefore the following discussion will be separated into these two main headings. Volcano attributes include spatial location and volcano type, while eruption attributes include timeframe data and quantitative eruption size characteristics. This database was structured using the 'normalisation' method, which is a process of removing duplication in the data, therefore minimising redundancy and the chance of entering inconsistent data (Connolly and Begg, 2005; Powell 2006).

### Volcano

The Volcano table contains information on almost 3,000 volcanoes from around the world. The information is derived from the Smithsonian Institution's GVP database and includes volcanoes that are known or suspected to have erupted within the Holocene or late Pleistocene. Also included are some volcanoes now in the fumarolic stage, which signifies either a long pause between eruptions or residual degassing of a volcano that has not erupted for many thousands of years (Siebert et al., 2010, p3).

Data fields include volcano name and Smithsonian volcano number (VNUM), which are sanctioned by the International Association of Volcanology and Chemistry of the Earth's Interior (IAVCEI). If data are found on an eruption from a volcano not currently in the GVP database, the Smithsonian will be responsible for generating a new VNUM, which we would then use to add the data to our database. Additional fields include a unique table identifier[a], latitude and longitude coordinates, region and sub-region codes and volcano type as assigned by the Smithsonian.

### Volcano Alternative Name

Alternative names, also sourced from the Smithsonian's database, are listed. These can be synonyms (i.e. different names for the same volcano), some of which are different spellings of the same name (e.g. Vulsini / Volsini) while others are significantly different from the official name (e.g. Santorini / Thera). They can also include names of features (e.g. cones or calderas). Although eruption records are linked via the main volcano name, the alternative names can be used to retrieve data from the appropriate volcano records.

### Eruption

The Eruption table lists the unique identifier ID field (required in every table) and the Unit Name, of which there may be more than one listed. This records how the eruptive deposit is referred to in the literature, and provides a way of ensuring that the data all relate to the same eruption. The other columns in the Eruption table are actually links to other data tables (called 'foreign keys'), discussed below. Foreign keys are joined to the main eruption table via a one-to-many relationship allowing multiple pieces of data relating to one eruption to be reported, e.g. more than one date for the eruption from different sources. There is no judgement on what data are included in the database; the user can therefore use their own discretion regarding which data they wish to use for their own analyses.

### Eruption Properties

This is a data-rich table which reports the following eruption characteristics, described separately below:

- VEI
- Magnitude
- Intensity
- Column height
- Tephra fall deposit volume and DRE (Dense Rock Equivalent) volume
- Ignimbrite[b] volume and DRE volume
- Primary intra-caldera material deposit volume and DRE volume

- Bulk deposit volume and DRE volume

Whilst this is a large amount of data to store in one table, these are all interrelated and require the same fields, namely a value and, where appropriate, either an error or qualifier (e.g. less than, greater than) or both. These quantitative characteristics are grouped into 3 categories to facilitate data handling: eruption size, deposit volume and DRE volume. The source of each data point is labelled as either 'literature', in which case it is linked to the reference source, 'calculated' or 'assumed' (see section 2.2.1.5 for more information).

**VEI** The Volcanic Explosivity Index (VEI) devised by Newhall and Self (1982) describes the size of explosive eruptions; values range from 0 (gentle) to 8 (colossal). VEI is reported in the Smithsonian Institution's GVP database and is the most widely reported eruption descriptor in volcanology. These data are used wherever available so as to avoid disparity between the databases. As there is no single piece of information that can fully describe the character of an explosive eruption, the VEI incorporates multiple criteria for assigning a value. The VEI is a semi-quantitative scale based on 8 criteria, but in most cases is based principally on the erupted tephra volume (ancient eruptions) or a combination of plume height and eruptive volume (observed eruptions). This enables the VEI to be assigned from both quantitative and qualitative data, even if some information is missing, allowing a great number of eruptions to be classified. However, the VEI implicitly assumes that eruption magnitude and intensity are related, which is not necessarily true (Carey and Sigurdsson, 1989; Pyle, 2000). The

LaMEVE database records VEI, magnitude (M) and intensity (I), so the relationships between these parameters can be assessed; this topic is discussed in this paper's companion.

**Magnitude** Magnitude is the preferred measure of eruption size used in the LaMEVE database. If not reported in the literature, it is calculated using the formula of Pyle (2000) with reference to deposit volumes:

$$M = log_{10}[erupted \ mass(kg)] - 7 \qquad (1)$$

which is equivalent to:

$$M = log_{10}\left[DRE(m^3) * magma \ density(kg/m^3)\right] - 7 \quad (2)$$

This measure has the added advantage over the VEI of being a more precise measure of eruption size. Typical estimates of volumes justify calculating magnitudes to the first decimal place in the database, allowing more discrimination. It is a quantitative parameter rather than a qualitative indicator that can be used to estimate magnitude-frequency relationships. The largest explosive eruptions on Earth are ~M=9 and so comparable to the magnitude of the largest earthquakes.

For eruptions with an assigned VEI value but no volume data or magnitude reported in the literature, eruption magnitude is assigned according to the relationship found between these two variables. (Figure 2 shows that, for 90% of eruptions in the VEI 4–6 range, the associated magnitude is within the same range (i.e. for VEI 5, 90% of the magnitudes are between 5 and 6). There is more variability for VEI 7 and 8 but the median value is still within the same magnitude range. Since
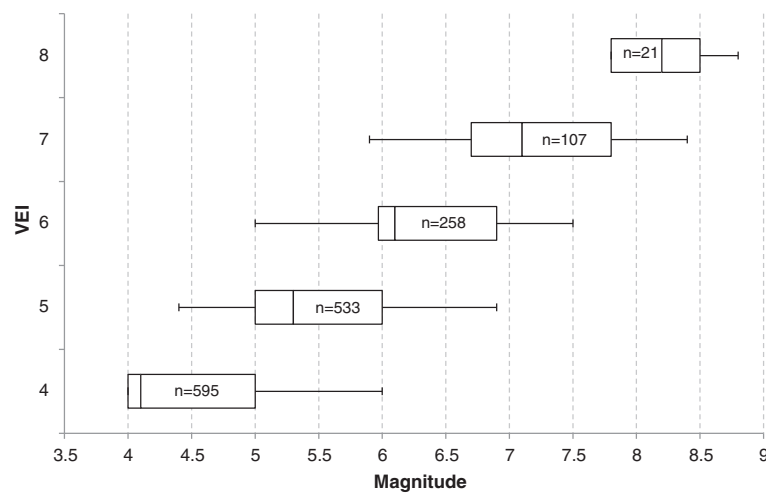


**Figure 2 Box and whisker plot of VEI versus Magnitude.** Boxes represent the range of the 5th and 95th percentiles with the dividing line being the median. Whiskers represent the minimum and maximum values. N values inside the boxes are the number of eruption records at that VEI.

**Table 1 The minimum and maximum magnitude values for each VEI classification in the LaMEVE database as well as the mean magnitude with associated variance**

| VEI | Count | Min. mag. | Max. mag. | Mean mag. | Variance |
|-----|-------|-----------|-----------|-----------|----------|
| 4 | 595 | 4 | 6 | 4.4 | 0.14 |
| 5 | 533 | 4.4 | 6.9 | 5.3 | 0.13 |
| 6 | 258 | 5 | 7.5 | 6.3 | 0.13 |
| 7 | 107 | 5.9 | 8.4 | 7.2 | 0.14 |
| 8 | 21 | 7.8 | 8.8 | 8.2 | 0.06 |

The mean magnitude value is entered as an estimate of magnitude when only VEI is reported.

most VEI values are determined from volumes of eruption tephra the correlation is as expected but re-assuring. Table 1 lists the mean magnitudes associated with each VEI value, which are then entered into the database if a magnitude has not been reported in the literature, and are labelled as 'calculated.'

Since VEI is, in the majority of cases, based on volume then Magnitudes in the range 4.0 to 4.9 should be equivalent to VEI 4. Table 2 shows that >80% of eruption records in the LaMEVE database have a magnitude value consistent with the VEI classification. The percentage of misclassifications (i.e. where Magnitude and VEI are not consistent with each other) is modest. Therefore, where VEI is missing but a magnitude value is reported, the rounded down magnitude value is entered for the VEI and labelled as 'calculated' (e.g. M=4.7 becomes VEI=4). Both values are included in the database to allow original source values to be reported wherever possible.

**Intensity and column height** Intensity is the rate at which magma is discharged and is primarily dependent on the pressure gradient between the magma chamber and the surface, magma viscosity, volatile content, and conduit dimensions (Carey and Sigurdsson, 1989). Intensity is calculated using the mass eruption rate as defined by Pyle (2000):

$$I = log_{10}[mass\ eruption\ rate(kg/s)] + 3 \qquad (3)$$

**Table 2 Number and percentage of records in the LaMEVE database where magnitude values are either less than, the same as, or greater than the VEI classification**

| VEI | M <VEI | | M =VEI | | M ≥VEI+1 | |
|-----|--------|------|--------|-------|----------|------|
| 4 | 0 | 0% | 553 | 92.9% | 42 | 7.1% |
| 5 | 18 | 3.4% | 483 | 90.6% | 32 | 6% |
| 6 | 14 | 5.4% | 237 | 91.9% | 7 | 2.7% |
| 7 | 14 | 13.1% | 89 | 83.2% | 4 | 3.7% |
| 8 | 2 | 9.5% | 19 | 90.5% | 0 | 0% |

It has been established that, for explosive eruptions, column height is related to intensity (Wilson et al., 1978; Sparks et al., 1997), reflecting the rate of thermal energy transfer into the atmosphere. Estimated column heights (in km) are determined either from direct observations for some historical eruptions, or by using the maximum clast size dispersal method of Carey and Sparks (1986). We note that this method has been recently updated by Burden et al. (2011), but has not yet been applied to the data in this database.

**Deposit volume and DRE** Where available, three distinct deposit volumes are recorded (tephra fall, ignimbrite, and primary intra-caldera material), which combined constitutes the bulk deposit volume. Volumes of associated lava flows are omitted from the database and excluded from calculations, thus given magnitudes are for the explosive phases of an eruption only and do not necessarily represent the magnitude of the entire eruption. The Dense Rock Equivalent (DRE) volumes are also stored. DRE corresponds to the unvesiculated erupted magma volume, i.e. the pre-eruption magma's volume. Where it is not reported in the literature, DRE is calculated using the following equation:

$$DRE(km^3) = \frac{tephra\ vol(km^3) * tephra\ density(kg/m^3)}{magma\ density(kg/m^3)} \qquad (4)$$

Tephra density, unless otherwise reported, is assumed to be 1000 kg/m$^3$. Magma density varies according to the magma type (see below). For some records estimates and measurements of tephra deposit density are given; in these cases the deposit-specific literature value is used.

**Data quality indicator** There is notable variation in the level of detail and consistency with which data are reported in the literature as well as in the clarity of explanation regarding data collection. This can be problematic when entering data into a database which follows a standardised format resulting in discrepancies between the reliability of different records which can be subtle. Therefore a 'data quality' indicator (ranging from 0 to 3) has been assigned to each eruption date and volume according to the criteria listed in Tables 3 and 4. This provides the user with a simple initial assessment of data reliability, and a way to exclude less reliable data from analyses. A full assessment of reliability should involve user evaluation of the original source information.

We assign the lowest data quality level to those magnitudes and volumes that are 'assumed.' These are cases where quantitative data are inferred from qualitative descriptions. For example, an eruption might be described as 'Plinian' or 'caldera forming'; such an

## Table 3 Criteria for applying data quality indicators to eruption dates

| DATING DATA QUALITY INDICATOR | |
|---|---|
| **VALUE** | **DESCRIPTION** |
| 0 | Only qualitative data provided (e.g. Holocene). |
| 1 | Date reported in the literature without error and or dating method not specified OR stratigraphically dated OR reported errors >30% OR unit has multiple published dates that differ by >30%. |
| 2 | Single date (historical or geochronology) with a method and errors cited OR unit has multiple published dates that differ by 10-30% OR historical ages uncertain by more than one year. |
| 3 | Single precise historical date within a year OR multiple published dates that differ by <10%. |

Assigning data quality indicators to eruption dates permits users to prioritise superior analytical techniques. For multiple ages the above criteria are thus applied to those ages where there is no basis for assigning one age or method as superior to another.

eruption is clearly explosive and therefore eligible for inclusion in the database (e.g. 'Tegalsruni' eruption unit from Merapi). The Smithsonian classifies these events as VEI 4 or larger, but does not assign a specific VEI. A magnitude of ≥4 is thus assigned to these eruptions in the LaMEVE database with the derivation method listed as 'assumed'. This permits the event to be included in a count of explosive eruptions despite an unknown actual eruption size. Eruption size and corresponding data quality can be updated as new information becomes available. When an assumed VEI/magnitude is entered the following corresponding assumed volumes are entered, as from Newhall and Self (1982):

## Table 4 Criteria for applying data quality indicators to volume data

| VOLUME DATA QUALITY INDICATOR | |
|---|---|
| **VALUE** | **DESCRIPTION** |
| 0 | No isopach map or evidence to support cited volume OR assumed value. |
| 1 | Simple isopach map or thickness data (< 10 measurements) OR volume calculation with no methodology given OR volume derived from method other than isopach map (e.g. size of caldera; VEI estimate; duration and intensity of column heights). |
| 2 | Isopach map drawn from >10 and <30 measured thicknesses with calculation method described. |
| 3 | Isopach map drawn from >30 measured thicknesses with calculation method described. |

There is no standardised method for constructing isopach maps from thickness data and several methods of extrapolation beyond the areas with data have been devised. The diversity of methods means there can be significant volume differences even with the same datasets when different methods are applied. Volumes are reported directly from sources using whichever method they applied, but preference is given to those calculated using the Pyle or Fierstein methods (Pyle 2000, Fierstein and Nathenson 1992). However, there are other methods and pre-1989 literature with various ways of integrating the volume from isopach maps. Some volumes include pyroclastic flow deposits where isopach maps are not necessarily readily drawn. Here the number of thickness measurements as above is used to assign a volume data quality indicator.

- VEI 4: tephra volume $\geq 0.1$ $km^3$
- VEI 5: tephra volume $\geq 1$ $km^3$
- VEI 6: tephra volume $\geq 10$ $km^3$
- VEI 7: tephra volume $\geq 100$ $km^3$
- VEI 8: tephra volume $\geq 1,000$ $km^3$

### Eruption Date and $C^{14}$ Date

These two tables are linked as they both refer to eruption date. While the Eruption Date table contains both start and end date fields, only the former is populated. End dates will be added in the future but there are very little reliable data for non-historical events and further thought is required on how to populate this entry. For historical eruptions, the day, month and year of eruption are often reported; there are currently 238 such eruptions in the LaMEVE database. Older eruptions are less precisely dated and often only the year is known. For such eruptions the dating method and associated error margin are recorded when known. As well as standard calendar years for recent eruptions, all eruption dates are recorded in years before present (BP) with 'present' taken as 1950; negative numbers in this field therefore reflect years after 1950.

For prehistoric events a number of dating techniques can be employed. For eruptions that occurred in the last ~50,000 years radiocarbon ($C^{14}$) dating is the most common dating method; just over 550 radiocarbon dates are currently listed in the database. However, radiocarbon years are not equivalent to standard calendar years (cal BP) because of fluctuations in atmospheric $C^{14}$ and thus a calibration must be applied to determine the actual eruption date. This calibration is sometimes done by the author of the source reference and can be reported with or without the uncorrected $C^{14}$ date. Where reported, the original $C^{14}$ dates and errors are recorded in the Eruption $C^{14}$ Date table. Where no calibrated date has been given, we have applied the calibration using a publically available tool. The method used is reported in the database, but generally the Calib program (v6.1.1) by Stuiver and Reimer (1993) is used for dates <26 Ky and the Fairbanks et al. program (2005) for events up to 50,000 cal BP. If no error has been reported for the $C^{14}$ date, an error margin of 12% of the date is entered into the calibration program, as derived from statistical analysis of those dates with reported errors. For uncalibrated dates that fall outside the range of even the Fairbanks program, and where no literature-sourced date is available, an age of 50 Ka (the limit of the Fairbanks et al. program) is entered into the Year BP field (with the standard 12% error) but is labelled as 'assumed;' as Year BP is a required field, we need to enter an approximate date for the eruption. In all cases, calibrations are reported using the 1-sigma range (68.3% confidence intervals), as 2-sigma ranges are not available using the

Fairbanks tool. Data from Bryson et al. (2006) of >2,000 calibrated radiocarbon dates from the late Pleistocene and Holocene have also been cited.

Other commonly used dating methods are Argon-Argon ($^{40}$Ar-$^{39}$Ar), Potassium-Argon (K-Ar), stratigraphy and tephrochronology. The first two are similar to radiocarbon dating insofar as they provide absolute dates, with some uncertainty, whereas the other two provide an approximate date relative to bracketing stratigraphic layers which usually have been precisely dated by another method.

### Rock Type Classification and Eruption Magma

These tables contain data on the geochemistry-based rock type classifications of the erupted magma for each eruption and the magma type range respectively. The codes used in the Rock Type Classification table are the same as those used by the Smithsonian (see Figure 3), and multiple codes are often entered for a single event because magma compositions can change over the course of an eruption. Each rock type is entered as either a standard or a minor (<10%) component. The magma type range is then determined and stored in the Eruption Magma table, or marked as 'unknown' if no data have been reported. If a single rock type has been entered into the Rock Type Classification table, this becomes the magma type. If a number of distinct rock types are reported for a single eruption event, a code to cover the range of these rock types (based on $SiO_2$ content) is entered; for example, an eruption producing both andesite (A) and rhyolite (R) would be given the magma code AR to represent a range from andesite to

rhyolite (e.g. the 30,000 BP caldera collapse eruption of Aso). There are currently 22 magma type codes entered in the relevant dictionary table: the 10 used by the Smithsonian, a further 12 based on combinations of these codes to cover eruptions with a range of rock types, and the 'unknown' category. This field can be expanded in future, if required, to encompass all possible combinations of rock types.

Each magma type has an *'assumed magma density'* ranging from 2,300 kg/m$^3$ for rhyolitic magmas to 2,700 kg/m$^3$ for basaltic magmas[c]. When the magma type is unknown, a median value of 2,500 kg/m$^3$ is entered. These density values are used in Equation 4 to calculate the DRE volume from deposit volumes. Whilst these are relatively arbitrary values the error introduced into DRE and M values is small compared to other sources of uncertainty. For example, for a tephra volume of 100 km$^3$, using a density value of 2,300 kg/m$^3$ gives a DRE of 43.49, whereas when the density is assumed to be 2,700 kg/m$^3$ the DRE is 37.04. Using either of these values to calculate the magnitude would result in the same value, correct to 3 decimal places (7.000). Furthermore, assumed density values are overwritten if the literature includes a calculation of magma density, and these reported values are used instead in the DRE volume calculation.

### Reference

An extensive table contains sources for all the LaMEVE data, and all data tables are linked to it so that the source for each piece of data can be traced. Included in the database are data from a number of pre-existing databases, such as the aforementioned Smithsonian
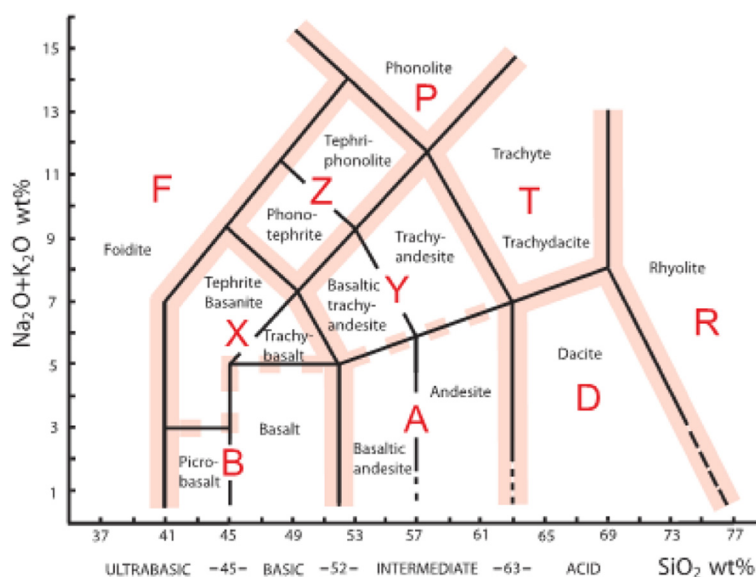


**Figure 3 Codes used in *Rock Type Classification* table, taken from Siebert et al. (2010).** These codes, or combinations of them, are also used in the *Eruption Magma* table.

Institution's GVP database and, more recently, a significant contribution of Japanese data following the translation of two main databases; one by the Quaternary Volcanoes in Japan 2008 (http://riodb02.ibase.aist.go.jp/strata/VOL_JP/EN/index.htm, last modified October 2011) and the other by Hayakawa 2010 (http://gunma.zamurai.jp/database/, last modified September 2010). The primary references for data derived from the Smithsonian's GVP will be available online from that database directly (Cottrell et al., in prep.). It also contains data gathered from a variety of published material, mostly peer-reviewed journal articles.

Commonly more than one paper is written about a particular eruption or deposit, which can result in different estimates for the same kind of data, for example an eruption date or volume. In the interests of being inclusive, all published data are included. However, a 'preferred' default value has been selected using the data quality indicators described in Tables 3 and 4, whereby the data point with the highest quality level defaults as the preferred value. If there is more than one entry with the same data quality level, the most recent publication is the default.

### Searching the Database

The database is available online through a website hosted by the British Geological Survey (BGS). In addition to information regarding the project's background and its future objectives, the website allows users access to the entire database via a customisable search facility. This search tool is flexible and multiple criteria can be entered. A GIS-enabled graphical tool allows users to select an initial Area of Interest (AoI), if desired. An attributes search can be performed either separately, or in addition to the area search, by entering criteria for a number of fields, currently:

- Volcano name
- Volcano type (select from list)
- Eruption start date (exact/partial date or range)
- Region (select from list)
- Eruption size (range)
- Composition (select from list)
- Bulk volume (range)
- DRE volume (range)
- Column height (range)
- Tephra fall volume (range)

The relevant records are displayed in a summary table with a map of the selected area with the volcano locations indicated (see Figure 4 for an example). Volcanoes are marked on the map with small red dots with the current page of results highlighted with dark blue dots.

By hovering over a blue dot it turns light blue and highlights the relevant results in the summary table. This also works in reverse. The search criteria used to generate the results are also listed.

By default results are ordered alphabetically by volcano name but there is also the option to sort by volcano type or region. The summary dataset can be downloaded as a spreadsheet or the search can be refined further via two methods. By selecting 'Refine Search' from a menu on the left, the original search screen is redisplayed and changes can be made. Alternatively, the 'Filter' tool can be used to select the volcano types and or regions of interest.

When an individual volcano is selected from the initial results page, a volcano summary, which includes a location map, is displayed at the top of the page (this is in addition to the full search results listed). The volcano-specific summary includes the following information:

- Volcano name
- Alternate name
- Region/Sub region
- VNUM (hyperlinked to the Smithsonian's website)
- Volcano type
- Latitude/Longitude

The 'Eruption Details' section (see Figure 5) lists all the available data in the database for the selected eruptive events. Each piece of data is appended with 'literature' or 'calculated' where appropriate and details of the literature source are available by clicking on the adjacent 'Reference' tab.

## Discussion
### Uses of the database

LaMEVE constitutes an extensive volcanic data repository available for researchers who would have previously had to do substantial literature searches to find information on past activity. It is anticipated that civil authorities responsible for crisis management will be interested in the data to inform future management of regional volcanoes.

Whilst the database currently holds a large number of records, it is inevitably incomplete and likely contains errors. The key to improving the database will be user input and future updates as new data are published, relevant existing data are identified and existing records are updated, corrected or modified. A notable advantage to having the database online is that it will allow users to easily access the data, identify and fill gaps and correct mistakes. Any changes or updates suggested by external users will be assessed by the database manager before being entered into the database.
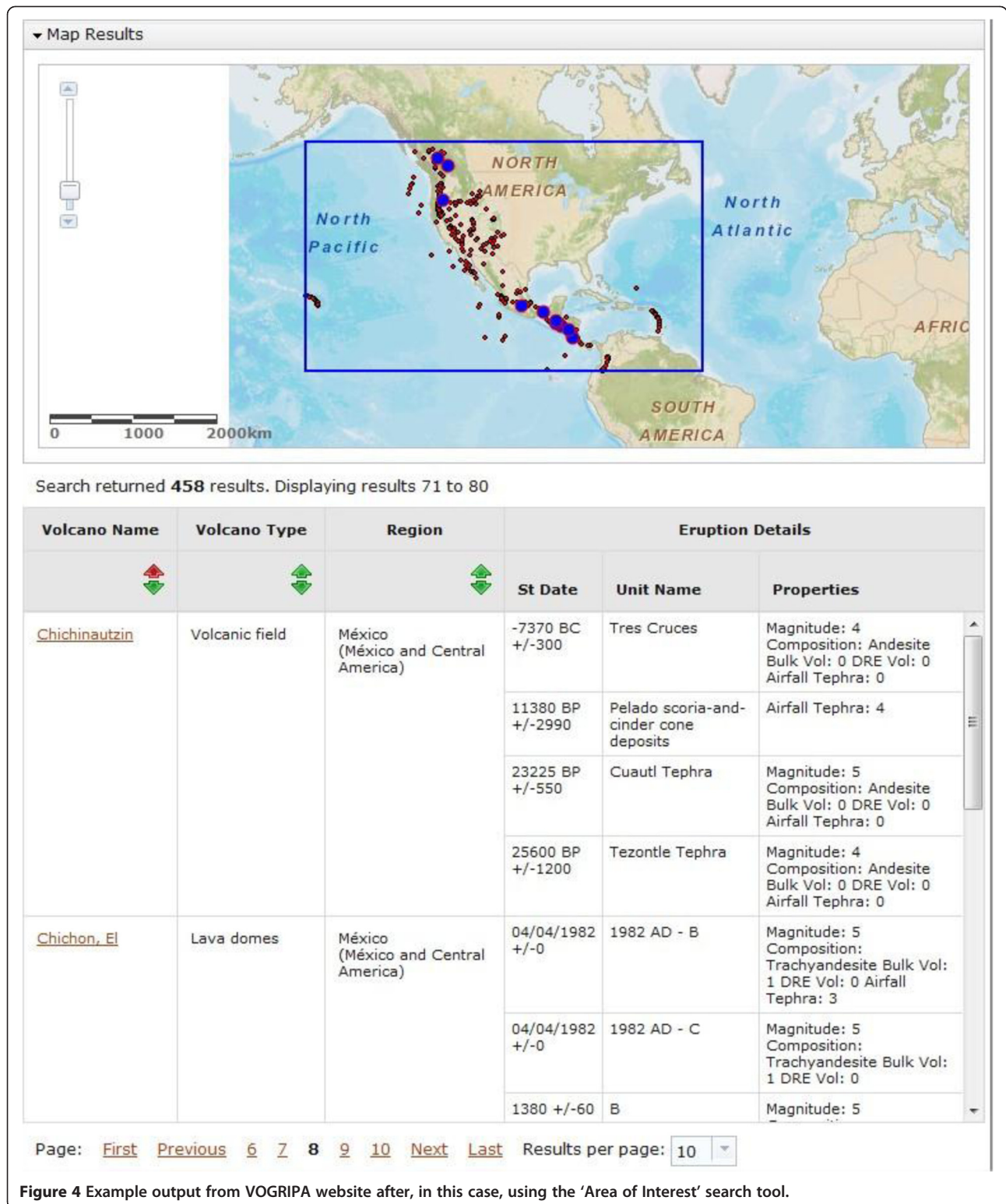
**Figure 4 Example output from VOGRIPA website after, in this case, using the 'Area of Interest' search tool.**

The database will also be used as a research tool to analyse global, regional and local patterns of volcanic activity to identify locations at high risk and gaps in knowledge about hazard and risk, as per the goals of VOGRIPA and GVM. Furthermore, the recurrence rates of different magnitude eruptions and particular hazardous phenomena, as well as the stationarity of global volcanism can be evaluated. Under-recording of events,
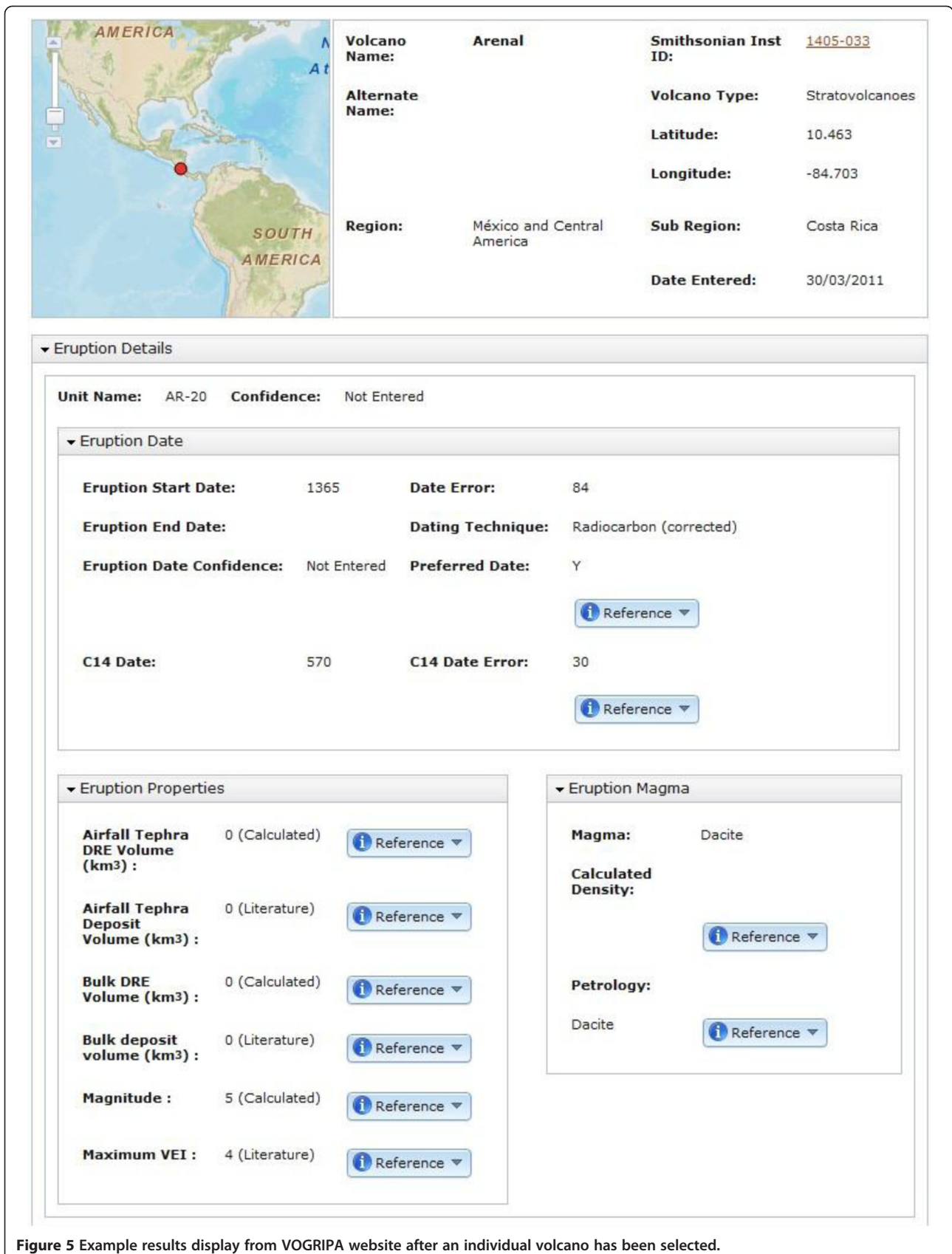
**Figure 5 Example results display from VOGRIPA website after an individual volcano has been selected.**

especially those of smaller magnitude, is a significant problem (Deligne et al., 2010; Furlan, 2010) and this can be investigated with the data. A preliminary analysis of the LaMEVE database looking at patterns in the data along with an assessment of under-recording, both temporally and spatially, will be presented in a forthcoming paper.

### Benefits

LaMEVE provides standardised, internally consistent data on global explosive volcanism. The database has considerable flexibility for the user who can, for example, make decisions on what data to use in analysis, e. g. decide which is the best date to use, how to combine multiple dates or whether to discard dates. We note that it is the responsibility of the user to explain and justify data usage.

LaMEVE can be queried for data that can feed into hazard and risk assessments allowing users to analyse hazard and risk within local, regional and global contexts; this should be particularly beneficial in under-studied areas. An example of an application of volcanic hazard and risk assessment is a study of 22 high priority GFDRR (Global Facility for Disaster Reduction and Recovery, World Bank) countries (Aspinall et al., 2011). Here the LaMEVE database has been used to calculate regional recurrence rates of explosive eruptions as a function of magnitude, which will be discussed in a later paper.

### Planned future development

The LaMEVE database is the first output of VOGRIPA. A number of databases will be developed to link with the existing structure to provide data specifically relating to the main volcanic hazards (e.g. pyroclastic flows, tephra fall, lahars, debris avalanches, ballistics). These databases will be cross-linked. This hazard-specific information will enable users to determine which hazards are most common at each volcano, based on past activity, and also the impacted spatial extent.

In order to translate hazard into risk simple measures of vulnerability will be added, including population densities, building types, locations of critical infrastructure and level of monitoring. This will enable multi-scale risk mapping to identify those areas at greatest risk from future volcanic activity and so highlight where future efforts should be focussed.

The LaMEVE database is currently based largely on historical information and terrestrial geological records of Quaternary volcanoes. However, there is the potential to add information from tephra deposits in marine sediments and from indications of explosive volcanism in ice cores (e.g. from sulphate anomalies). Challenges in incorporating these data include the estimation of eruption magnitude and determination of the source volcano.

The LaMEVE database is to be made publically accessible at the time this paper is published as v1.0. For Holocene eruptions, LaMEVE data not already included in the GVP database will be imported by GVP where it will be maintained and updated in collaboration with VOGRIPA and GVM. The GVP database has been restructured for this purpose and will be available for search and download. This partnership prevents divergence and redundancy, and the division of labour will accelerate progress on both databases. The Pleistocene eruptions will be subsequently regularly updated by VOGRIPA. The current plan calls for annual updates (i.e. v2.0 released in 2013) but if other updates are required, they will be denoted as sub-versions (e.g. v1.1). There will be three mechanisms of adding new data, updating existing entries and correcting errors. Firstly, as VOGRIPA and GVM are funded until the end of 2014, GVM researchers will continue to search for and incorporate new data and make corrections when errors are identified. Secondly, researchers can send new data and modifications and draw attention to errors by contacting the database manager identified on the website. Thirdly, researchers can enter data after being given clearance as a GVM user. This data will still be checked for validity and consistency by the database manager. Only the current version of the database will be available to the public; the evolving version will not be made public prior to specified release dates when it will replace the earlier version.

### Difficulties

The principal difficulty in creating a database with such a wide scope is compiling information. There are two main issues. First, the data may not actually exist, either due to a lack of geological investigations or because data are unpublished. Second, not all published materials are easily accessible (either because of journal subscription or language issues). There are also potentially relevant data published in the 'grey literature,' which may also be inaccessible to academics. These concerns emphasise the value of opening the database to additional scientific contributions as undoubtedly more data exist than have been found thus far.

It is evident from even a basic analysis of the data in the database that under-recording of eruptions over time, particularly those of lower magnitudes, is a considerable problem. This will be discussed in detail in a forthcoming companion paper, and also in Deligne et al. (2010) using just the Holocene record.

Comparisons between the data are complicated by the fact that there are numerous data types as well as a mix of collection methods. For example, eruption dates can

be historical or geological, which have notably different accuracies. Furthermore, a number of methods are used to determine certain eruption characteristics, such as eruption size, which complicates comparisons. Differences in data accuracy are unavoidable, particularly over the timescale covered in the database; it is the responsibility of users to take this into consideration.

Any database of this kind will, and should, evolve as more data become available and errors are corrected. Studies which use this database will need to be explicit about which version was used. Every published database will be archived to allow comparison of analyses. The development, updating and archiving of databases will require considerable effort and resources and should be sustained.

## Conclusions

The LaMEVE database contains nearly 3,000 Quaternary volcanoes and just under 2,000 explosive eruptions with magnitude ≥4 spanning the last 1.8 My. It constitutes the first part of the larger VOGRIPA database and is accessible online at www.bgs.ac.uk/vogripa. The most common types of dates stored are historical events (238) and radiocarbon dates (558). An analysis of patterns, relationships and under-recording issues in the dataset will be published separately.

The database constitutes a major resource for those conducting research or looking for information on large magnitude explosive eruptions from around the world as it synthesises data from a wide range of sources. It is available via an easy-to-use web interface with the facility to download data. The database will be maintained and updated within the GVM project with opportunities for the volcanological community to add new data and identify errors and make corrections. The design permits future expansion, so that hazard-specific information can be stored as well as vulnerability data, which will enable global risk assessments to be carried out.

## Availability and requirements

The database is available at www.bgs.ac.uk/vogripa with no restrictions on access.

## Endnotes

[a]An automatically generated unique code (typically numerical) assigned to each row in the table to enable cross-referencing to each record.

[b]Ignimbrite' refers to both pyroclastic flow and pyroclastic surge deposits.

[c]From Figure 3, the assumed densities are applied as following to each rock types: P/T/R = 2,300 kg/m$^3$; D = 2,400 kg/m$^3$; A/Y/Z = 2,500 kg/m$^3$; F/X/B = 2,700 kg/m$^3$.

## Additional file

> **Additional file 1: Entity relationship diagram of the LaMEVE database.**

**Author details**
[1]School of Earth Sciences, University of Bristol, Queens Road, Bristol BS8 1RJ, UK. [2]British Geological Survey, Nicker Hill, Keyworth, Nottingham, NG12 5GG, UK. [3]Smithsonian Institution, National Museum of Natural History, Washington, DC 20560, USA. [4]Current address: GNS Science, PO Box 30368, Avalon, Lower Hutt 5040, New Zealand. [5]Current address: Faculty of Engineering, Universidad Mariana, San Juan de Pasto, Nariño, Colombia. [6]Current address: Instituto de Geofísica, Universidad Nacional Autónoma de México, Ciudad Universitaria, C.P. 04510, Coyoacán, Distrito Federal, México. [7]Current address: Lancaster Environment Centre, Lancaster University, Lancaster, LA1 4YQ, UK. [8]Department of Geology, University of South Florida, 4202 East Fowler Avenue, Tampa, Florida 33620, USA. [9]British Geological Survey, Murchison House, West Mains Road, Edinburgh, EH9 3LA, UK. [10]Geological Survey of Japan, AIST, Site 7, 1-1-1, Higashi, Tsukuba, Ibaraki 305-8567, Japan.

**References**
Aspinall W, Auker M, Hincks TK, Mahony S, Nadim F, Pooley J, Sparks RSJ, Syre E (2011) Volcano hazard and exposure in GDRFF priority countries and risk mitigation measures - GFDRR Volcano Risk Study. Bristol University Cabot Institute and NGI Norway for the World Bank, Washington DC, NGI Report 20100806; 309pp, 3 May 2011

Bryson RU, Bryson RA, Ruter A (2006) A calibrated radiocarbon database of late Quaternary volcanic eruptions. Earth Discussions 1(2):123–134

Burden RE, Phillips JCC, Hincks TK (2011) Estimating volcanic plume heights from depositional clast size. J Geophys Res 116:B11206

Carey S, Sigurdsson H (1989) The intensity of plinian eruptions. Bull Volcanol 51 (1):28–40

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

Connolly T, Begg C (2005) Database systems: A practical approach to design, implementation, and management. Pearson Education Limited, Harlow

Deligne NI, Coles SG, Sparks RSJ (2010) Recurrence rates of large explosive volcanic eruptions. J Geophys Res 115:B06203

Fairbanks RG, Mortlock RA, Chiu T-C, Cao L, Kaplan A, Guilderson TP, Fairbanks
    TW, Bloom AL, Grootes PM, Nadeau M-J (2005) Radiocarbon calibration curve
    spanning 0 to 50,000 years BP based on paired 230Th/234U/238U and 14C
    dates on pristine corals. Quat Sci Rev 24(16–17):1781–1796
Fierstein J, Nathenson M (1992) Another look at the calculation of fallout tephra
    volumes. Bull Volcanol 54(2):156–167
Furlan C (2010) Extreme value methods for modelling historical series of large
    volcanic magnitudes. Stat Model 10(2):113–132
Quaternary Volcanoes in Japan (2008) Geological Survey of Japan, http://riodb02.
    ibase.aist.go.jp/strata/VOL_JP/EN/index.htm. Accessed regularly
Hayakawa Y (2010) Hayakawa's 2000-year and one million-year database,
    http://gunma.zamurai.jp/database. Accessed regularly
Mason BG, Pyle DM, Oppenheimer C (2004) The size and frequency of the largest
    explosive eruptions on Earth. Bull Volcanol 66(8):735–748
Newhall CG, Self S (1982) The Volcanic Explosivity Index (VEI): an estimate of
    explosive magnitude for historical volcanism. J Geophys Res
    87(C2):1231–1238
Powell G (2006) Beginning database design. Wiley, Indianapolis
Pyle DM (2000) Sizes of volcanic eruptions. In: Sigurdsson H, Houghton BF,
    McNutt SR, Rymer H, Stix J (eds) Encyclopedia of Volcanoes. Academic Press,
    London
Siebert L, Simkin T, Kimberly P (2010) Volcanoes of the world, 3rd edn. University
    of California Press, Berkeley
Sparks RSJ, Bursik MI, Carey SN, Gilbert JS, Glaze LS, Siggurdsson H, Woods AW
    (eds) (1997) Volcanic plumes. Wiley, New York
Stuiver M, Reimer PJ (1993) Extended 14C data base and revised CALIB 3.0 14C
    age calibration program. Radiocarbon 35(1):215–230
Wilson L, Sparks RSJ, Huang TC, Watkins ND (1978) The control of volcanic
    column heights by eruption energetics and dynamics. J Geophys Res 83
    (B4):1829–1836