BOOM! Tephrochronological dataset and exploration tool of the Southern (33–46° S) and Austral (49–55° S) volcanic zones of the Andes

Consuelo Martínez Fontaine, Vanessa Peña-Araya, Chiara Marmo, Marine Le Morvan, Guillaume Delpech, Karen Fontijn, Giuseppe Siani, Lucile Cosyn-Wexsteen

BOOM! Tephrochronological dataset and exploration tool of the Southern (33–46° S) and Austral (49–55° S) Volcanic Zones of the Andes

Consuelo Martínez Fontaine[a, b, 1], Vanessa Peña-Araya[c], Chiara Marmo[a], Marine Le Morvan[e], Guillaume Delpech[a], Karen Fontijn[d], Giuseppe Siani[a], Lucile Cosyn-Wexsteen[a]

[a]Géosciences Paris-Saclay (GEOPS), Université Paris-Saclay, CNRS UMR 8148, 91405 Orsay, France

[b]Laboratoire des Sciences du Climat et de l'Environnement (LSCE), Laboratoire mixte CNRS-CEA, Avenue de la Terrasse, 91198 Gif-sur-Yvette Cedex, France

[1]Present address : Centro de investigación en Tecnologías para la Sociedad C+, Universidad del Desarrollo. Av. Plaza 680, Edificio S, 7610658 Las Condes, Chile. E-mail: cdpmartinez@udd.cl

[c]Inria, Laboratoire CNRS, Université Paris-Saclay - Bât. 660 91405 Orsay Cedex, France

[d]Department of Geosciences Environment and Society, Université Libre de Bruxelles, Av. F Roosevelt 50, 1050 Brussels, Belgium

[e]Inria Saclay – Île-de-France, Parietal team, 1 Rue Honoré d'Estienne d'Orves, 91120 Palaiseau, France

Corresponding author: Consuelo Martínez Fontaine

Abstract

Tephrochronology studies the deposits of explosive volcanic eruptions in the stratigraphic record. The Southern (SVZ, 33–46° S) and Austral (AVZ, 49–55° S) Volcanic Zones of the Andes are two very active volcanic zones where tephrochronology is of great use. There, it can be used to improve chronologies of paleoenvironmental records in Patagonia, an area providing valuable records at global scale; as well as to identify areas likely to be affected by volcanic eruptions in the future, essential for producing volcanic hazard maps. The close proximity of many volcanic centers with recurrent explosive activity, which have very similar geochemical compositions, and very often poor age constraints, represent a challenge for the study of tephrochronology in the region. In addition to this, the ever-growing amount of tephrochronological information in the area, dispersed in different types of publications which vary greatly in format, makes the integration of the data produced by different actors, and consecutively its interpretation, increasingly difficult. Here we address this issue by compiling the BOOM! dataset, which integrates ~30 years of research on 32 active volcanic centers and 132 different eruptions, which took place during the last 20,000 years. To help users find and reuse data in the large dataset, we developed an online platform which provides user-friendly tools for exploring it, and helps users download subsets of it. To integrate this very heterogeneous information, special attention was given to include information which allows users to evaluate data quality and comparability, as well as to provide tools in the explorer for users to filter data by different criteria. The integration of this dataset opens new perspectives for the development of novel visualizations of tephrochronological data, for example, to better understand the multidimensional uncertainties associated with it. For example, uncertainties associated with analytical precision, with age estimates of both tephra deposits and volcanic eruptions, and of tephra classification. Additionally, it allows for the use of robust statistical tools to correlate tephra deposits, including those based on machine learning algorithms, which are here explored.

Keywords: Tephrochronology, Southern Volcanic Zone of the Andes, Austral Volcanic Zone of the Andes, Geochemistry, Radiocarbon, Machine learning

1 Introduction

Tephrochronology is a discipline of geosciences which uses the deposits of explosive volcanic eruptions as stratigraphic and chronological markers. These deposits are called tephras, and by studying them, researchers can reconstruct the eruptive history of volcanic centers: how often volcanic activity occurs, the magnitude of the eruptions, and the dispersion of the volcanic products. The Southern (SVZ, 33–46° S) and Austral (AVZ, 49–55° S) Volcanic Zones of the Andes are two very active volcanic zones, where ~65 volcanic centers have been identified. Many of these centers have had recurrent explosive activity in historical times, e.g., Llaima (Naranjo & Moreno, 2005), Puyehue-Cordón Caulle (Naranjo et al., 2017), Calbuco (Sellés & Moreno, 2011), Hudson (Naranjo et al., 1993) and Lautaro volcanic centers (Mayr et al., 2019). While for others, evidence of recurrent explosive activity during the last ~20,000 years has also been identified, e.g., Mocho-Choshuenco (Rawson et al., 2015), Michinmahuida (Amigo et al., 2013) and Mount Burney volcanic centers (Smith et al., 2019). More so, some of these volcanic centers are very explosive, with distal dispersion of tephras being registered hundreds of kilometers from their source, as the H1 eruption from Hudson volcanic center (Naranjo & Stern, 1998; Kilian et al., 2003; Stern, 2008; Del Carlo et al., 2018; Smith et al., 2019). Because of the latter, one important concern in these areas is protecting surrounding villages and other important infrastructure (e.g., power plants or dams) from potentially dangerous eruptions. Part of this work is done by the Chilean Geology and Mining agency SERNAGEOMIN, which produces volcanic hazard maps of the different volcanic centers and regions in the country and distributes them to the population (e.g., Bertin et al., 2018). In order to produce these maps, SERNAGEOMIN relies on tephrochronological and volcanological information from their own surveys as well as from peer-reviewed scientific publications (e.g., Sellés & Moreno, 2011; Bertin et al., 2018), which helps them identify areas likely to be affected by future eruptions. Besides this very important application, and because of the recurrent explosive activity of the SVZ and AVZ, there is great potential of using tephrochronology to improve the chronologies of paleoenvironmental records in Patagonia (Fontijn et al., 2014). If tephras deposited during the same eruptive event are identified in different sedimentary archives (e.g., lake sediment cores, marine sediment cores, peat cores, or

archaeological sites), the chronologies of different records (e.g., paleoclimatological, paleoceanographical, or archaeological records) can be synchronized. Paleoenvironmental records in Patagonia provide valuable climatic records, such as the variations in the Southerly Westerly Winds (e.g., Moreno et al., 2018); valuable oceanographic records, such as changes in oceanic circulation in the Southern Ocean (e.g., Siani et al., 2013); as well as unique records describing interactions between human populations, climate change and volcanic activity at high latitudes (e.g., Villarosa et al., 2006; Prieto et al., 2013). In order to correctly integrate these records, having good chronologies is essential, thus highlighting the potential of using tephrochronology in the area.

In the last four decades, tephrochronology together with volcanology have increased our understanding of the eruptive history of the SVZ and AVZ, revealing higher than previously thought recurrence rates and explosivity of many of the volcanic centers in the area (e.g., Chaitén and Michinmahuida volcanoes (Amigo et al., 2013; Watt et al., 2013; Moreno et al., 2015; Alloway et al., 2017a, b; Martínez Fontaine et al., 2021)). This work, and especially the study of sedimentary archives which favor the preservation of tephras (such as lake cores (e.g., Bertrand et al., 2014)), has revealed a very intricate tephrostratigraphic record in the area, given by the close proximity of the volcanic centers together with their high eruptive recurrence rates. In order to disentangle this record, tephrochronologists try to fingerprint tephras using their physical characteristics, geochemical composition, stratigraphic position, and age (e.g., Lowe, 2011). By doing this, they can correlate tephras deposited in different locations and ideally identify their volcanic source and the specific eruption during which they were deposited. In the SVZ and AVZ, doing this is not trivial. For example, many volcanic centers and eruptions have very similar geochemical compositions, or in other cases, the age estimates of tephra deposits have high uncertainties, or they do not exist at all, making it hard to distinguish tephras based on these criteria. This issue is enhanced by the *unFAIRness* of the tephrochronological data (e.g., physical characteristics, geochemical composition, stratigraphic position, and age of tephras) in the area, i.e., data is not readily Findable, Accessible, Interoperable, or Reusable (Wilkinson et al., 2016; Abbott et al., 2022). At present, the available tephrochronological and volcanological information of the SVZ and AVZ can be found dispersed in journals, undergraduate or doctoral theses, publications from government institutions (SERNAGEOMIN), and personal collections,

which vary greatly in format, and with some information sometimes very difficult to find. This, together with the increased data stream associated with the development of analytical techniques (such as Electro Probe microanalysis (EMPA, also referred to as EMP), typically used for analyzing glass (or mineral) major element composition of volcanic ash shards, e.g., Lowe (2011)), makes the integration of the data produced by different actors (researchers and practitioners from government institutions), and consecutively its interpretation, more and more difficult. The increasing need for organization of the vast and growing amount of tephrochronological and volcanological data in the region, is an issue that has already been addressed in previous publications (Fontijn et al., 2014, 2016; Alloway et al., 2017a, 2017b). Here we address it by making two main contributions. First, by compiling a dataset which standardizes and integrates ~30 years of research mainly in tephrochronology (complemented with volcanological information), on 32 active volcanic centers and 132 different eruptions of the SVZ and AVZ during the last 20,000 years. Additionally, we developed an online platform[1] which provides user-friendly tools for the exploration of the large dataset, and helps users download subsets of it. Both contributions aim to be a step towards making these data FAIR (Wilkinson et al., 2016).

2 Methods

2.1 Understanding the problem: Interviews with domain experts

In order to characterize the work and needs of people using tephrochronology in the SVZ and AVZ, we conducted a series of interviews with researchers and practitioners who are part of this community, including volcanologists, tephrochronologists, palaeoclimatologists, among others. This information was then used as input to design the dataset structure and the exploration tool. Between June and August 2020, we interviewed 19 people: two researchers in volcanology, one researcher in paleoceanography and a team of 16 people from the Chilean National Volcanic Vigilance Network, part of SERNAGEOMIN. However small, the interviewed sample represents to some extent the diversity of the tephrochronological community in the area. Each interview lasted between one and three hours and was

[1] https://boom-vis.lisn.upsaclay.fr/

divided in two main phases. In the first phase, we asked them about their demographic information (age, research area, position) and their daily work using tephrochronological data. More specifically, we focused on understanding what scientific questions they try to answer, which data they use to do so, how they acquire, organize and analyze these data, and what problems they commonly encounter in this process. In the second phase, we asked them to show us a concrete example from their work in order to confirm or complement the answers given during phase one, and particularly to understand how they use visualizations to organize data and answer their scientific questions. Finally, we showed them an early prototype of the exploration tool in order to get initial feedback to iterate its design. All interviews were audio recorded with the participants' consent (available in section S1 of the supplementary material). A summary of the answers from the interviews can be found in Table S1 in the supplementary material.

The interviews had two main results. On the one hand, we confirmed that the unFAIRness of the data was an important issue for the participants. All of them expressed how difficult and time-costly it was for them to collect, organize and analyze tephrochronological data given the dispersion of data sources and the lack of common publication standards. On the other hand, the interviews allowed us to list the information more routinely used by tephrochronologists working in the area, as well as pinpointing the problems they encounter when using these data (Table S1). Then, we designed a structure that contains this information in a way that, together with the explorer, addresses those issues. In particular, the main problems we addressed were: difficulties organizing the bulk of data, ignoring the existence of data, not understanding a priori if data from different sources are comparable or not, and problems visualizing data in its multidimensionality (e.g., visualizing at once the geographical position of the tephra deposits, their geochemical composition, stratigraphy and chronology). More detailed descriptions on the design of the BOOM! dataset and explorer, and how they address these issues, are presented in sections 3 and 4, respectively. The information from the interviews substantially aided the development of the structure of the dataset and the exploration tool.

2.2 Data collection

We collected tephrochronological and volcanological information on 32 active volcanic centers belonging to the SVZ and AVZ, between volcanoes Llaima (~38.7° S) and Mount Burney (~52.3° S). Here, volcanic centers are considered active if evidence of volcanic activity occurring during the last ~20,000 years has been identified. This time frame is relevant for volcanology as it captures the volcanic centers that are likely to still erupt in the future. However, it is also informed by practical limitations due to the presence of the Patagonian Ice Sheet (PIS) in the area during the last glacial period (~20,000–60,000 years Before Present; (Davies et al., 2020)). During the last glacial period, the PIS covered most of the Andes south of ~38° S, and thus, unconsolidated tephra deposits older than ~20,000 years have been largely eroded.

Three types of data were collected:

i. Physical characteristics and geochemical analyses (e.g., EMPA) of **pyroclastic material** (tephras themselves). A full list of the geochemical analyses included in the dataset is detailed in Table S2.
ii. Radiocarbon (^{14}C) ages of **organic matter** *associated* with the tephras, typically recovered from within the deposits or from palaeosols immediately underlying them (used to estimate when tephras were deposited).
iii. Geochemical composition and $^{40}Ar/^{39}Ar$ ages of **effusive material** (e.g., lava, breccia, prismatically joint blocks).

Even though this dataset is aimed to help the tephrochronological community, volcanological data (i.e., data of effusive material) was also collected to complement the tephrochronological information. Because effusive material builds the volcanic edifice itself, there is little uncertainty regarding its volcanic source. On the contrary, tephra deposits can be identified thousands of kilometers away from any volcanic center, and often their source has not been robustly identified yet. Thus, the geochemical composition of effusive materials was included in the dataset as a geochemical "ground truth", to which to compare the geochemical composition of tephra deposits for which the volcanic source has not been yet identified. This information is particularly useful for volcanic centers for which little tephrochronological information is available. On the other hand, $^{40}Ar/^{39}Ar$ ages provide age constraints

when no ^{14}C ages or other stratigraphic constraints are available. Errors associated to $^{40}Ar/^{39}Ar$ ages are usually one order of magnitude larger than those of ^{14}C ages and generally do not allow to differentiate between eruptions in the SVZ and AVZ. However, in recent years $^{40}Ar/^{39}Ar$ errors are getting smaller thanks to the development of more precise analytical techniques.

From the publications in the literature, we prioritized those which included analyses of the geochemical composition of tephras, together with age or stratigraphic constraints, as long as the composition reported was magmatic (in this particular geologic setting, this corresponds to ~40–80 wt.% SiO_2). Otherwise, and because of the data scarcity for many volcanic centers, we included all publications we could find in the literature which met these criteria. However, we always tried to include information necessary for users to evaluate the quality and comparability of the data (sections 4 and S5) and flagged the data when some issue was identified (section S5.4, Table S4). In total, we collected data from 72 scientific publications published in peer-reviewed journals, five publications from SERNAGEOMIN, and two doctoral theses (the whole list of publications can be found below). These publications included works on tephrochronology (31), volcanology (28), geological mapping (6), palynology (6), glaciology (2), volcanic hazard mapping (1), archaeology (1), limnology(1), pedology (1) and paleoceanography (1). The information included in the dataset was extracted either directly from the publication tables, text and supplementary material, or alternatively through private requests to the authors when the data discussed in the publication was not readily available. The authors were informed of the purpose of the data request beforehand. From this process, ~16,800 sample analyses were included in the dataset, identified as sourced in 132 different eruptive events, from 32 different volcanic centers. Tephra deposits described in the literature for which the volcanic source has not been identified yet were also included in the dataset, which correspond to ~11% of the sample analyses.

Publications in peer reviewed journals:

Abarzúa & Moreno, 2008; Abarzúa et al., 2004; Alloway et al., 2015; Alloway et al., 2017a; Alloway et al., 2017b; Amigo et al., 2013; Bertrand et al., 2008a; Bitschene et al., 1993; Bouvet de Maisonneuve et al., 2012; Brahm et al., 2018; Bucchi et al., 2015; Carel et al., 2011; Casati et al., 2019; Clapperton et al., 1995;

Constantini et al., 2011; Del Carlo et al., 2018; D'Orazio et al., 2003; Fontijn et al., 2016; Futa & Stern, 1988; Geoffroy et al., 2018; Gerlach et al., 1988; Haberle & Lumley, 1998; Harambour, 1988; Heusser et al., 1989 Heusser et al., 2000; Hickey-Vargas et al., 1989; Jacques et al., 2014; Kilian et al., 2003; Kratzmann et al., 2008; Lara et al., 2004; Lara et al., 2006; Lohmar et al., 2012; López-Escobar et al., 1992; López-Escobar et al., 1993; López-Escobar et al., 1995; Martínez Fontaine et al., 2021; Mayr et al., 2019; Miranda et al., 2013; Weller et al., 2014; Moreno & León, 2003; Moreno et al., 2015; Moreno, 2004; Morgado et al., 2015; Morgado et al., 2019; Motoki et al., 2006; Naranjo & Moreno, 1991; Naranjo et al., 1993; Naranjo & Stern, 2004; Naranjo & Moreno, 2005; Naranjo & Stern, 1998; Naranjo et al., 2017; Pesce & Moreno, 2014; Rawson et al., 2015; Rawson et al., 2016; Reubi et al., 2011; Schindlbeck et al., 2014; Simmons et al., 2020; Singer et al., 2008; Smith et al., 2019; Stern et al., 2011; Stern et al., 2015; Stern, 1992; Stern, 2008; Tagiri et al., 1993; Villarosa et al., 2006; Wastegård et al., 2013; Watt et al., 2011a; Watt et al., 2011b; Watt et al., 2013; Weller et al., 2014; Weller et al., 2015; Weller et al., 2017; Weller et al., 2019.

Publications from SERNAGEOMIN :

Naranjo & Stern, 1998; Naranjo & Moreno, 2005; Lara et al., 2006; Moreno et al., 2010; Sellés & Moreno, 2011; Bertin et al., 2018.

Doctoral theses: Lohmar, 2008; Mella, 2008.

2.3 Data Visualization: Development of the BOOM! explorer

The BOOM! explorer is complementary to the dataset itself and aims to provide tools for users to make the most of it. As mentioned before, we designed the explorer to address the main problems participants expressed during the interviews (section 2.1, Table S1). In turn, this helped us better understand the dataset we were collecting and modify its structure in order to better answer to community requests. For example, when looking for ways to visualize if data from different sources were comparable or not, we realized that important information to evaluate this was missing in the dataset, such as the type of register the sample corresponds to (pyroclastic material, organic matter, or effusive material) or the type of analysis performed in the sample (bulk versus micro analytical). Taking this into consideration, this

information was included in the dataset and users can choose which type of register and analysis to visualize in the explorer. This process led an iterative redesign of the explorer, which included: modifying the visual representation of the data, including additional User Interface interactions (e.g., filtering data by different criteria), and reorganizing the visual elements in the explorer. This iterative redesign was also supported by informal meetings with practitioners, some of them from the same group we had previously interviewed, where they told us their impressions of the tool, expressed their need for additional features, and gave us general feedback. The development of the explorer additionally helped us identify gaps in the dataset; include additional information, for example, a particular type of data of a particular volcanic center or eruptive event in order to better characterize them; as well as identify outliers and flag data. A detailed description of how the explorer can be used to explore the dataset is exemplified in section S6 by a use case.

Given the above process, we decided to build a light visualization tool, which can be easily installed in a local environment. In this way, other researchers can make modifications/improvements to the explorer. The BOOM! explorer is published under an MIT License[2], which allows to reproduce it and build from it, under the same license as the original one. The tool was developed in Javascript, using node.js[3] as the back-end environment. For the visualization components, we used the Leaflet library[4] for the map, D3.js[5] for the timelines, and dc.js[6] with Crossfilter[7] for the geochemical composition scatterplots. The dataset is delivered as a CSV file (see section 3 for more details), but in order to allow fast queries in the visualization, we exported it to a SQLite[8] dataset that is accessed from node.js. This transformation also includes precomputed queries that help with the performance of the tool. The BOOM! explorer (https://boom-vis.lisn.upsaclay.fr/) is hosted by the Laboratoire Interdisplinaire des Sciences du

[2] https://mit-license.org/
[3] https://nodejs.org/
[4] https://leafletjs.com/
[5] https://d3js.org/
[6] https://dc-js.github.io/dc.js/
[7] https://crossfilter.github.io/crossfilter/
[8] https://www.sqlite.org

Numérique, LISN (Université Paris-Saclay, CNRS & Inria), France, which guarantees its hosting and access as long as the server exists.

3 BOOM! Dataset structure

Tephrochronology identifies deposits from explosive volcanic eruptions in the stratigraphic record, samples, and characterizes them based on their physical characteristics, geochemical composition, stratigraphy and age. Then, based on this characterization and the comparison with available tephrochronological (and volcanological) information in the area, tephrochronologists ideally identify the volcanic source and the specific eruption during which the tephra was deposited, i.e., they *classify* it. The BOOM! dataset follows this structure. It is composed of a collection of descriptions and laboratory analyses of tephra deposits identified in the SVZ and AVZ (complemented with data on some lava flows). In the BOOM! dataset, each of these descriptions or analyses is called a "sample observation", and one sample of a tephra deposit may be described by one or many sample observations. For example, in order to characterize a tephra deposit, researchers can describe the physical characteristics of the deposit as a whole (e.g., its color, grain size, and thickness) and at the same time analyze the geochemical composition of dozens of individual volcanic glass shards obtained in just one sample. Then, one sample observation corresponds to the analyzed geochemical composition of one individual glass shard (as exemplified in Figure 1). The sample as a whole, on the other hand, is described by the physical characteristics of the tephra deposit and the geochemical composition of all the analyzed glass shards. Additionally, tephra deposits are ideally not described isolated, but as part of a stratigraphic column, and their relative stratigraphic position is used to identify their volcanic source. In the BOOM! dataset, each stratigraphic column is called a "section", and each interval of that column is called a "sub section". In total, 1,303 sections are contained in the BOOM! dataset. However, it is important to note that for 876 sections, only one tephra deposit is described, and no name is given for the section by the authors. The remaining 427 sections are given a name in the original publication and are described by between one and 46 sub sections, characterized by between one and ten samples, each described by between one and 105 sample observations, corresponding to a total of 2,899 samples and 16,768 sample observations.

These descriptions of tephra deposits in the BOOM! dataset have been, for the most part, already classified, i.e., their volcanic source and the specific eruptive event during which they were deposited have been identified by the authors. Thus, they can be used to characterize different volcanic centers and individual eruptive events in the SVZ and AVZ and serve as a reference to compare unclassified samples to. In the BOOM! dataset, an eruptive event is characterized by all the samples (and respective sample observations) which have been correlated with that eruptive event (Figure 1). Since these samples have been collected in different locations, in different types of sedimentary archives, and provide different types of information (physical characteristics, geochemical composition, age), together they can be used to provide a more complete characterization of an eruption. For example, when did the eruption occur, what was its explosivity, the dispersion of its products, and the type of the eruption it corresponded to. In the same way, all the samples attributed to a particular volcanic center can be used to describe it. These samples may or may not be attributed to a particular eruptive event yet, in which case they are labeled as *unknown* eruptions. Nevertheless, all the samples attributed to a volcanic center can be used to describe it, though grouping them by eruption provides more information to understand their eruptive history: the recurrence of volcanic activity, its general explosivity, if temporal changes in volcanic activity have been registered, etc.

In practice, the BOOM! dataset is distributed as two .CSV files: a main file named BOOMDataset, and a secondary file named MeasurementRuns. In the BOOMDataset file, each sample observation corresponds to a row, which can be characterized by a maximum of 80 attributes (columns), which in turn can be grouped by the type of information they provide (Figure 1, Figure 2, Table S3):

1. **Identification attributes**: these attributes correspond to information used to identify the sample observation, but which do not further describe the tephra deposit itself. They can be subdivided into ID attributes, Position attributes, Reference attributes, and Analysis attributes.
2. **Characterization attributes:** they correspond to the actual observations (data) of the samples, which are used to characterize –and ideally fingerprint– tephra deposits. These attributes can be further subdivided into: Stratigraphy attributes (the stratigraphic position of the tephra deposit

and age estimate(s)); Physical characteristics attributes (color, thickness, and grain size of the tephra deposit); and Geochemical composition attributes (major elements (both *raw* and normalized), trace elements and isotope ratios compositions of either pyroclastic or effusive material).

3. **Interpretation attributes**: They correspond to information that is inferred based on the characterization of the tephra deposit (together with previous tephrochronological information). In particular: the volcanic source of the deposit (Volcano); the specific eruptive event during which the volcanic products (tephra or lava) were deposited (Event); and the Magnitude and Volcanic Explosivity Index (VEI) of the eruptive event. The fact that these are *interpretation* attributes is very relevant in tephrochronology. On the one hand, because tephra can be deposited far away from their volcanic source (e.g., H1 eruption from Hudson volcano has been identified ~700–900 km away from its source (Naranjo & Stern, 1998; Kilian et al., 2003; Stern, 2008; Del Carlo et al., 2018; Smith et al., 2019)), identifying which volcanic center and during which eruption a tephra was deposited is not always trivial. Additionally, the tephrochronological record is incomplete and biased towards the stratigraphic sections that are best preserved or more easily accesible. Thus, the interpretation of the eruptive event, volcanic center, magnitude and VEI may be updated as new information becomes available. In fact, the volcanic source of some of the tephra samples in the BOOM! dataset have been questioned and re-interpreted over time. For example, the volcanic source of MIC2 and COR1 tephra deposits identified by Naranjo & Stern (2004) as sourced from the Michinmahuida and Corcovado volcanic centers, have been questioned in later publications, such as Watt et al., (2009, 2013), Amigo et al. (2013) and Alloway et al. (2017a, b). In these cases, the reinterpretation of the samples is indicated by a flag (see below).

4. **Metadata**: A final group of attributes corresponds to Flags and Comments, which give additional information to interpret the data. On the one hand, the Flag attribute is used to communicate issues identified when including the data in the BOOM! dataset. Sample observations can be flagged for seven types of reasons (detailed in Table S4): problems with the sample ID

(SampleID_Issue), its position (Position_Issue), the Digital Object Identifier (DOI) of the publication where the data was obtained from (DOI_Issue), its age (Age_Issue), its geochemical composition (Geochemistry_Issue), the volcanic source it is attributed to (VolcanicSource_Issue), or the name given to the eruptive event (EventName_Issue). In every case, the type of flag is indicated under the Flag attribute, and the particular reason why the sample observation is flagged is detailed under the Flag Description attribute. On the other hand, the Comments attribute corresponds to additional information that can help users interpret the data, which is not described by any other attribute.

As mentioned before, in total ~16,800 sample observations are included in the BOOM! dataset, interpreted as sourced from 32 volcanic centers and 132 different eruptive events. The amount of information related to each volcanic center varies from a total of less than 20 sample observations for volcanic centers such as Huequi, Corcovado, Yanteles, or Cay, to thousands of sample observations for volcanic centers such as Llaima, Quetrupillán, Mocho-Choshuenco, Puyehue-Cordón Caulle and Chaitén. For each volcanic center, sample observations attributed to between one (Aguilera volcano) and twenty-six (Mocho-Choshuenco) different eruptive events are included in the dataset. Because of the heterogeneity of the information included in the BOOM! dataset (the different types of sedimentary archives, sources of the information, dates of publication, and of uses of the tephrochronological information), not every sample observation is described by all of the 80 attributes previously described. In fact, in the BOOMDataset file, 58.3% of the attribute information is “missing” (Figure 2).

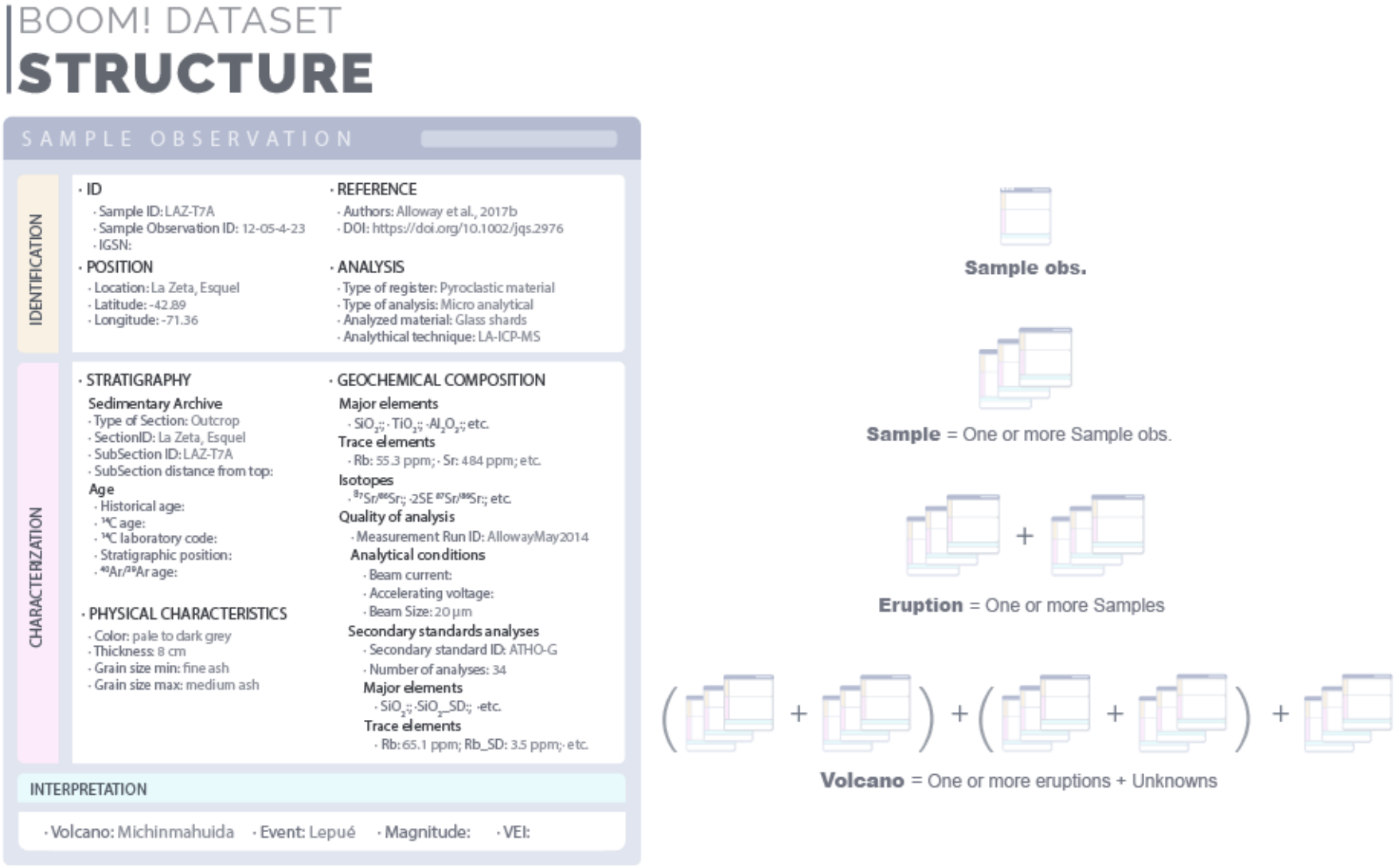


Figure 1 BOOM! dataset structure. To the left are shown the attributes in the BOOMDataset and MeasurementRuns files, grouped by the type of information they provide, as described in the text. Attributes from the MeasurementRuns file correspond to the information below "Quality of analysis". The metadata attributes are not illustrated. The information shown corresponds to sample LAZ-T7A (Alloway et al. 2017b). The full description of the attributes and units is described in Table S3.

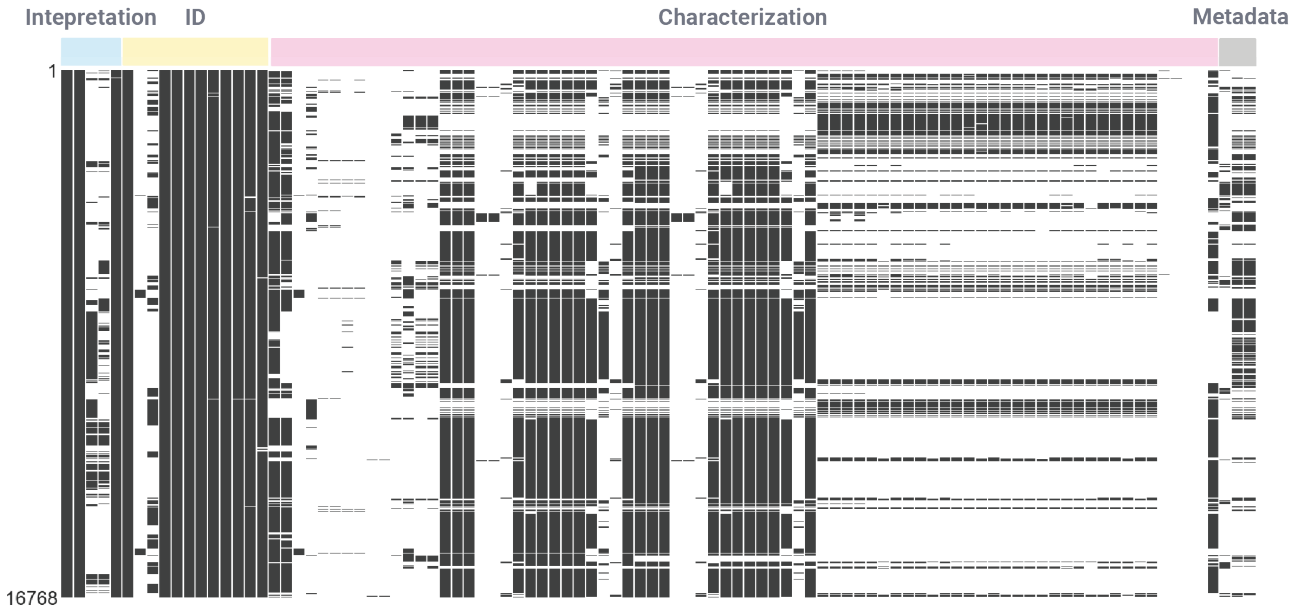


Figure 2 Representation of the "missing" values in the BOOMDataset file. Each column represents an attribute and each row a sample observation. The attributes are grouped according to the type of information they provide: Interpretation attributes, ID attributes, Characterization attributes, and Metadata. Dark grey cells represent filled values and white cells represent "missing" values. The figure was created using the Python library Missigno[9].

In addition to the main BOOMDataset file, a second CSV file named MeasurementRuns is also distributed (Figure 3). This file contains information which is part of the Geochemical composition attributes previously described (attributes under "Quality of analysis" in Figure 1). It corresponds to information used by geochemists to evaluate the quality and comparability of the geochemical analyses performed by different laboratories. We decided to store this information in a different file because of the way the geochemical composition of most of the samples in the BOOM! dataset were analyzed. Most of the data in the dataset corresponds to analyses of the geochemical composition of individual volcanic glass shards (~88% of the sample observations in the BOOMDataset file), which were analyzed either by EMPA (for major elements) or by Laser Ablation - Inductively Coupled Plasma - Mass Spectrometry LA-ICP-MS, for trace elements. Individual glass shards are analyzed in "batches", and during one batch (called a "measurement run" in the BOOM! dataset), dozens or hundreds of individual glass shards from different samples can be analyzed. To evaluate the accuracy and precision of the analyses performed during one

[9] https://github.com/ResidentMario/missingno

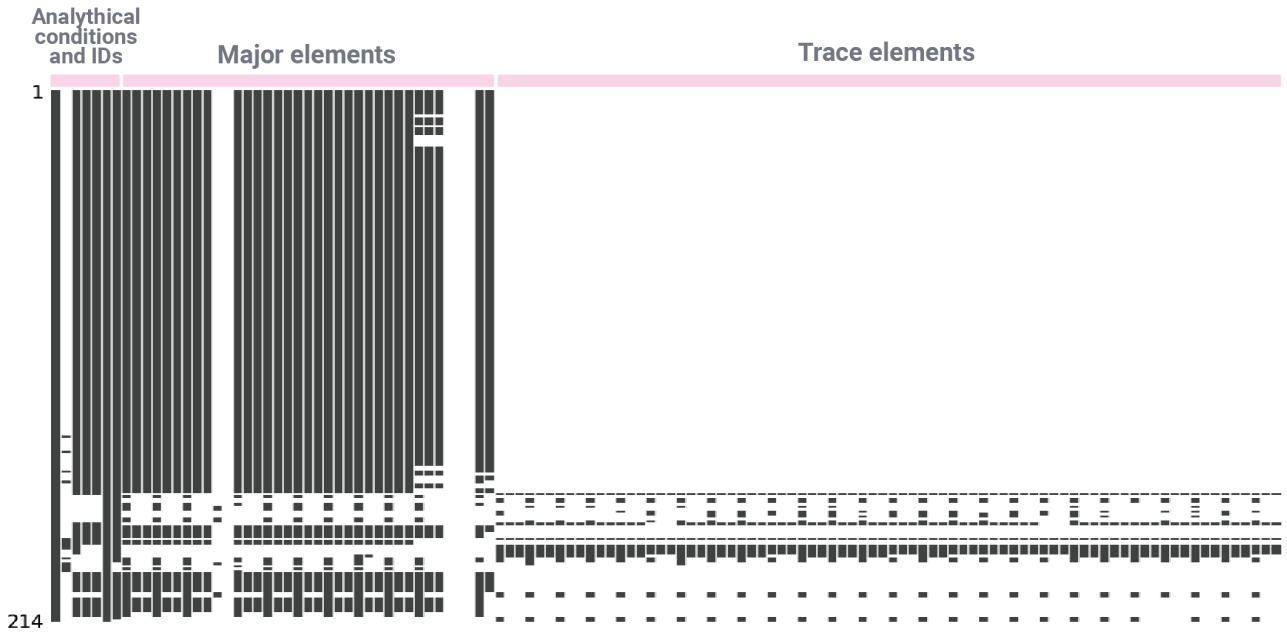


Figure 3 Representation of the "missing" values in the MeasurementRuns file. Dark grey cells represent filled values and white cells represent missing values. The figure was created using the Python library Missigno.

measurement run, geochemists repeatedly analyze the composition of different secondary standards along with the glass shards to be studied. These secondary standards ideally have certified reference values (e.g., Jochum et al., 2006, 2011, 2016) which can be compared with the analyzed values during each measurement run to estimate their accuracy and precision. In particular, by comparing the mean value of each analyzed element of each secondary standard to the certified values, the accuracy of the analysis of each element during that particular measurement run can be estimated (more details on how to do this with the BOOM! dataset is explained in section 4.2.2). In a similar way, the precision can be estimated as the standard deviation of the analyzed composition of each element on each secondary standard analyzed during a particular measurement run. Taking this into consideration, in order to provide information for other researchers to evaluate the comparability of their geochemical analyses, researchers generally publish: the label given by them to a particular measurement run; which samples and sample observations were analyzed during that run; which secondary standards were analyzed; the number of analyses of each secondary standard in that run; the mean value and the standard deviation of each analyzed element of each secondary standard; as well as the analytical totals if major elements

were analyzed. In the MeasurementRuns file, each row corresponds to the information of one secondary standard during a particular run, thus one run can correspond to several rows. The reason for analyzing different secondary standards during one measurement run is that they have different geochemical compositions (e.g., GOR132-G (~46 wt.% SiO_2), StHs6/80-G (~63.7 wt.% SiO_2), ATHO-G (~75.6 wt.% SiO_2) (Jochum et al., 2006)), and the performance of an analytical technique is associated with that composition. Taking the latter into consideration, to assess the quality of the analyses, geochemists usually consider the accuracy and precision of the secondary standard or multiple secondary standards with a geochemical composition relatively similar to that of the analyzed sample. In addition, in order to evaluate the comparability of EMPA analyses, the analytical conditions during each measurement run are also considered (e.g., beam size, accelerating voltage, and beam current). Thus, this information is also stored in the MeasurementRuns file. In both the MeasurementRuns and BoomDataset files, each batch analysis label is indicated under the measurement run attribute, which links both files. The MeasurementRuns file contains 81 different measurement runs where between one and seven secondary standards were analyzed. The measurement runs described in the MeasurementRuns file corresponds to 10,891 of the sample observations in the BOOMDataset, attributed to 26 volcanic centers and 61 eruptive events. In the same way as with the BOOMDataset file, because of the heterogeneity of the data, ~69% of the cells in the MeasurementRuns file are "missing" (Figure 3).

The two files of the dataset are hosted on the ESPRI server of the IPSL, France, which guarantees its hosting and access as long as the server exists. Both sets can be downloaded directly from the IPSL catalog[10] as a Web Map service (WMS), Web Feature service (WFS) or as a CSV file; or using the BOOM! explorer to download subsets of it (see section S6 for more details).

[10] https://doi.org/10.14768/47b4525f-ff39-4940-a963-4d2673f2362e

4 Data comparability

As already mentioned, the BOOM! dataset integrates tephrochronological (and volcanological) information extracted from 79 different publications. Even though all these publications provide relevant information for tephrochronology, they can be very heterogeneous: they correspond to different types of publications (scientific publications, government reports, and doctoral theses); from different disciplines (tephrochronology, volcanology, archaeology, limnology, geological mapping, palynology, glaciology, pedology, and paleoceanography); which were published through 30 years of research (1988-2021 C.E.); and which additionally can have very different research goals (e.g., produce volcanic hazard maps, do research on volcanology and tephrochronology, and improve chronologies of sedimentary archives for paleoclimate, palynology or paleoceanography studies). The idea behind collecting the BOOM! dataset, is that all this information together can be used to better describe eruptive events and volcanic centers in the SVZ and AVZ and help disentangle the tephrochronological record there. However, because of the heterogeneities in the source of the information, different publications might study the same phenomenon in a slightly different way. This does not necessarily mean that the information cannot be integrated, but it is necessary to understand what the information actually represents in order to correctly interpret it. Taking this into account, several of the attributes included in the dataset, correspond to information for the user to evaluate the comparability of the data. In the following we focus on attributes associated with age estimates and geochemical composition comparability. For further description of other relevant attributes, such as type of register, type of section, and flags, the reader is referred to the supplementary material (section S5).

4.1 Age estimate comparability

A very important part of fingerprinting tephra deposited during different eruptive events in a very active volcanic region, such as the SVZ and AVZ, is estimating the age of the eruption. In the timeframe comprehended in the dataset (last 20,000 years), the age of a tephra is generally estimated by ^{14}C dating organic matter which has been identified *associated* to the tephra deposits. This is done because the pyroclastic material itself generally cannot be directly dated with enough precision to differentiate

between eruptions (for example, by the $^{40}Ar/^{39}Ar$ method, for which errors are generally on the order of thousands of years, whereas eruptions can occur every as little as tens of years (e.g., Singer et al., 2008)). Because of this, researchers try to identify organic matter which is ideally embedded *within* the tephra deposit itself, and alternatively immediately above or below it. This organic matter can be ^{14}C dated with a precision of tens of years, and in that way, provide a more precise reference for when the eruption occurred. Both the stratigraphic position of the organic matter with respect to the tephra (above, within, or below), and the specific material that has been dated, are important aspects to consider when evaluating the comparability of the different age estimates. For instance, if a piece of charcoal is found within a tephra deposit, it is assumed that the charcoal is the result of hot pyroclastic material burning living trees during an explosive eruption. Thus, by dating the charcoal, a more or less precise age of when the eruptive event happened, can be obtained. Alternatively, if no organic matter is found within the tephra deposit itself, soil that has formed either above or below the deposit, might be dated. Unfortunately, the ^{14}C age of the soil might not represent the *true* age of the tephra deposit. Soil is formed, among others, by the degradation of organic matter, and during its formation it incorporates organic matter that is contemporary to the soil formation process, but it can also incorporate older organic matter that was already present when soil formed; as well as younger organic matter which can infiltrate from above, for example through a permeable tephra layer. Thus, soil ^{14}C ages or bulk organic sediment ^{14}C ages, correspond to a mean value of organic matter of different ages. Additionally, soil can be found either below or above the tephra deposit, and thus, it might have formed long before the eruption, or it might have taken years to develop after the tephra deposition. In these cases, the ^{14}C ages from the soil are generally interpreted as maximum and minimum ages of the tephra deposit, respectively. In a similar way, if organic macro remnants, such as charcoal or wood, are found below or above the tephra deposit and dated, those ages are also considered as maximum and minimum ages. Because of the latter, a good practice, which can help reduce the age estimate uncertainties is replicates, i.e., dating more than once a determined sample. In the BOOM! dataset, however, most ^{14}C ages have not been replicated, which can be associated with increased expenses, but also with material availability.

Another important aspect to consider when estimating the age of eruptive events, is the type of sedimentary archive where the tephra -and associated organic matter- were identified. Depending on the depositional environment, different types of organic matter for ^{14}C dating will be available. For example, in the ocean, generally wood or charcoal are not identified and planktic foraminifera are dated instead to provide an age estimate of tephras. However, these ages cannot be directly compared with on land ages, as the concentration of ^{14}C in the ocean and the atmosphere are not in equilibrium. Because of this, a correction must be applied to marine ages in order to compare them with atmospheric ages. This correction changes in time and space, and is a subject of ongoing research (e.g., Siani et al., 2013; Merino-Campos et al., 2019). A similar situation occurs with ^{14}C ages in lake sediment cores. In this case, ideally terrestrial macrofossils found within the lake sediment cores are dated, which represent the atmospheric ^{14}C age, as long as the sediment has not been reworked. In many cases, however, these are not available and bulk sediments are date instead. Because of different processes occurring in lakes, bulk sediment ^{14}C content is generally not in equilibrium with the atmosphere. Among these processes: terrestrial input of *old* organic matter by river inflow, groundwater ^{14}C content, dissolution of aged carbonate or variations in biological activity in the lake (e.g., Geyh et al., 1998; Yu et al., 2007). Because of this, the difference between the contemporary atmosphere and lacustrine sediment ^{14}C age can change both in time and space and is characteristic of each lake. Thus, marine ^{14}C ages and bulk sediment ^{14}C ages in lake sediment cores, provide a loose reference of the time of deposition of tephra deposits, but should in no case be directly compared with on land ages without a proper correction.

Taking the latter into consideration, in order to correctly interpret the 14C ages in the BOOM! dataset in terms of time of deposition of tephras, in addition to the 14C age and respective analytical error, users should also consider the analyzed material (e.g., charcoal, wood, soil, organic macro remnants, planktic foraminifera), the type of analysis (micro analytical (e.g., charcoal, wood) or bulk (e.g., soil, bulk sediment)), the stratigraphic position of the analyzed material (e.g., X cm above or below the tephra), the type of section (e.g., marine core, lake core, outcrop), and whether or not replicates for the sample have been published.

4.2 Geochemical composition comparability

An important aspect when fingerprinting tephras is assessing their geochemical composition. Because many volcanic centers and eruptive events have distinct compositions, by analyzing the geochemical composition of tephras, researchers can discriminate between potential sources. To do this, tephrochronologists most often analyze the major element composition of volcanic products (~82% of the sample observations in the BOOM! dataset). However, especially in the SVZ, many volcanic centers and different eruptive events have overlapping major element compositions, thus additional information is needed to discriminate the volcanic source of tephras. Because of the latter, in some cases trace elements are also analyzed (~23% of the sample observations in the BOOM! dataset), and less frequently, isotopic ratios (for example $^{87}Sr/^{86}Sr$ and $^{143}Nd/^{144}Nd$), which provide a good discrimination tool, however are less routinely analyzed (0.1% of the sample observations in the BOOM! dataset). When comparing the geochemical composition of different volcanic products, there are several aspects to consider to understand if data from different publications are comparable or not. In the first place, it is important to understand what kind of volcanic product (pyroclastic or effusive material), and particular material (e.g., individual glass shards, bulk tephra, melt inclusions, etc.) has been analyzed. Different materials can represent different processes associated with eruptive activity, and as such, can have different geochemical compositions, even if they correspond to the same eruptive event (for more details, see section 4.2.1). Secondly, it is important to check if the data interpreted meet community quality criteria and if not, filter the data accordingly (section 4.2.2). Finally, when comparing the major element composition of different sample observations, it is important that the major element compositions are normalized to a volatile-free composition in order to compare them. Additionally, as different major elements are analyzed in different publications (for example FeO_T versus Fe_2O_{3T}), it is also important to check that the normalizations are comparable (section 4.2.3). In the following sections we describe these three aspects in detail.

4.2.1 Analyzed volcanic products

Pyroclastic and effusive material represent different expressions of eruptive activity, and because of that, even if they correspond to the same eruptive event, they will not necessarily have the same geochemical

composition. In addition, different materials can be analyzed within the same type of register, which should also be taken into consideration when interpreting the data. For example, in a pyroclastic deposit, generally individual glass shards or crystals are analyzed, which represent different magmatic processes and thus they will have very different geochemical compositions, even though they correspond to the same tephra deposit (e.g., Morgado et al., 2015). In the BOOM! dataset, the analyzed materials were entered as they were presented in the publications not to lose the raw information and as a way to make explicit the heterogeneity on how data is published. Analyzed materials within pyroclastic deposits include: tephra, bulk tephra, pumice, bulk pumice, scoria, bulk scoria, lapilli, accretionary lapilli, individual accretionary lapilli, bulk ash, bulk glass, glass shards, matrix glass, melt inclusions, closed melt inclusions, open melt inclusions, and juvenile lithics; whereas for effusive material, analyzed material comprehend: lava, prismatically jointed blocks (PJB), and breccia.

Another important aspect to note when using the geochemical composition of volcanic products for fingerprinting tephras, is what in the BOOM! dataset is called the “type of analysis”. Pyroclastic deposits and effusive material can be analyzed either micro-analytically (e.g., individual glass shards or minerals) or instead, the sample can be crushed and analyzed as a whole. The latter case is called a “bulk” analysis, and it represents a mean value of minerals, glass shards and even lithics, and most often they display narrow trends in bivariate plots. When available, micro-analytical analyses are preferred to bulk analyses because the glass (melt) composition might be more sensitive to small amounts of fractional crystallization and thus may reflect a more variable geochemical composition. Thus, providing more options for fingerprinting tephras, for example, of different eruptions from the same volcanic source (Lowe, 2011). When bulk analyses are performed, on the other hand, the composition of different eruptive events and even different volcanic centers can be very similar and does not always allow to discriminate the source of volcanic products.

In order to correctly interpret the geochemical composition of eruptive events and volcanic centers in the BOOM! dataset, the user should take into consideration: the type of register (effusive material, pyroclastic material), type of analysis (micro-analytical, bulk) and the analyzed material (e.g., lava,

pumice, wood). These qualitative attributes can help users understand if the geochemical compositions they are interpreting represent similar processes or not, and thus if they can be directly compared.

4.2.2 Quality control of analyses

Generally, when researchers obtain the results from the geochemical analyses they perform (but also when collecting information to compare their samples with), they check if the data meets some community criteria, to understand if the analyses are *good* analyses. Among these: if the analysis was correctly performed, if what is said to be analyzed was in fact analyzed, and if the accuracy and precision of the analysis is within an acceptable range. In most cases, researchers will publish only the data that meet these criteria. However, different researchers, disciplines, or fields of application, can follow slightly different criteria. Thus, it is important for the users to have access to the information necessary for them to evaluate the latter, so they can filter data according to their own criteria. In the BOOM! dataset, we have included several attributes for users to evaluate this, which we describe in the following.

A first order quality control of analyses of geochemical composition is a check of the analytical conditions in which they were performed. Depending on the technique employed, this information is more or less routinely communicated. Because most of the BOOM! dataset corresponds to EMPA analyses of glass shards (~76% of sample observations), we have included only the analytical conditions for analyses performed by EMPA, which correspond to beam size, accelerating voltage, beam current of a particular measurement run, stored in the MeasurementRuns file. Common values are described in Table S3.

An additional important aspect to evaluate when assessing the quality of a geochemical analysis, particularly of major element analyses, is that the analytical total is within accepted values. Major (and minor) elements correspond to the geochemical elements found more abundantly in igneous rocks, and their abundances are typically expressed in weight percentage (wt.%) of oxides: SiO_2, TiO_2, Al_2O_3, FeO and/or Fe_2O_3, MnO, MgO, CaO, Na_2O, K_2O, P_2O_5. Additionally, in some cases Cl and F are also analyzed and considered with minor elements. The sum of the analyzed major (and minor) elements composition is called the “analytical total” (“Total” in the BOOM! dataset), and typically is between 90–100 wt.%. The difference between the analytical total and a 100 wt.% can derive from different situations, among them:

561 analytical error (including due to surface roughness of the samples, poor sample positioning, and loss of
562 alkalis during the analysis), from post-depositional hydration (e.g., alteration of the volcanic products),
563 or from the presence of non-degassed volatiles in the groundmass of the volcanic products. For a more
564 detailed discussion on *acceptable* analytical totals depending on the volatile content of the volcanic
565 products, the reader is referred to Pearce et al. (2008). The value of the analytical total that is considered
566 as a *good* analysis also depends on the analytical technique employed, which is associated with the type
567 of analysis performed (bulk or micro analytical). On the one hand, when bulk samples are analyzed, for
568 example, by X-ray fluorescence (XRF), Atomic Absorption Spectroscopy (AAS), or Inductively Coupled
569 Plasma Emission Spectrometry (ICP-AES), in addition to the major element composition, a value called
570 Loss on ignition (LOI) is sometimes estimated. The LOI corresponds to the volatile content of the sample,
571 for example H_2O, CO_2, SO_2, which can derive from different degrees of post-depositional hydration or syn-
572 eruptive characteristics of the volcanic products. In this case, the analytical total plus the LOI should
573 approach 100 ± 2 wt.%, otherwise it is considered a *bad* analysis. In the dataset, ~1.9% of the sample
574 observations corresponding to bulk analyses of major elements of pyroclastic or effusive material, are
575 outside this range, which corresponds to ~0.2 % of the sample observations in the whole BOOM! dataset.
576 On the other hand, when glass shards are analyzed by EMPA, the volatile content is usually not directly
577 analyzed. In this case, analytical totals lower than 100 ± 2 wt.% do not necessarily reflect poor quality
578 analyses or alteration of the samples, but more often reflect characteristics of the volcanic deposit and
579 environment, which are not accounted for by the major elements alone. Because of the latter, there is no
580 consensus regarding acceptable totals for EMPA glass data, and thus, in the dataset this value is stored
581 when presented in the literature (~86% of the sample observations analyzed for major elements in the
582 BOOM! dataset have been published along with their analytical totals) for the user to judge for
583 themselves. For ~95% of the sample observations analyzed for major elements with EMPA, the analytical
584 total is between 95 and 102 wt%. Samples with analytical totals lower than 90 wt% were not included in
585 the dataset.

586 Another important quality control researchers perform is making sure that what was intended to be
587 analyzed, actually was. To do this, they generally check that the geochemical composition is within

588 plausible ranges. In the case of bulk analyses, this means that the composition is magmatic, i.e., that SiO_2
589 is between ~40–80 wt.%. In the case of micro-analytical analyses, it depends on what it is analyzed,
590 generally glass or mineral crystals. Sometimes, the groundmass in pumice, scoria, individual glass shards,
591 etc., can be very microlite-rich, particularly for relatively mafic samples (e.g., of basaltic andesite
592 composition) and so it can be difficult to analyze the glass phase using a defocused beam by EMPA or LA-
593 ICP-MS; on the other hand, when crystals are analyzed, if they are very small (<20 μm) it might be difficult
594 to analyze their composition at specific locations (e.g., core vs. rim). In order to check if the analyses were
595 performed on the desired material, researchers check the composition of major elements and evaluate
596 whether it is feasible for the expected phase (e.g., glass vs. a mineral phase). In the BOOM! dataset, data
597 was included as long as it was magmatic and indicated as either bulk tephra or individual glass shards in
598 the original publication. The user can choose to filter the information based on their own criteria.

599 One of the most robust ways of reducing the uncertainties in correlations based on the geochemical
600 composition of tephras is side-by-side analysis. This means, analyzing in the same laboratory and with
601 the same methodology unknown and reference tephras (Lowe et al., 2017). In the BOOM! dataset, this
602 has strictly been done in only publication (Smith et al., 2019). Smith et al. (2019) reanalyzed reference
603 tephras previously identified on land by Stern (1992, 2008) and by Weller et al. (2015), to check for
604 potential correlations with the tephras identified by them in a lake sediment core. On the contrary, most
605 publications a side-by-side analysis is not done. However, in many publications unknown tephras
606 identified in several sections are analyzed which allow authors to robustly correlate them and
607 characterize eruptions methodologies (e.g., Rawson et al., 2015; Fontijn et al., 2016; Alloway et al., 2017a,
608 b). These analyses are done in the same laboratory and following the same methodologies, reducing the
609 uncertainty in their correlation. In general though, samples which have been analyzed in previous
610 publications are not reanalyzed, probably because of the increase in expenses or complications obtaining
611 the samples. Alternatively, one important way of evaluating if analyses performed in different
612 laboratories are comparable, as well as of evaluating their accuracy and precision (see below), is
613 analyzing secondary standards along with the unknown samples and reporting the results.

An important evaluation to make to correctly interpret the available geochemical information, is assessing the accuracy and precision of the analyses. In the case of isotope ratios included in the BOOM! dataset ($^{87}Sr/^{86}Sr$ and $^{143}Nd/^{144}Nd$), an analytical error (generally 2σ) is directly provided by the laboratories and is included in the dataset, which allows users to understand the analytical precision of that data. In the case of major and trace elements, however, assessing this is less straightforward. The most appropriate way to assess the accuracy and precision of these analyses, is to consider the composition of the secondary standards analyzed along with the samples, as described in section 3. From the total of sample observations analyzed for major and trace elements, for ~64% (10,438 of the sample observations), it was possible to obtain the full secondary standard information (i.e., name of the standards analyzed in each measurement run, number of analyses performed during each measurement run, analyzed mean value and standard deviation for each analyzed element for each standard). In ~2.8% of the cases (453 sample observations) only the mean analyzed values were published. Thus, for ~33.2% of the sample observations analyzed it is not possible to evaluate the accuracy and precision of the analyses. Additionally, from the 33 secondary standards analyzed in the publications included in the BOOM! dataset, certified values are available for only 19 of them, which corresponds to ~55% of the sample observations, from which ~53% has the full information (mean, standard deviation and n). Eight of the certified secondary standards correspond to glass: GOR128-G, GOR132-G, KL2-G, ML3B-G, T1-G, StHs6/80-G, ATHO-G (Jochum et al., 2006), NIST SRM 610 (Jochum et al., 2011), and eleven to whole rock powders AGV-2, BCR-2, BHVO-1, BHVO-2, BIR-1, JA-1, JA-2, W-2 (Jochum et al., 2016), S-Y[11], OREAS184[12], OREAS700[13]. For the whole rock analyses included in the dataset, only the mean values were published.

In order to provide a broad idea of the accuracy and precision of the analyses included in the BOOM! dataset, we have estimated the accuracy and precision of the analyses for which the authors provided the full information, and the certified values are available online. The accuracy is estimated as the mean analyzed value versus the certified value of each element analyzed on each secondary standard during

[11] Certified values available at: https://www.nrcan.gc.ca/our-natural-resources/minerals-mining/mining-resources/sy-4-diorite-gneiss/8025
[12] Certified values available at: https://www.oreas.com/crm/oreas-700/
[13] Certified values available here: https://www.oreas.com/crm/oreas-184/

each measurement run, and the precision as the relative standard deviation (RSD) of the *n* analyzed values of each element on each secondary standard during each measurement run (RSD = (1SD/mean)*100). Figure 4a and b show the results of these estimations, grouped by secondary standard. Jupyter notebooks written in Python are available in the github repository[14] of the project for users to explore the overall dataset accuracy and precision, as well as for each measurement run or sample. Overall, the highest accuracies and precisions are observed for SiO_2, which range between ~0.96 and ~1.02, and ~0.09% and 1.54%, respectively; and the lowest accuracies and precisions are observed for P_2O_5, which range between ~0.16 and ~2.00, and between 3% and 206%, respectively.

As mentioned in section 3, more than one secondary standard is generally analyzed by run, which are intended to represent the geochemical composition of the studied samples. Because of that, more than one value for accuracy and precision for each analyzed element is associated to each batch (for example as in Figure 5a and b). In this case, researchers will generally evaluate the accuracy and precision of a sample observation by comparing it with the secondary standard(s) with the geochemical composition closest to that of the unknown sample. Thus, the precise analytical error is not calculated, rather it is generally estimated as the highest value of precision and accuracy for each element. For example, along with sample CLD025A (~55–57 wt.% SiO_2), analyzed during measurement run 11112 by Fontijn et al. (2016), secondary standards GOR132-G (45.5 wt.% SiO_2), KL2-G (50.3 wt.% SiO_2), T1-G (58.6 wt.% SiO_2), StHs6/80-G (63.7 wt.% SiO_2) and ATHO-G (75.6 wt.% SiO_2) were analyzed. In this case, standards KL2-G and T1-G have the composition closest to the sample. During that run, the analysis of TiO_2, for example, is estimated as ~5% for standard KL2-G, and ~3% for standard T1-G, whereas precisions are 2% and 5%, respectively. Thus, the estimated accuracy and precision for the analysis of TiO_2 are both ~5%.

Users should take into consideration all of the above when interpreting the data and the dataset, and filter regarding their own criteria.

[14] https://github.com/consuelola/BOOM/Notebooks

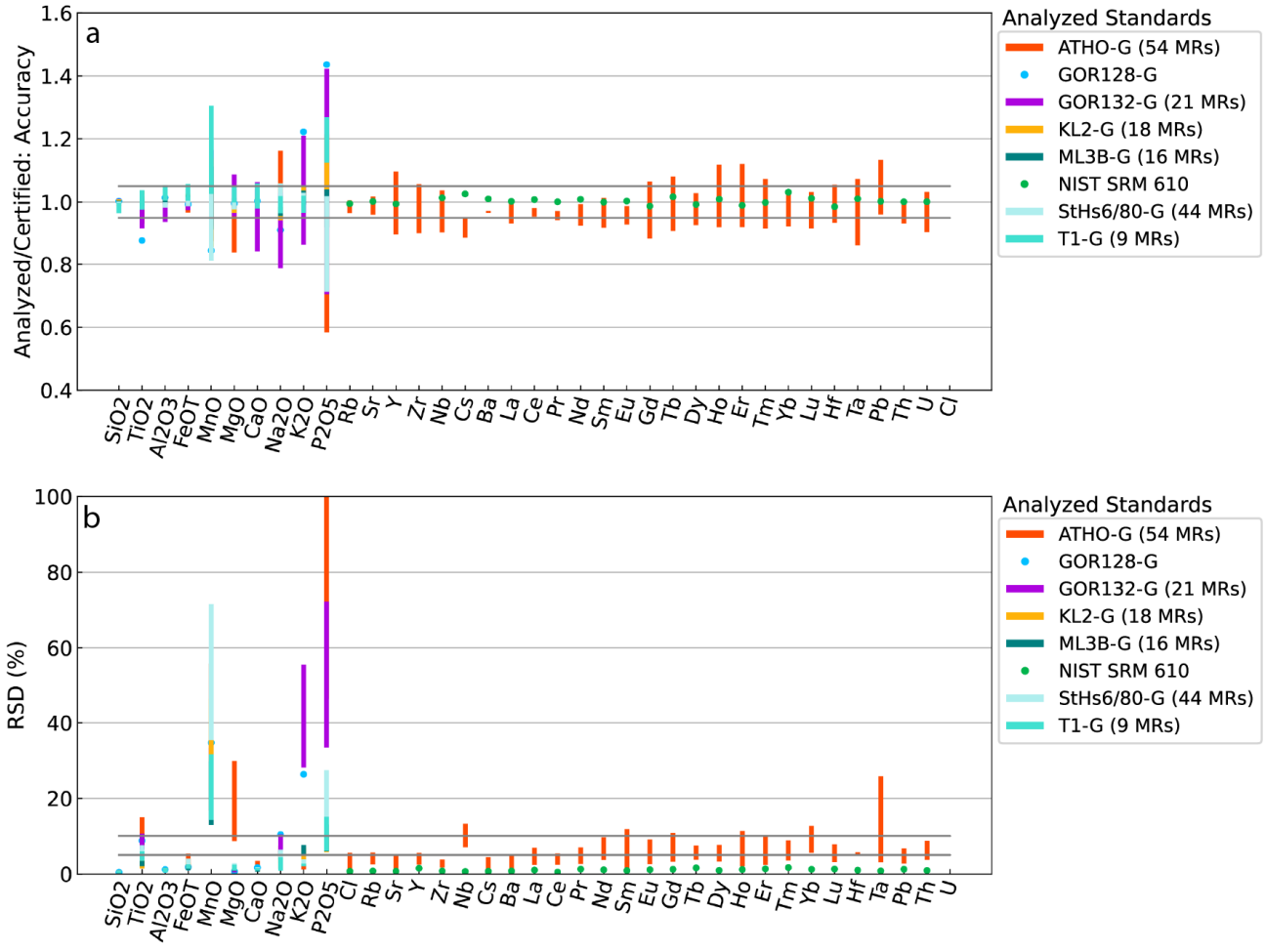


Figure 4. Accuracies (a) and precisions (b) of the geochemical composition analyses included in the BOOM! dataset, estimated as described in the text. When the same secondary standard was analyzed in more than three measurement runs, the 1σ range are plotted as a vertical line and the number of measurement runs considered is indicated in the legend as #MRs. When a secondary standard was analyzed during three or less measurement runs, accuracy and precision are plotted as a dot. Bold grey horizontal lines correspond to 0.95 and 1.05 values for the accuracy and 5 and 10% for the precision. Only secondary standards with certified values are shown, analyzed during measurement runs for which the name of the secondary standard for which the full information is published (as described in the text): ATHO-G, GOR128-G, GOR132-G, KL2-G, ML3B-G, T1-G, StHs6/80-G, (Jochum et al., 2006), NIST SRM 610 (Jochum et al., 2011).

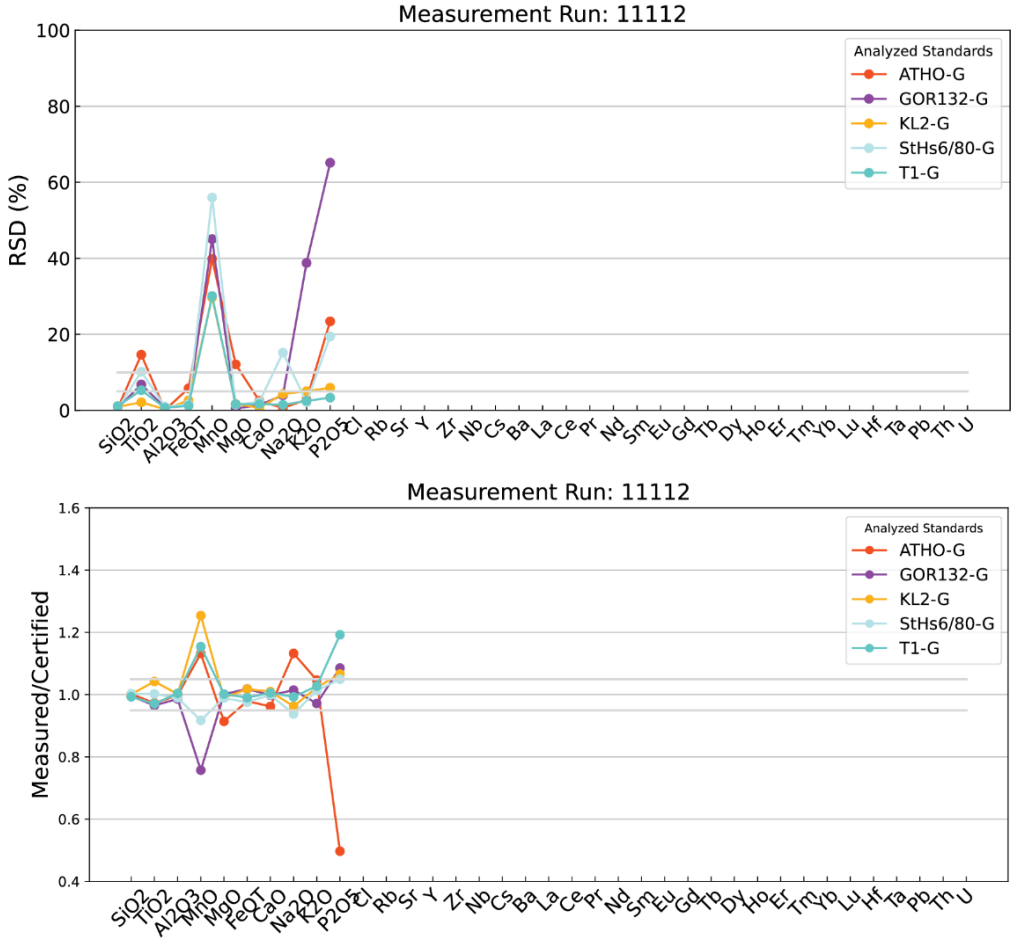


Figure 5 Accuracies and precisions of the geochemical composition analyses of each secondary standard analyzed during measurement run 11112 (Fontijn et al., 2016), as an example. a: Accuracy for major elements, calculated as mean analyzed value of each element of each secondary standard analyzed during measurement run 11112, versus the certified values (Jochum et al., 2006). b: Precision of the geochemical analyses, calculated as mean analyzed value of each element of each secondary standard analyzed during measurement run 11112, divided by the standard deviation of the n analyses during that measurement run. Bold grey horizontal lines correspond to 0.95 and 1.05 values for the accuracy and 5 and 10% for the precision.

4.2.3 Major element normalization

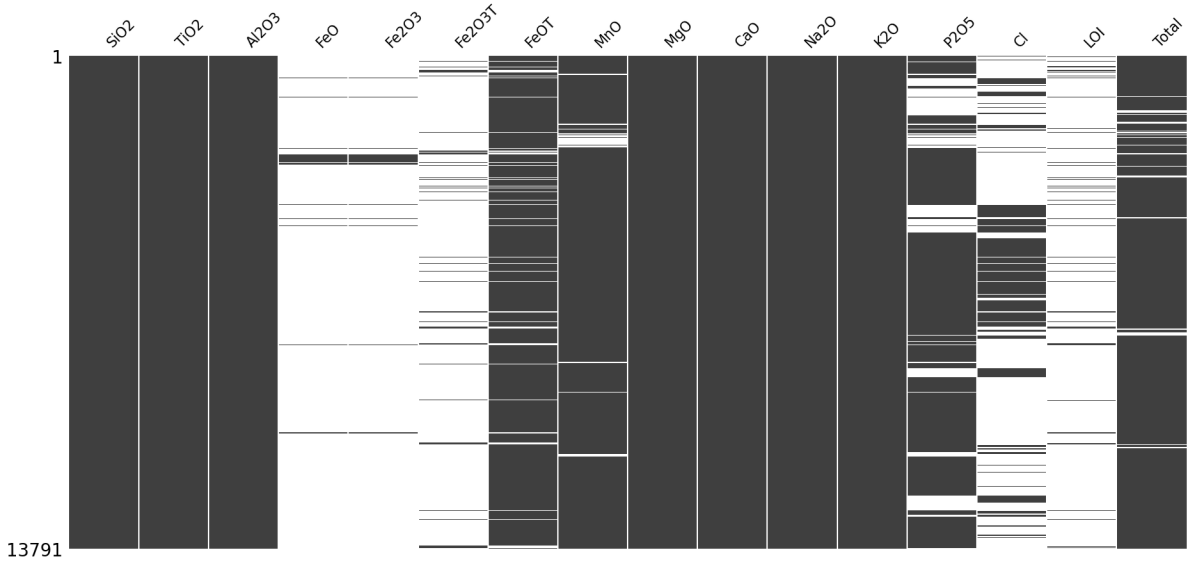
As already mentioned, the major element composition of volcanic products is expressed in wt.%, and ideally the sum of all the major oxides and LOI (when analyzed) should approach 100 wt.%. However, because of post-depositional alteration, variable volatile contents in magmas, and analytical errors, the analytical totals are generally not exactly 100 wt.%. In order to eliminate the effects of these processes in the composition of oxides, researchers normalize the major element composition to a 100% volatile-free (also named anhydrous) basis, i.e., by the sum of the major oxides without LOI (when analyzed). In order for users to be able to directly compare the major element compositions of sample observations in the

BOOM! dataset, both the *raw* and normalized major elements are stored. Because of the heterogeneities in the data sources in the dataset, when normalizing major elements two aspects should be taken into consideration, which we describe in the following.

In the first place, in order to normalize the major element compositions of samples by a comparable amount, special attention must be paid to how iron (Fe) was reported. Fe can exist in volcanic rocks in two states of oxidation (Fe^{2+}, as FeO; and Fe^{3+}, as Fe_2O_3), which exist in different proportions in different types of rocks. Most analytical techniques employed to analyze volcanic rocks are not able to distinguish between these two states of Fe, and thus it is presented in most publications as a *total* Fe content. For example, when individual volcanic glass shards are analyzed by EMPA, Fe is usually reported as total Fe as FeO, which might be communicated by authors as FeO, FeO_T or FeO*. In contrast, when bulk samples are analyzed, for example by XRF or Inductively Coupled Plasma - Optical Emission Spectrometry (ICP-OES), Fe is typically reported as total Fe as Fe_2O_3, and it might be presented in the publication either as Fe_2O_3 or Fe_2O_{3T}. Additionally, other techniques such as AAS and wet chemical analysis (titration) can discriminate between Fe^{2+} and Fe^{3+}, however they are not commonly used. In this case, publications calculate the total Fe as $FeO_T = FeO + 0.899*Fe_2O_3$, or as $Fe_2O_3T = 1.1*FeO + Fe_2O_3$. Thus, when normalizing major element compositions of heterogeneous sources, it is important to calculate the analytical total using the same expression for Fe, as using one or the other will result in different normalized compositions. In the BOOM! dataset, ~90% of the major element compositions were analyzed by EMPA. Because of this, the normalization in the dataset is done considering Fe as FeO_T for calculating the analytical total. When FeO_T was not directly presented in the original publication, it was calculated either as: $FeO + 0.899*Fe_2O_3$, when both FeO and Fe_2O_3 were analyzed (~3% of major element observations); or as $0.899*Fe_2O_{3T}$, when Fe was analyzed as Fe_2O_{3T} (~5% of major element observations). To save the original data and its heterogeneity, Fe is described in the dataset by four attributes: FeO, Fe_2O_3, FeO_T and Fe_2O_{3T}, and only the original data is filled in the non-normalized (raw) version of the dataset.

Second, different publications might analyze different sets of major elements and thus, samples might not be normalized by a comparable analytical total. In particular, elements found in lower concentrations

(such as P_2O_5, MnO or Cl) are not always analyzed (Figure 6), because of the increased amount of time required to analyze them with reasonable precision, versus the information they provide (when using EMPA). However, because of their low concentrations, whether these elements are analyzed or not, will have a small influence on normalizations. We have performed a small test in order to check if the normalized values of samples where different major elements have been analyzed are still comparable. We re-normalized the major element compositions of sample observations where all major elements included in the BOOM! dataset were analyzed by a new analytical total, calculated as the sum of all major elements analyzed, excluding Cl, P_2O_5 and MnO. Then, we calculated the variation between the normalized and the re-normalized compositions of each element (e.g., $SiO_{2\ renormalized} - SiO_{2\ normalized}$) for each sample observation and compared it to the maximum achievable precision for that element. For the latter, we considered the 95% CL of the certified values of the secondary standards analyzed along with the samples (Figure 7). We only tested sample observations for which both the analytical total are included in the BOOM! dataset and the secondary standards analyzed have certified values (~61% of sample observations analyzed for major elements). Because the secondary standards used for whole-rock and glass analyses are different, we compare the sample observations with their respective secondary standards. In all cases, except for SiO_2, Al_2O_3, and FeO, the difference of the analyzed composition between the normalizations is lower than the 95%CL of the geostandards analyzed along with the samples (Figure 7, S1, S2). Nevertheless, in the case of SiO_2, Al_2O_3, and FeO, ~99% of the sample observations the difference is within or lower than the 95% CL of the standards for the range of compositions of the sample observations. For Al_2O_3, the difference between the two normalizations is lower than 0.13 wt.% in~ 99% of the cases, which is lower than the 0.2 wt.% analytical error of Al_2O_3 for most secondary standards analyzed along with the samples, except for BHVO-2G, for which it is 0.1 wt.%. Thus, even though in the BOOM! dataset sample observations are included for which different major elements where analyzed, their normalized compositions are still comparable.



738

739 **Figure 6** Heterogeneity of the analyzed major elements contained in the BOOM! dataset. The rows
740 correspond to the sample observations. Dark grey cells represent "filled" values and white cells represent
741 "missing" values. In this case, 13,791 sample observations have major element data, but not all elements
742 have been analyzed for each sample. The figure was created using the Python library Missigno.

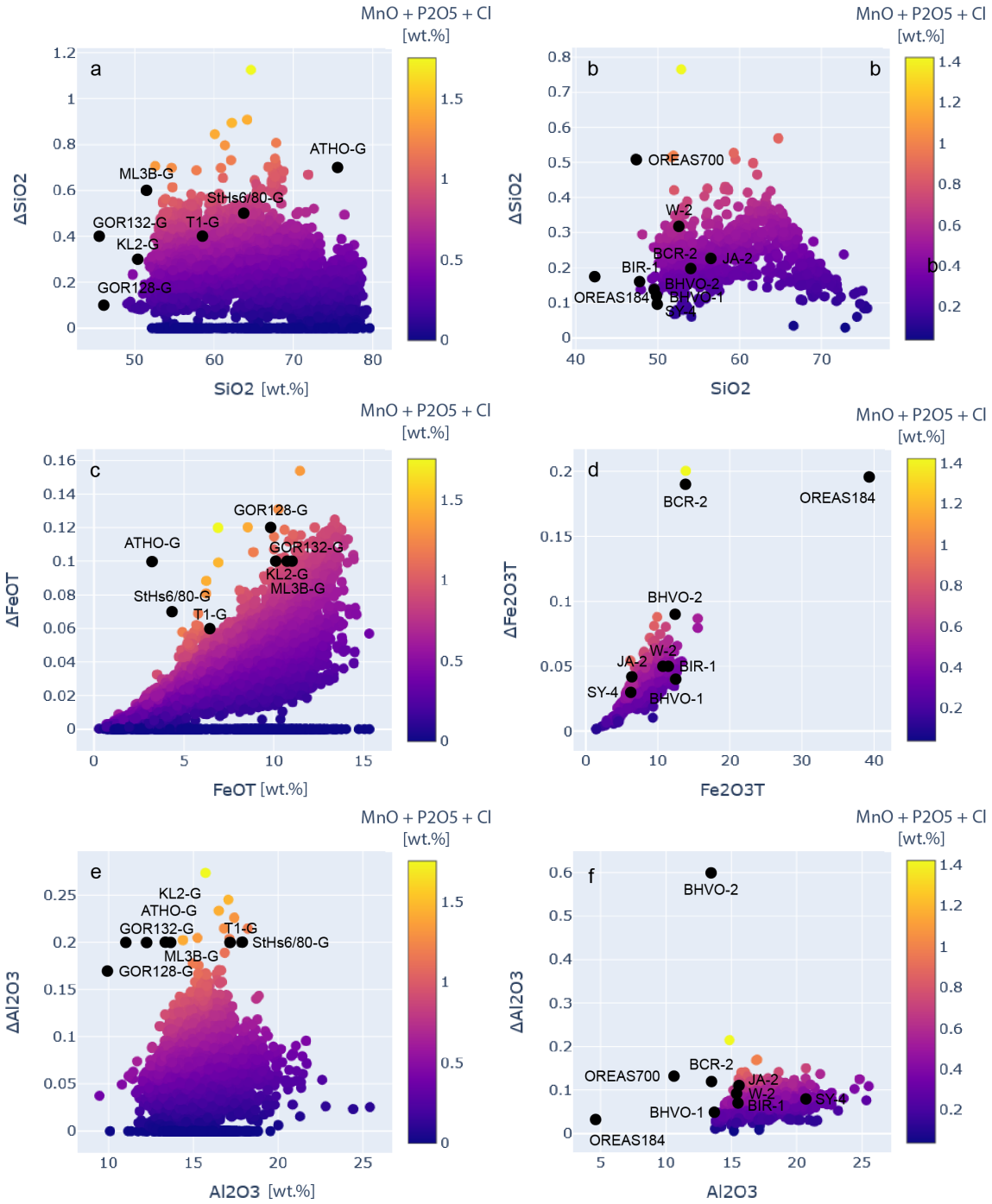


Figure 7. Impact of normalizing major element compositions of sample observations in the BOOM! dataset by different analytical totals. Δ: Normalized value considering the analytical total as the sum of all analyzed major elements - Normalized value considering the analytical total withouth Mn0, P_2O_5 and Cl. a, c, e: glass shard analyses; b, d, f: whole-rock analyses. Only samples for which the anaytical total is published in the original publication are plotted. The 95% CL of the certified secondary standards analyzed along with the samples are plotted as a reference of the highest achievable precision of the analyses. GOR128-G, GOR132-G, KL2-G, ML3B-G, T1-G, StHs6/80-G, ATHO-G and T1-G (Jochum et al., 2006), BCR-2, BHVO-1, BHVO-2, BIR-1, JA-2, W-2 (Jochum et al., 2016), SY-4 (https://www.nrcan.gc.ca/our-natural-resources/minerals-mining/mining-resources/sy-4-diorite-gneiss/8025), OREAS184 (https://www.oreas.com/crm/oreas-184/), OREAS700 (https://www.oreas.com/crm/oreas-700/).

5 Using the BOOM! Dataset to explore potential correlations

The BOOM! dataset and explorer were developed to help in the difficult task of disentangling the tephrochronological record in the SVZ and AVZ, which in practice means robustly correlating tephras deposited during the same eruptive event. The latter is generally done by fingerprinting them, i.e., identifying unique characteristics that distinguish tephra deposits. Doing this is often quite difficult, because of the multidimensionality of the tephrochronological information, as communicated during the interviews (Table S1). Because of this, researchers often use different tools to explore and visualize different type of data (GIS software, Excel, R, Adobe Illustrator, etc.), and so, characterizing each volcanic center and eruptive event can become very tedious, especially when dealing with large amounts of data, such as in the BOOM! dataset. With this in mind, the BOOM! explorer (https://boom-vis.lisn.upsaclay.fr/) was created as a complementary tool to the dataset. By using it, users can visualize at the same time the geographical distribution of volcanic centers and the volcanic products originated during specific eruptive events, as well as their geochemical composition, and the available information to assess the chronologies of the eruptive events (Figure 8). The BOOM! explorer can at the same time be used to compare the available data for each volcanic center and known eruptive event and evaluate the "finger-printability" of the tephras. In section S6 of the supplementary material we exemplify how the BOOM! explorer can help users in correlations with an example. It is important to note that the BOOM! explorer is not meant to be a comprehensive tool, and users can always download the dataset, add supplementary information, and produce different visualizations fitted to their specific needs; as well as download the source code of the explorer and make changes fitted to their needs. In this regard, the explorer can also be used as a catalog of the data in the BOOM! dataset. By using it, users can perform a first inspection of the dataset, understand the amount and type of information available for each volcanic center and eruptive event, and chose which information to download. Additionally, we have produced a Jupyter notebook called Correlations, available in the github[15] repository, which users can use to plot their own unknowns and compare them with the BOOM! dataset.

[15] https://github.com/consuelola/BOOM/assets/Notebooks

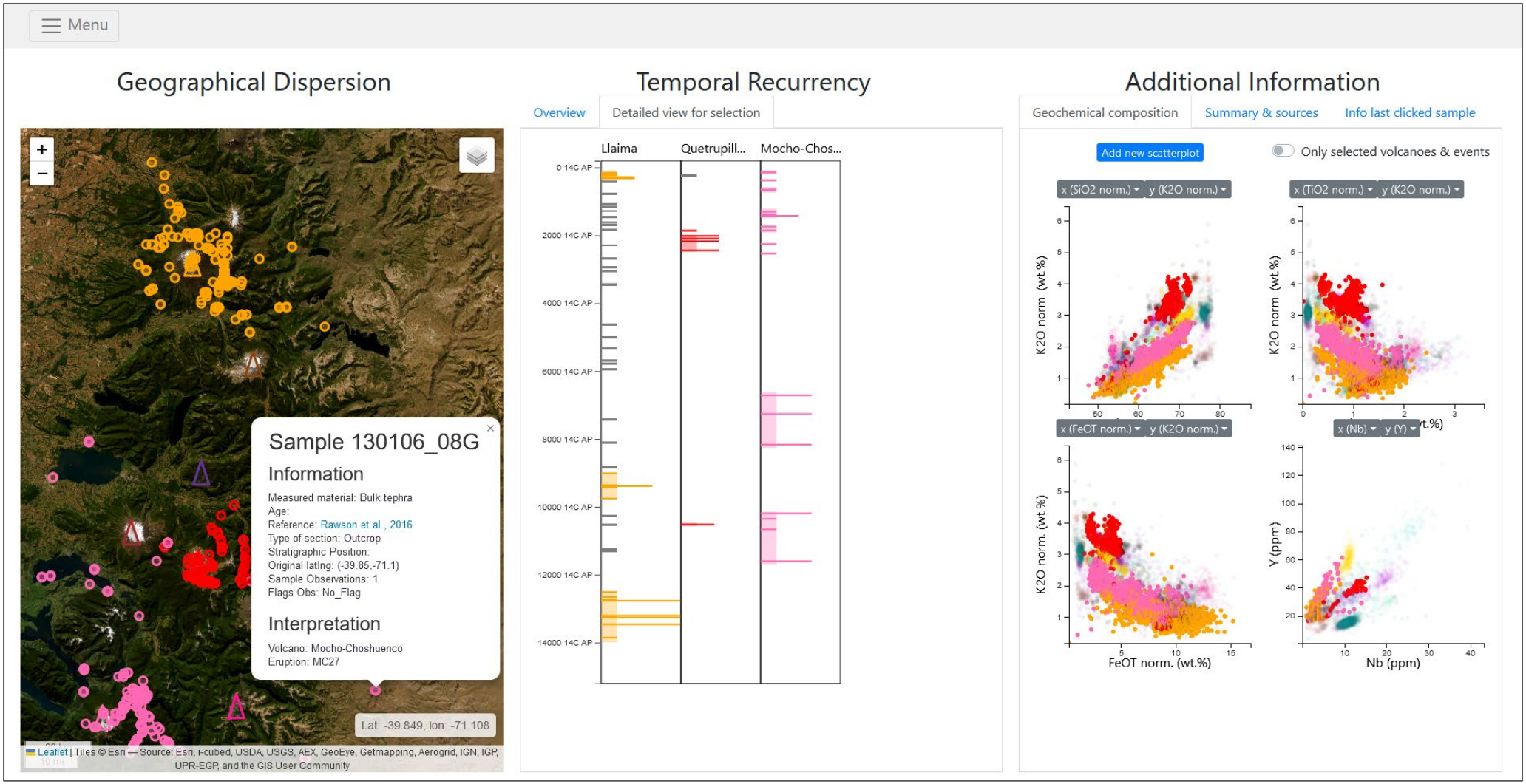


Figure 8 Screenshots of the sections of the exploration tool (https://boom-vis.lisn.upsaclay.fr), showing different visualizations of the data in the dataset. In the figure, volcanic centers Llaima, Quetrupillán, Mocho-Choshuenco have been selected. For more details on how to use the explorer, the reader is referred to section S6.

Even though being able to visually compare tephrochronological data in its multidimensionality is very helpful, sometimes it is not enough to robustly distinguish tephras. The latter is especially true in areas where the geochemical composition of many volcanic centers and eruptive events is very similar, as is often the case in the SVZ and AVZ. This is worsened by having large amounts of data from different eruptive events. The latter is not uncommon, and because of it, statistical tools are sometimes used by tephrochronologists to find ways to more robustly fingerprint tephras (Lowe et al., 2011; 2017). In section 5.1, we briefly evaluate the use of geochemical composition together with machine learning algorithms to classify samples in the dataset labeled as *unknown* volcano. This exploration is not aimed to be exhaustive either, but rather to provide a starting point and encourage the use of the BOOM! dataset to explore the application of machine learning for tephra classification.

5.1 Machine learning application

The BOOM! dataset offers a great opportunity for the application of machine learning in the discipline of tephrochronology. To demonstrate this potential, we trained models on the BOOM! dataset to perform automatic tephra correlation. In machine learning terminology, this corresponds to a classification task, where the input are sample observations describing the geochemical composition of tephras, and the outcome to predict is the volcanic system which originated the tephra deposit. For simplicity, in this experiment we classified each sample observation rather than the sample as a whole, which can be done, for example, by a majority vote (e.g., Bolton et al., 2020). In practice, when the geochemical composition of more than one sample observation is available for one tephra deposit, researchers consider the geochemical trend described by all the sample observations to classify the sample. Thus, future applications should take this into consideration.

5.1.1 Preprocessing

Prior to training the models, a few preprocessing steps are necessary. First of all, only sample observations for which the volcanic source has been identified were considered, which corresponds to ~89% of the dataset. This is necessary because the volcanic source “labels” are needed both for training the models and for evaluating their performances. From the latter, we discarded sample observations for

which the volcanic source is debated in different publications (sample observations with a VolcanicSource_Issue flag, corresponding to 86 sample observations). Additionally, we only considered sample observations for which the major or trace elements composition of the volcanic products were analyzed, i.e., sample observations which correspond to ^{14}C ages of organic matter were ignored. From these, we discarded sample observation which are geochemical composition outliers (31 sample observations), those which have analytical totals lower than 95 wt.%, and those with LOI higher than 5 wt.%, as they might correspond to altered samples (3,242 sample observations, for more details the reader is referred to section 4.2.2). After this treatment, we only kept sample observations from volcanic centers left with more than 10 sample observations, from at least two different samples. This is required in order to have sufficient data for training and testing the models. The latter left a total of 13,925 sample observations from 2,167 samples from 27 different volcanic centers. As a result of the preprocessing, the volcanic centers Cay, Macá, Yanteles, Corcovado, Cordón Cabrera and Subsidiary Vcha dome were not included in the dataset. The whole preprocessing is done in the preprocessing function in utils.py, available in the github repository[16]. We call the resulting dataset the majors *or* traces dataset, as it includes all sample observations for which either major elements, trace elements or both have been measured.

5.1.2 Methods

We trained four machine learning classifiers or estimators: Logistic Regression, k-Nearest Neighbors (kNN), Random Forest (RF) and Gradient Boosting (GB). These classifiers were trained with volcanic centers as target values and major and trace element compositions as features. Because of the high percentage of missing data (~62 %), imputation of the missing values was performed prior to model fitting for all models, except for Gradient Boosting which natively handles missing values. Four different imputation methods were considered: mean imputation (with sklearn's SimpleImputer) which imputes a constant value, in this case the mean value of each feature; k-Nearest Neighbors imputation (with sklearn's KNNImputer) which imputes by the mean value of the nearest neighbors, in this case 15

[16] https://github.com/consuelola/BOOM/Machine_learning/src

neighbors; iterative conditional imputation (with sklearn's IterativeImputer) relying on either a Bayesian Ridge (BR) regressor or a Random Forest regressor. This method is related to the well-known imputation method MICE (Multiple Imputation by Chained Equations, van Buuren, 2018). It consists in imputing features iteratively, in a round-robin fashion, using a model (in our case one of BR or RF) that takes as target the feature to impute, and as input data of all the other features. This results in a total of 13 imputer-estimator pairs tested. For all imputation algorithms and models, we used the scikit-learn[17] python library (Pedregosa et al., 2011). The code is available on the run_experiments.py function on github.

The BOOM dataset is subject to class imbalance, i.e., the number of sample observations for each volcanic center varies greatly. For example, six of the 27 volcanic centers are represented by less than 50 sample observations in the dataset (Yate, Apagado, Hornopirén, Huequi, Lautaro and Aguilera), while five volcanic centers are represented by more than 1,000 sample observations (Llaima, Quetrupillán, Mocho-Choshuenco, Puyehue-Cordón Caulle, and Chaitén). To account for this class imbalance, the performances of the thirteen models were evaluated by comparing both their accuracies and balanced accuracies. The accuracy indicates the proportion of correctly classified sample observations over the entire test set, whereas the balanced accuracy weights sample observations according to the inverse prevalence of their class (in this case, the volcanic center). Balanced accuracy therefore gives more weight to rarer classes than accuracy, and allows a better understanding of whether these rarer classes tend to be well-classified or not. It may, however, be affected by a higher variance: because of the little data available for some of the volcanic centers (for example, 11 sample observations for Huequi) it can be more sensitive to changes in the hyperparameters of the models chosen in each fold, to the randomness of the imputing mechanism, and to the train-test divisions of the sample observations.

For model evaluation, we performed a 10-fold cross validation, which gives us 10 performance values (here accuracy and balanced accuracy) for each imputer-estimator pair tested. On each fold, a grid search is performed on the training set with an inner 5-fold cross-validation to select the best hyperparameters

[17] https://scikit-learn.org

for the estimator. The model is then retrained with the hyperparameters identified on the whole train set before being evaluated on the test set. Note that we optimized the hyperparameters of the estimators but used default hyperparameters for the imputation algorithms to alleviate the computational burden.

As many samples are composed of several sample observations, special care was taken not to separate sample observations belonging to the same sample in the train and test sets. This is very important, as in practice, either the volcanic source is known for all sample observations in a sample, or unknown for all of them. Moreover, sample observations coming from the same sample are most likely to be more similar than sample observations coming from the same volcanic source in general. The failure to keep together all sample observations from a given sample in the same set would thus not reflect a practical use case and artificially inflate performances. To ensure that we do not split sample observations inappropriately, we used a custom GridSearchCV_with_groups object as the usual GridSearchCV object from scikit-learn does not allow yet to specify groups.

Model training was expensive in terms of memory (~228GB) and processing time (~24 hours on 30 cores), mainly due to the iterative conditional imputation with RF regressor which used a lot of memory and took two orders of magnitude longer than all other imputation schemes.

5.1.3 Results

The mean, standard deviation (SD), and maximum error of the ten accuracies and balanced accuracies for each model are shown in Table 1 and Figure 9, and the detailed results for each fold can be accessed in the github repository in the results/major_or_traces folder. Additionally, the confusion matrix, feature importances, and bivariate plots to visualize the classifications were produced for each of the thirteen models and are also available in the github repository, in the figures/major_or_traces folder.

The BOOM! dataset allows the learning of successful models for tephra correlation - The best model obtained is the iterative conditional imputation with BR regressor and RF as prediction model. It achieves 92% (SD: 3) accuracy on average over the cross-validation folds, and 79% (SD: 6) balanced accuracy. This performance highlights a clear signal in the geochemical data. While the successful application of machine learning has been demonstrated in previous studies (Bolton et al., 2020; Pignatelli & Piochi, 2021; Uslular

et al., 2022), it is the first time that it is demonstrated at this scale (13,925 sample observations from 2,167 samples and 27 different volcanic centers) on heterogeneous data. Indeed, the geochemical analyses were performed by many different laboratories, and many sample observations have been analyzed for either major or trace elements, which can affect the performance of machine learning models. It is thus positive and encouraging that models good enough to be of practical interest (>90% accuracy) can be learned from this data set. Note that here we focused on a few classical and effective models, without trying to obtain the best possible performances. It is thus very likely that improvements are possible, for example by using ensembling techniques, or by improving the handling of missing values.

Table 1. Performances of the thirteen models tested in the major *or* traces set, i.e., the dataset containing sample observations on which either major elements, trace elements or both have been measured. For each of the imputer-estimator pairs, the mean accuracy and balanced accuracy of the 10-folds of the cross-validation are presented. Imputers are indicated in the columns and estimators in the rows LR: Logistic Regression; kNN: k-Nearest Neighbors; RF: Random Forest; GB: Gradient Boosting; BR: Bayesian Ridge.

			Iterative conditional imputation with		
	Simple imputer with mean	**kNN Imputer**	**BR regressor**	**RF regressor**	
LR	0.74/0.71	0.75/0.69	0.76/0.72	0.70/0.67	
kNN	0.87/0.76	0.85/0.73	0.90/0.80	0.83/0.71	
RF	0.90/0.77	0.87/0.73	0.92/0.79	0.86/0.73	
GB					0.90/0.75

The imputation algorithm matters – Our results (Figure 9a and 9c) show that the choice of imputation algorithm plays an important role in the performance of the models, both with regards to accuracy and balanced accuracy. Interestingly, the ranking of imputation methods remains similar across models. The kNN Imputer and iterative conditional imputation with RF led to the worst performances, while the mean imputation and iterative conditional imputation with BR had the best performances. This is surprising as

905 iterative conditional imputation with RF is a state-of-the-art imputation method. This situation might be
906 related to the structure of the "missingness" patterns (Figure S5). The dataset has a block structure, with
907 10,305 sample observations for which only the major elements have been analyzed, 2,408 sample
908 observations for which only trace elements have been analyzed, and "only" 1,212 sample observations
909 for which both major and trace elements have been analyzed. In order to impute the major elements in
910 the "traces only" block and conversely, imputation algorithms need to learn the relationships between
911 the majors and traces from the sample observations for which both have been analyzed, which represent
912 here a relatively small fraction of the data. This is a rather difficult setting, which may explain that mean
913 imputation outperforms iterative conditional imputation with RF or the kNN imputer. In contrast, the
914 iterative conditional imputation with BR regression outperforms mean imputation for all models.
915 Previous work using machine learning for identifying the volcanic source of volcanic products have either
916 not had any missing values in their dataset (Bolton et al., 2020) or have solely used a simple imputation
917 with either the mean (Pignatelli et al., 2021) or with zero (Uslular et al., 2020). Our results highlight the
918 importance of having good imputation schemes to improve overall performances, especially with
919 datasets compiled from very heterogeneous data, such as the BOOM! dataset. In what follows, we
920 consider the accuracies and balanced accuracies obtained with the iterative conditional imputation with
921 BR regression.

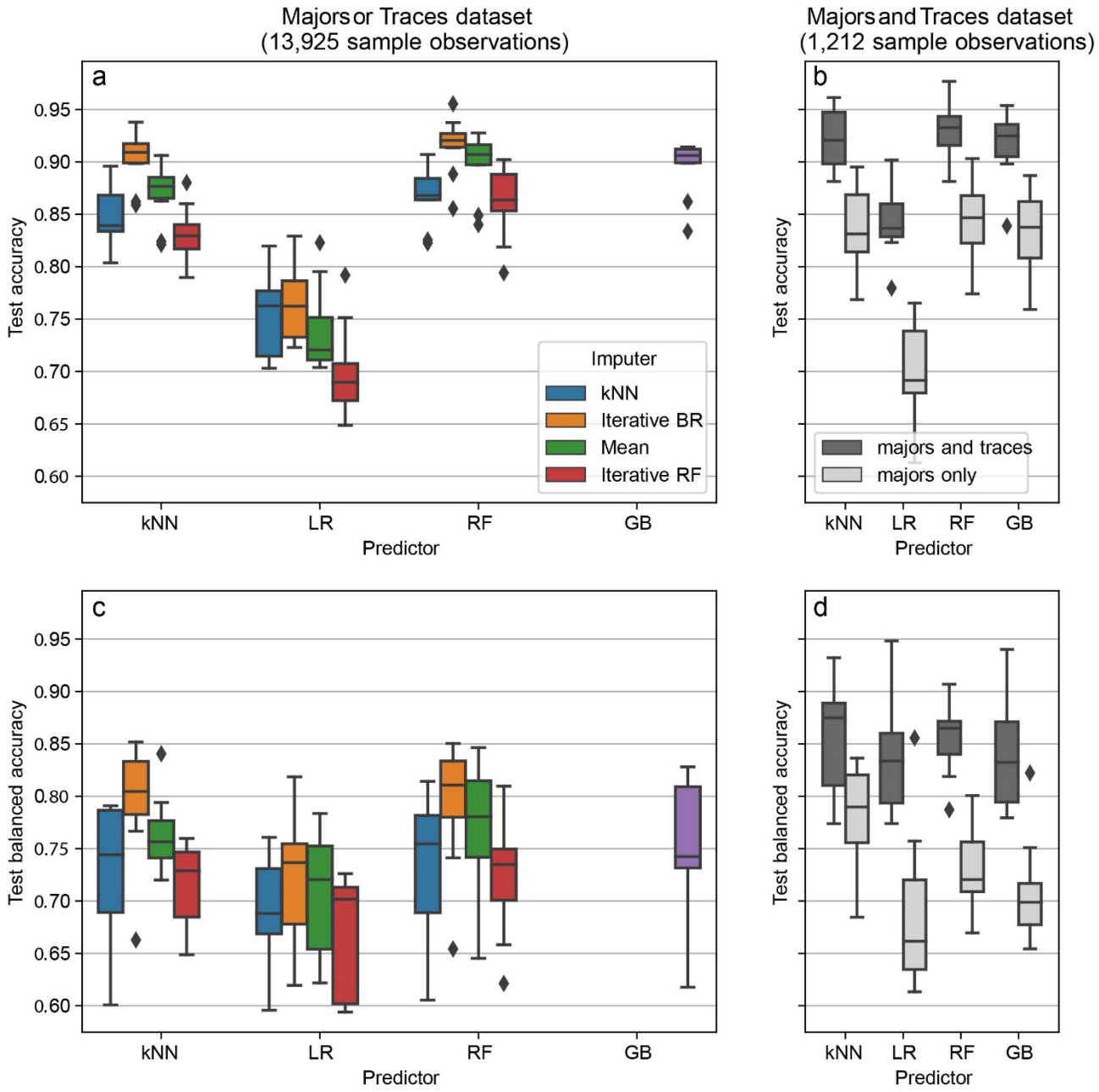


Figure 9. Boxplots showing the performances of the thirteen imputation-estimator models tested. Boxplots were obtained considering the performances on each of the 10 folds of the cross validation. a and c correspond to the accuracy and balanced accuracy on the "major *or* traces" set, i.e., the dataset containing sample observations on which either major elements, trace elements or both (13,925 sample observations) have been measured. In the legend the different imputation mechanisms are indicated: kNN: k-Nearest Neighbors; Iterative BR: iterative conditional imputation with Bayesian Ridge regressor; Mean: Simple Imputer with the mean value of each element; Iterative RF: iterative conditional imputation with Random Forest regressor. On the x axis, the different predictors tested are indicated: kNN: k-Nearest Neighbors; LR: Logistic Regression; RF: Random Forest; GB: Gradient Boosting. b and d correspond to the performances obtained on the "major *and* traces set", i.e., the dataset considering only sample observations for which *both* major elements *and* at least one trace element have been measured. In b and d, the comparison of the performances considering only the major elements of the major and traces set, versus the performances considering major and trace elements of the major and traces set, are plotted.

936 **Random Forest performs best, closely followed by KNN and Gradient Boosting –** Overall, RF had the
937 best performances (acc. 92% / balanced acc. 79%), followed by kNN (90% / 80%), GB (90% / 75%), and
938 finally LR, whose performance was significantly lower (76% / 72%). These results are not surprising as
939 RF and GB have been shown to work well with a wide range of datasets (Fernández-Delgado et al., 2014).
940 In particular, RF is among the best performing algorithms in previous work assessing the use of machine
941 learning for identifying the volcanic source of volcanic products, (Bolton et al., 2020; Pignatelli & Piochi,
942 2021; Uslular et al., 2022). In contrast, GB was tested exclusively in the South Aegean Active Volcanic Arc,
943 with similar performances to that of RF (Uslular et al., 2022). kNN displays more variable results in the
944 literature. Good performances were observed in a study with Alaska tephras (Bolton et al., 2020) and
945 kNN had the best performance in the Neapolitan region (Pignatelli & Piochi, 2021), yet it obtained
946 relatively bad performances in the South Aegean Active Volcanic Arc (Uslular et al., 2022). Finally, LR
947 displayed the worst performances considering both balanced and unbalanced accuracies, being lower
948 than all other models outside of 1SD. This reflects the fact that a linear model seems inappropriate on
949 such a dataset.

950 **Performances on rarer classes are satisfying –** Balanced accuracies are systematically lower than
951 accuracies, reflecting the fact that the learned classifiers are better on classes that have more sample
952 observations. This behavior is expected as the more samples we have of a class, the more accurately we
953 can characterize it. As can be observed in the confusion matrix for the model with the highest accuracy
954 (Figure 10), the five volcanic centers with more than 1000 sample observations are well classified by the
955 model. Only Llaima is sometimes classified as Villarrica or Mocho-Choshuenco, however it is relatively
956 well classified as well. Rarer classes are relatively well classified as well. In fact, from the five less
957 represented classes, Apagado, Hornopirén, and Huequi are very well classified. The few poorly classified
958 classes over the entire dataset (Sollipulli, Villarrica, Carrán-Los Venados, Osorno, Yate, and Aguilera), are
959 represented by a number of sample observations ranging from ~40 to 120, except for Villarrica with 858
960 sample observations. The latter indicates that classes that are less represented are more likely to be badly
961 classified. Thus, including more information in the future from the poorly classified samples could
962 increase the performances of the models.

Assessing the usefulness of traces in addition to major elements – In tephrochronology trace elements are generally analyzed to increase the chances of fingerprinting tephras. However it is unclear to which point measuring trace elements in addition to major elements can help better identify the volcanic source of a deposit. To answer this question, we restricted our attention to sample observations for which *both* major *and* trace elements are available, as both are needed to evaluate the effect of adding traces in a model. As most samples have only a subset of trace elements measured, we included all samples that have at least one trace element measured. It resulted in a dataset comprising 1,212 sample observations, belonging to 908 samples and representing 25 volcanoes. On this dataset, we trained the same models as above to predict the volcanic source, however, in this case, we first considered only the major elements (in the github repository this experiment corresponds to major_and_traces_restricted), and then, both the major and trace elements (major_and_traces in the github repository). All training details are similar to above. When only major elements are considered, the dataset has few missing values (0.1%) so we just imputed them by the mean. However, there are many missing values in the trace elements (30%), therefore we compared the various imputation methods on the majors and traces dataset. The results (Figure S6) show that mean imputation is one of the best methods for this dataset. Thus, we chose to compare the results with majors only and with majors and traces when both datasets are imputed by the mean (Figure 9b and 9d). Note that these results are affected by a higher variance than previously as we are using a dataset with roughly 1,000 sample observations, compared to more than 10,000 previously.

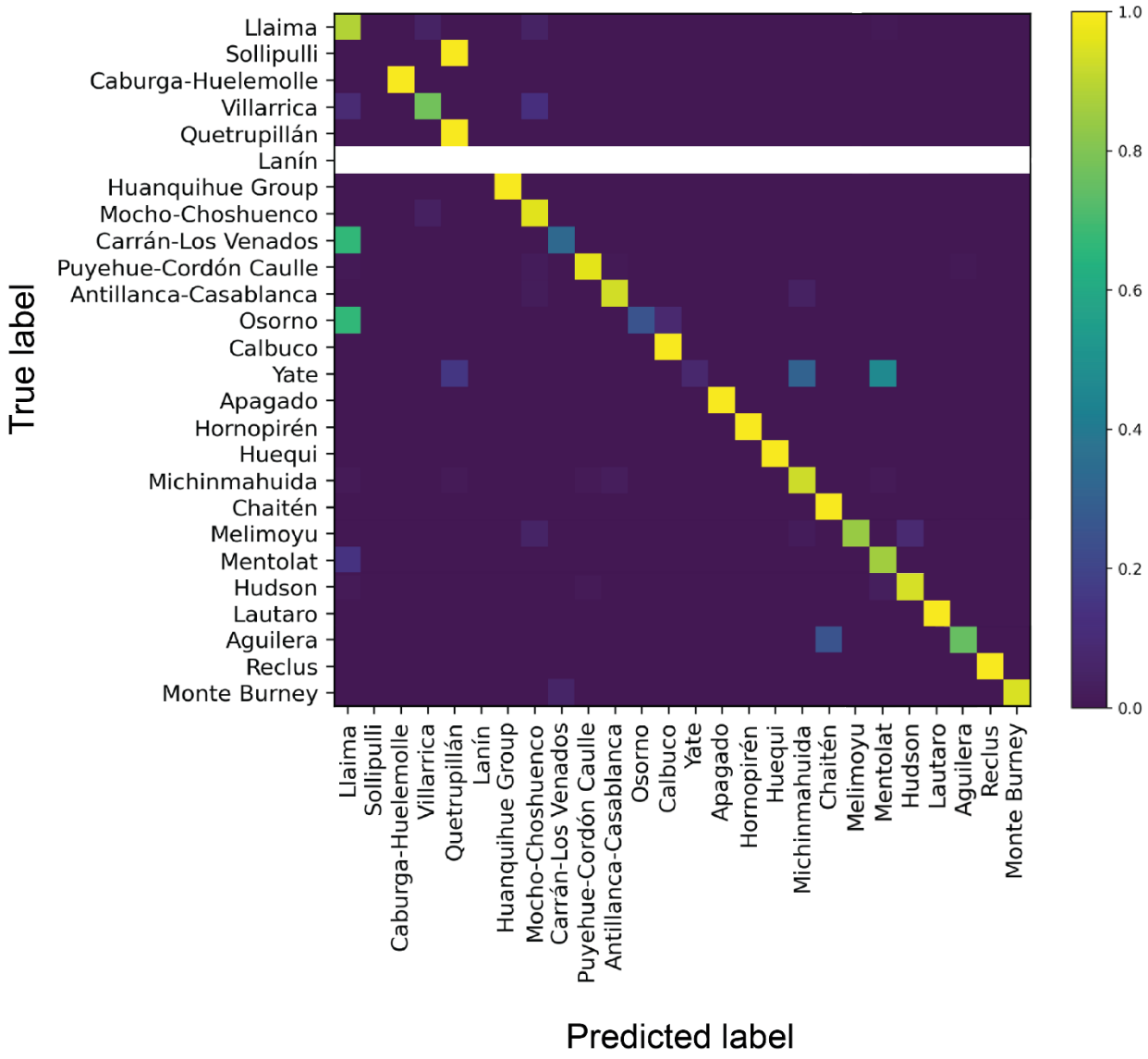


Figure 10. Confusion matrix showing the proportion of predicted sample observations by each volcanic center. Rows with no data correspond to classes in which no sample observations were left in the train set. The colored bar to the right of the figure indicates the proportion of classified sample observations, yellow being 1, thus all sample observations were classified as that volcanic center; and purple being 0, thus no sample observations were classified as that volcanic center.

Figure 9 (b, d) shows that using traces in addition to majors improves performances substantially across all models. The accuracy of kNN, RF and GB rises by 10 percent from 80-85% to 90-95%, while that of LR increases by 15 percent. The effect of adding traces as predictors is even stronger when looking at the balanced accuracy, except for kNN. While the balanced accuracy greatly varies across models when using majors elements only, it is roughly on par across models when adding traces. This may indicate that some rarer classes may be better classified when traces are used in addition to majors. Surprisingly, in terms

of balanced accuracy, even LR performs almost as well as RF or kNN when traces are considered, reaching 83.4%. This is an interesting observation as LR is a simple and easy to inspect model which tephrochronology experts might prefer over models that are more difficult to interpret. Overall, these results highlight the potential of traces for better (automated) fingerprinting of tephras.

The results discussed here constitute a first exploration showing encouraging performances. With more than 90% accuracy and around 80% balanced accuracy, the tests here performed highlight that machine learning with the BOOM! dataset has a great potential to improve correlations in the SVZ and AVZ. Nevertheless, models would benefit from an improvement in their balanced accuracy to help disentangle all tephrochronological records in the region, including the under-represented ones. As mentioned before, imputation algorithms have a noteworthy effect on performances when considering the whole dataset. Future work should thus focus on understanding the best imputation strategies, given the missingness structure of the dataset. In addition, it is important to include in the future a method to make pooled predictions, i.e., to predict the source not only for each sample observation independently, but for the whole sample at once. This question has been explored by Bolton et al. (2020) and performances of all models improved when considering the predictions at the sample level rather than at the sample observation level. Finally, using the BOOM! dataset, a more in-depth study can be developed to further the use of machine learning not only to identify the volcanic source of tephras, but to identify the exact eruption originating the tephra. In this sense, additional information included in the dataset could be included in the models, such as the ages, the type of volcanic product, and the location where they were identified. The code for the experiment here performed is available on github and can be used to improve our baselines.

6 Conclusions and perspectives

The BOOM! dataset here presented integrates an unprecedented amount of tephrochronological (and volcanological) information in the SVZ and AVZ in a machine-readable way. It contains information on 32 different volcanic centers and 132 eruptive events that occurred during the last 20,000 years, extracted from 79 different scientific publications, in ten different disciplines, produced through ~30 years of research. The machine-readable dataset allowed for the development of the BOOM! explorer, which provides novel visualizations of the tephrochronological information in the area. In particular, the BOOM! explorer allows users to visualize data in its multidimensionality (geographical distribution, chronology, and geochemical composition), filter data according to different criteria of comparability, and easily compare the available information for a given volcanic center or eruption, as well as directly linking the information to the source publication. The development of the explorer in parallel to the dataset greatly aided the development of its structure, by providing clear ways to organize the bulk of data and identify information that was missing in order to meet community needs.

In order to effectively integrate this heterogeneous data, it was key to incorporate information for users to evaluate its quality and comparability, which is a repeated request from the tephrochronological community. However, much of this information is not always communicated explicitly in publications, or in a way that is easily integrated with data from other publications, which impedes the reuse of data and can lead to misinterpretation of it. For example, in many publications it is not communicated in which state of oxidation of Fe is analyzed, rather it is deduced from the analytical technique. The latter can lead to confusion, especially when researchers from different disciplines reuse the information. In other cases, the position, type of register, analytical totals, or secondary standards analyzed along with the samples are not indicated in publications and could not be obtained when contacting the authors. Other types of information are not always communicated in a machine-readable way, making it very time consuming to integrate it with other data. In particular, including the stratigraphy of sections was a common request by the community, however in many publications this is communicated only in figures showing stratigraphic columns without the exact stratigraphic position and thickness of tephras. Future work in

tephrochronology would ideally publish these data in a machine-readable way to increase its re usability and the efficiency of tephrochronological work in the area.

The dataset and explorer are not meant to be comprehensive, and the work here presented is expected to be regarded as a basis from which to build upon. Regarding the dataset, more information could be incorporated to better describe tephra deposits. For example, more information to describe the physical characteristics of the tephra deposits, such as macroscopic description of it and photos; a more thorough description of the stratigraphy of sedimentary archives and their isotopic stratigraphy, when existing; as well as schematic representations of stratigraphic columns; supplementary dating methods, e.g., from dendrochronology, and more information to describe ^{14}C ages, for example reservoir ages and the choice of macrofossil dated, when suited, as well as the calibrated ages by the last published calibration curve; regarding the geochemical composition, analyses of mineral compositions or additional isotopic ratios such as $\delta^{18}O$. Regarding the explorer, additional or alternative visualizations could be developed to better visualize the uncertainty of the geochemical composition analyses and ^{14}C ages, the uncertainties related to age estimates of eruptive events. Additionally, an important request of the community to pursue in the future is finding ways to better communicate the uncertainty associated with the interpretation of the volcanic source of tephras, and how this is susceptible to change in the future, as new tephrochronological information becomes available. Finally, an important request which could not be met in this version of the dataset and the explorer, is that people can upload their own data and compare it to the dataset, which would be of great use for the community.

Even though compiling the BOOM! dataset and providing the explorer is a great contribution to the tephrochronological community in the area, an important challenge for the future is finding ways for the dataset to incorporate more information. In its current form as described here, the version of the dataset is static. Taking the latter into consideration, a collaboration with the National Volcanic Network (RNVV) of SERNAGEOMIN is already in place to look for ways of transferring the dataset and explorer to secure its maintenance and evolution in the future.

7 Data Availability

The BOOM! dataset is hosted on the ESPRI server of the IPSL, France, which guarantees its hosting and access as long as the server exists. Both files of the dataset can be downloaded directly from the IPSL catalog (https://doi.org/10.14768/47b4525f-ff39-4940-a963-4d2673f2362e) as a Web Map service (WMS), Web Feature service (WFS) or as a CSV file. Additionally, either the entire dataset or subsets of it can be downloaded as a CSV file through the BOOM! explorer (https://boom-vis.lisn.upsaclay.fr).

8 Acknowledgements

This study benefited from the Chilean scholarship for PhD *Becas Chile, convocatoria* 2017. For data hosting and processing, it benefited from the IPSL mesocenter ESPRI facility, which is supported by CNRS, UPMC, Labex L-IPSL, CNES and *Ecole Polytechnique*. We especially thank Karim Ramage and Vincent Douet from ESPRI for their kind support. Additionally, we would like to especially thank the team from the RNVV from SERNAGEOMIN and Sebastian Watt, who kindly participated in the interviews and gave continuous feedback for the development of the platform. We also thank every researcher who kindly shared details on the information they discuss in their papers in order to enrich the dataset: Brent Alloway, Charles Stern, Daniel Sellés, Derek Weller, Eduardo Morgado, Francisco Bucchi, Harriet Rawson, Hugo Moreno, José Naranjo, Luis Lara, Maarten Van Daele, Rebecca Smith, Sebastien Bertrand, Simon Haberle, and Silke Lohmar. Finally, we thank Tamara Basualto for helping with the graphic design in the paper, and Sebastián Miranda, Camila Miranda and Taylor Quilty for helping us come up with BOOM!

9 References

Abarzúa A.M. & Moreno P.I. 2008. Changing fire regimes in the temperate rainforest region of southern Chile over the last 16,000 yr. *Quaternary Research,* 69 (1), 62–71. https://doi.org/10.1016/j.yqres.2007.09.004

Abarzúa A.M., Villagrán C. & Moreno P.I. 2004. Deglacial and postglacial climate history in east-central Isla Grande de Chiloé, southern Chile (43°S). *Quaternary Research,* 62 (1), 49–59. https://doi.org/10.1016/j.yqres.2004.04.005

1091 Abbott, P., Bonadonna, C., Bursik, M., Cashman, K., Davies, S., Jensen, B., Kuehn, S., Kurbatov, A., Lane, C.,
1092 Plunkett, G., Smith, V., Thomlinson, E., Thordarsson, T., Walker, J.D., & Wallace, K. 2022.
1093 Community Established Best Practice Recommendations for Tephra Studies-from Collection
1094 through Analysis (4.0.2) [Data set]. *Zenodo*. https://doi.org/10.5281/zenodo.6568306

1095 Alloway, B.V., Pearce, N.J.G., Moreno, P.I., Villarosa, G., Jara, I., De Pol-Holz, R. & Outes, V. 2017a. An 18,000
1096 year-long eruptive record from Volcán Chaitén, northwestern Patagonia: Paleoenvironmental
1097 and hazard-assessment implications. *Quaternary Science Reviews*. 168, 151–181.
1098 https://doi.org/10.1016/j.quascirev.2017.05.011

1099 Alloway, B.V., Moreno, P.I., Pearce, N.J.G., De Pol-Holz, R., Henríquez, W.I., Pesce, O.H., Sagredo, E.,
1100 Villarosa, G. & Outes, V. 2017b. Stratigraphy, age and correlation of Lepué Tephra: a widespread
1101 c. 11 000 cal a BP marker horizon sourced from the Chaitén Sector of southern Chile. *Journal of*
1102 *Quaternary Science*. 32 (6), 795–829. https://doi.org/10.1002/jqs.2976

1103 Alloway, B.V., Pearce, N.J.G., Villarosa, G., Outes, V. & Moreno, P.I. 2015. Multiple melt bodies fed the AD
1104 2011 eruption of Puyehue-Cordón Caulle, Chile. *Scientific Reports*, 5: 17589.
1105 https://doi.org/10.1038/srep17589

1106 Amigo, Á., Lara, L.E. & Smith, V.C. 2013. Holocene record of large explosive eruptions from Chaitén and
1107 Michinmahuida Volcanoes, Chile. *Andean Geology*, 40 (2), 227–248.
1108 https://doi.org/10.5027/andgeov40n2-a03

1109 Bertin, L., Moreno, H. & Becerril, L., 2018. *Peligros del Campo Volcánico Carrán-Los Venados, Región de los*
1110 *Ríos*, *Carta Geológica de Chile, Serie Geológica Ambiental,* N° 33*. Servicio Nacional de Geología y*
1111 *Minería – Chile* (SERNAGEOMIN)*, Subdirección Nacional de Geología*.

1112 Bertrand, S., Castiaux, J. & Juvigné, E. 2008a. Tephrostratigraphy of the late glacial and Holocene
1113 sediments of Puyehue Lake (Southern Volcanic Zone, Chile, 40°S). *Quaternary Research*. 70 (3),
1114 343–357. https://doi.org/10.1016/j.yqres.2008.06.001

1115 Bertrand, S., Charlet, F., Charlier, B., Renson, V. & Fagel, N. 2008b. Climate variability of southern Chile
1116 since the Last Glacial Maximum: a continuous sedimentological record from Lago Puyehue (40°S).
1117 *Journal of Paleolimnology*, 39, 179–195. https://doi.org/10.1007/s10933-007-9117-y

Bertrand, S., R. Daga, R. Bedert, & K. Fontijn. 2014. Deposition of the 2011–2012 Cordón Caulle tephra (Chile, 40°S) in lake sediments: Implications for tephrochronology and volc anology. Journal of Geophysical Research: Earth Surfce, 119, 2555–2573, https://doi.org/10.1002/2014JF003321

Bitschene, P., Arias, N., Arizmendia, A., Giacosa, R., Grizinik, M., Fernández, M., Márquez, M. & Nillni, A. 1993. Vulcanology and enviromental impact of the August 1991 eruption of the Hudson Vulcano (Patagonian Andes; Chile). *Zentralblatt Geologie u. Paläontologie*. Stuttgart. H 1-2: 493-500.

Bolton, M.S.M., Jensen, B.J.L., Wallace, K., Praet, N., Fortin, D., Kaufman, D., De Batist, M. 2020. Machine learning classifiers for attributing tephra to source volcanoes: an evaluation of methods for Alaska tephras. *Journal of Quaternary Science,* 35 (1–2), 81–92. https://doi.org/10.1002/jqs.3170

Bouvet de Maisonneuve, C., Dungan, M.A., Bachmann, O. & Burgisser, A. 2012. Insights into shallow magma storage and crystallization at Volcán Llaima (Andean Southern Volcanic Zone, Chile). *Journal of Volcanology and Geothermal Research*. 211–212, 76–91. https://doi.org/10.1016/j.jvolgeores.2011.09.010

Brahm, R., Parada, M. A., Morgado, E., Contreras, C. & McGee, L. E. 2018. Origin of Holocene trachyte lavas of the Quetrupillán volcanic complex, Chile: Examples of residual melts in a rejuvenated crystalline mush reservoir. *Journal of Volcanology and Geothermal Research*, 357, 163–176. https://doi.org/10.1016/j.jvolgeores.2018.04.020

Bucchi, F., Lara, L. E. & Gutiérrez, F. 2015. The Carrán–Los Venados volcanic field and its relationship with coeval and nearby polygenetic volcanism in an intra-arc setting. *Journal of Volcanology and Geothermal Research*, 308, 70-81. https://doi.org/10.1016/j.jvolgeores.2015.10.013

Carel, M., Siani, G. & Delpech, G. 2011. Tephrostratigraphy of a deep-sea sediment sequence off the south Chilean margin: New insight into the Hudson volcanic activity since the last glacial period. *Journal of Volcanology and Geothermal Research,* 208 (3–4), 99–111. https://doi.org/10.1016/j.jvolgeores.2011.09.011

Casati, E., D'Amico, M., Šefrna, L., Trombino, L., Tunesi, A. & Previtali, F. 2019. Geo-pedological contribution to the reconstruction of Holocene activity of Chaitén volcano (Patagonia, Chile). *Journal of South American Earth Sciences*, 94: 102222. https://doi.org/10.1016/j.jsames.2019.102222

Clapperton C.M., Sugden D.E., Kaufman D.S. & McCulloch R.D. 1995. The last glaciation in central Magellan Strait, southernmost Chile. *Quaternary Research*, 44 (2), 133–148. https://doi.org/10.1006/qres.1995.1058

Costantini, L., Pioli, L., Bonadonna, C., Clavero, J. & Longchamp, C. 2011. A Late Holocene explosive mafic eruption of Villarrica volcano, Southern Andes: The Chaimilla deposit, *Journal of Volcanology and Geothermal Research*, 200 (3–4), 143–158. https://doi.org/10.1016/j.jvolgeores.2010.12.010

Davies, B.J., Darvill, C.M., Lovell, H., Bendle, J.M., Dowdeswell, J.A., Fabel, D., García, J.L., Geiger, A., Glasser, N.F., Gheorghiu, D.M., Harrison, S., Hein, A.S., Kaplan, M.R., Martin, J.R.V., Mendelova, M., Palmer, A., Pelto, M., Rodés, Á., Sagredo, E.A., Smedley, R.K., Smellie, J.L., Thorndycraft, V.R., 2020. The evolution of the Patagonian Ice Sheet from 35 ka to the present day (PATICE). *Earth-Science Rev*iews 204: 103152. https://doi.org/10.1016/j.earscirev.2020.103152

Del Carlo, P., Di Roberto, A., D'Orazio, M., Petrelli, M., Angioletti, A., Zanchetta, G., Maggi, V., Daga, R., Nazzari, M. & Rocchi, S., 2018. Late Glacial-Holocene tephra from southern Patagonia and Tierra del Fuego (Argentina, Chile): A complete textural and geochemical fingerprinting for distal correlations in the Southern Hemisphere. *Quaternary Science Reviews*, 195 (1), 153–170. https://doi.org/10.1016/j.quascirev.2018.07.028

D'Orazio, M., Innocenti, F., Manetti, P., Tamponi, M., Tonarini, S., González-Ferrán, O., Lahsen, A. & Omarini, R. 2003. The Quaternary calc-alkaline volcanism of the Patagonian Andes close to the Chile triple junction: geochemistry and petrogenesis of volcanic rocks from the Cay and Maca volcanoes (~45°S, Chile). *Journal of South American Earth Sciences*, 16 (4), 219–242. https://doi.org/10.1016/S0895-9811(03)00063-4

Fernández-Delgado M., Cernadas E., Barro S., & Amorim D. 2014. Do we need hundreds of classifiers to solve real world classification problems? *Journal of machine learning research,* 15 (1): 3133–3181.

Fontijn, K., Lachowycz, S.F., Rawson, H., Pyle, D.M., Mather, T.A., Naranjo, J.A. & Moreno-Roa, H. 2014. Late Quaternary tephrostratigraphy of southern Chile and Argentina. *Quaternary Science Reviews*, 89, 70–84. https://doi.org/10.1016/j.quascirev.2014.02.007

Fontijn, K., Rawson, H., Van Daele, M., Moernaut, J., Abarzúa, A.M., Heirman, K., Bertrand, S., Pyle, D.M., Mather, T.A., De Batist, M., Naranjo J.A. & Moreno, H., 2016. Synchronisation of sedimentary records using tephra: A postglacial tephrochronological model for the Chilean Lake District. *Quaternary Science Reviews*. 137, 234–254. https://doi.org/10.1016/j.quascirev.2016.02.015

1176 Futa K. & Stern C. R. 1988. Sr and Nd isotopic and trace element compositions of Quaternary volcanic
1177 centers of the southern Andes. *Earth and Planetary Science Letters*, 88 (3–4), 253–262.
1178 https://doi.org/10.1016/0012-821X(88)90082-9

1179 Geoffroy, C. A., Alloway, B. V., Amigo, Á., Parada, M. A., Gutierrez, F., Castruccio, A., Pearce, N. J. G., Morgado,
1180 E. & Moreno, P.I. 2018. A widespread compositionally bimodal tephra sourced from Volcán
1181 Melimoyu (44°S, Northern Patagonian Andes): Insights into magmatic reservoir processes and
1182 opportunities for regional correlation. *Quaternary Science Reviews,* 200, 141–159.
1183 https://doi.org/10.1016/j.quascirev.2018.09.034

1184 Gerlach, D.C., Frey, F.A., Moreno-Roa, H. & López-Escobar, L. 1988. Recent Volcanism in the Puyehue–
1185 Cordon Caulle Region, Southern Andes, Chile (40.5°S): Petrogenesis of Evolved Lavas. *Journal of*
1186 *Petrology,* 29 (2), 333–382. https://doi.org/10.1093/petrology/29.2.333

1187 Geyh, M.A., Schotterer, U., Grosjean, M. 1998. Temporal changes of the ^{14}C reservoir effect in lakes.
1188 *Radiocarbon,* 40 (2), 921–931. https://doi.org/10.1017/S0033822200018890

1189 Haberle, S.G. & Lumley, S.H. 1998. Age and origin of tephras recorded in postglacial lake sediments to the
1190 west of the southern Andes, 44°S to 47°S. *Journal of Volcanology and Geothermal Research*, 84 (3–
1191 4), 239–256. https://doi.org/10.1016/S0377-0273(98)00037-7

1192 Harambour, S.V. 1988. Sobre et hallazgo del mítico volcán Reclus, ex Mano del diablo, Hielo Patagónico
1193 Sur, Magallanes, Chile. *Revista Geológica de Chile*, 15 (2), 173–179.

1194 Heusser, C.J., Heusser, L.E. & Hausser A. 1989. A 12,000 yr B.P. tephra layer at Bahia Inutil (Tierra del
1195 Fuego, Chile). *Anales del Instituto de la Patagonia*, 19 (1), 39–49.

1196 Heusser, C. J., Heusser, L. E., Lowell, T.V., Moreira, A. & Moreira, S. 2000. Deglacial palaeoclimate at Puerto
1197 del Hambre, subantarctic Patagonia, Chile. *Journal of Quaternary Science*, 15 (2), 101–114.
1198 https://doi.org/10.1002/(SICI)1099-1417(200002)15:2<101::AID-JQS500>3.0.CO;2-Y

1199 Hickey-Vargas, R.L., Moreno Roa, H., López-Escobar, L. & Frey, F.A. 1989. Geochemical variations in
1200 Andean basaltic and silicic lavas from the Villarrica-Lanin volcanic chain (39.5° S): an evaluation
1201 of source heterogeneity, fractional crystallization and crustal assimilation. *Contributions to*
1202 *Mineralogy and Petrology*. 103, 361–386. https://doi.org/10.1007/BF00402922

1203 Jacques, G., Hoernle, K., Gill, J., Wehrmann, H., Bindeman, I. & Lara, L.E. 2014. Geochemical variations in
1204 the Central Southern Volcanic Zone, Chile (38–43°S): The role of fluids in generating arc magmas.
1205 *Chemical Geology,* 371, 27–45. https://doi.org/10.1016/j.chemgeo.2014.01.015

Jochum, K.P., Stoll, B., Herwig, K., Willbold, M., Hofmann, A.W., Amini, M., Aarburg, S., Abouchami, W., Hellebrand, E., Mocek, B., Raczek, I., Stracke, A., Alard, O., Bouman, C., Becker, S., Dücking, M., Brätz, H., Klemd, R., de Bruin, D., et al. 2006. MPI-DING reference glasses for in situ microanalysis: New reference values for element concentrations and isotope ratios. *Geochemistry, Geophysics, Geosystems*, 7 (2): Q02008. https://doi.org/10.1029/2005GC001060

Jochum, K.P., Weis, U., Stoll, B., Kuzmin, D., Yang, Q., Raczek, I., Jacob, D.E., Stracke, A., Birbaum, K., Frick, D.A., Günther,D. & Enzweiler, J. 2011. Determination of Reference Values for NIST SRM 610—617 Glasses Following ISO Guidelines. *Geostandards and Geoanalytical Research*, 34 (4), 397–429. https://doi.org/10.1111/j.1751-908X.2011.00120.x

Jochum, K.P., Weis, U., Schwager, B., Stoll, B., Wilson, S.A., Haug, G.H., Andreae, M.O. & Enzweiler, J. 2016. Reference Values Following ISO Guidelines for Frequently Requested Rock Reference Materials. *Geostandards and Geoanalytical Research*, 40 (3), 333–350. https://doi.org/10.1111/j.1751-908X.2015.00392.x

Kilian, R., Hohner, M., Biester, H., Wallrabe-Adams, H.J., Stern, C.R. 2003. Holocene peat and lake sediment tephra record from the southernmost Chilean Andes (53-55°S). *Revista Geológica de Chile* 30 (1), 23–37. http://dx.doi.org/10.4067/S0716-02082003000100002

Kratzmann, D.J., Carey, S., Scasso, R. & Naranjo, J.A. 2008. Compositional variations and magma mixing in the 1991 eruptions of Hudson volcano, Chile. *Bulletin of Volcanology*. 71, 419–439. https://doi.org/10.1007/s00445-008-0234-x

Lara, L. E., Naranjo, J. A. & Moreno, H. 2004. Lanín volcano (39.5°S), Southern Andes: geology and morphostructural evolution. *Revista Geológica de Chile*, 31 (2), 241–257. https://doi.org/10.4067/S0716-02082004000200004

Lara, L.E., Moreno, H., Naranjo, J.A., Matthews, S. & Pérez de Arce, C. 2006. Magmatic evolution of the Puyehue-Cordón Caulle Volcanic Complex (40° S), Southern Andean Volcanic Zone: From shield to unusual rhyolitic fissure volcanism. *Journal of Volcanology and Geothermal Research*, 157 (4), 343–366. https://doi.org/10.1016/j.jvolgeores.2006.04.010

1232 Lara, L. E. & Moreno, H. 2006. *Geología del complejo volcánico Puyehue-Cordón Caulle, Región de Los Lagos*.
1233 Servicio Nacional de Geología y Minería, Carta Geológica de Chile, Serie Geológica Básica, 99,
1234 Escalara 1:50.000, Santiago.

1235 Lohmar, 2008. Pétrologie des grands dépôts d'ignimbrites des volcans Villarrica (Licán et Pucón) et
1236 Llaima (Ignimbrite Curacautín), dans les Andes du Sud (Chili). PhD Thesis, Université Blause
1237 Pascal - Clermont Ferrand II, France.

1238 Lohmar, S., Parada, M., Gutiérrez, F., Robin, C. & Gerbe, M.C. 2012. Mineralogical and numerical
1239 approaches to establish the pre-eruptive conditions of the mafic Licán Ignimbrite, Villarrica
1240 Volcano (Chilean Southern Andes). *Journal of Volcanology and Geothermal Research*, 235–236, 55–
1241 69, https://doi.org/10.1016/j.jvolgeores.2012.05.006

1242 López-Escobar, L., Parada, M.A., Moreno, H., Frey, F.A. & Hickey-Vargas, R.L. 1992. A contribution to the
1243 petrogenesis of Osorno and Calbuco volcanoes, Southern Andes (41°00'-41°30'S): a comparative
1244 study. *Revista Geológica de Chile,* 19 (2), 211–226.

1245 López-Escobar, L., Kilian, R., Kempton, P.D. & Tagiri, M. 1993. Petrography and geochemistry of
1246 Quaternary rocks from the Southern Volcanic Zone of the Andes between 41°30′ and 46°00′S,
1247 Chile. *Revista Geológica de Chile*, 20, 33–55.

1248 López-Escobar, L., Parada, M.A., Hickey-Vargas, R.L., Frey, F.A., Kempton, P.D. & Moreno, H. 1995. Calbuco
1249 Volcano and minor eruptive centers distributed along the Liquiñe-Ofqui Fault Zone, Chile (41°–
1250 42° S): contrasting origin of andesitic and basaltic magma in the Southern Volcanic Zone of the
1251 Andes. *Contributions to Mineralogy and Petrology*, 119, 345–361.
1252 https://doi.org/10.1007/BF00286934

1253 Lowe, D.J. 2011. Tephrochronology and its application: A review. *Quaternary Geochronology*, 6, 107–153.
1254 https://doi.org/10.1016/j.quageo.2010.08.003

1255 Lowe, D.J., Pearce, N.J.G., Jorgensen, M.A., Kuehn, S.C., Tryon, C.A. & Hayward, C.L. 2017. Correlating
1256 tephras and cryptotephras using glass compositional analyses and numerical and statistical
1257 methods: Review and evaluation. *Quaternary Science Reviews,* 175, 1–44.
1258 https://doi.org/10.1016/j.quascirev.2017.08.003

1259 Martínez Fontaine, C., Siani, G., Delpech, G., Michel, E., Villarosa, G., Manssouri, F. & Nouet, J. 2021. Post–
1260 glacial tephrochronology record off the Chilean continental margin (~41° S). *Quaternary Science*
1261 *Reviews,* 261: 106928. https://doi.org/10.1016/j.quascirev.2021.106928

Mayr, C., Smith, R.E., García, M.L., Massaferro, J., Lücke, A., Dubois, N., Maidana, N., I., Meier, W. J-H., Wissel, H. & Zolitschka, B. 2019. Historical eruptions of Lautaro Volcano and their impacts on lacustrine ecosystems in southern Argentina. *Journal of Paleolimnology*, 62, 205–221. https://doi.org/10.1007/s10933-019-00088-y

McNamara, K., Rus, A.C., Cashman, K.V., Castruccio, A. & Abarzúa, A.M. 2019. Comparison of lake and land tephra records from the 2015 eruption of Calbuco volcano, Chile. *Bulletin of Volcanology*, 821: 10 (2019). https://doi.org/10.1007/s00445-019-1270-4

Mella, M. 2008. Petrogêneses do complexo vulcânico Yate (42, 30ºS), Andes do Sul, Chile. Tese de Doutorado, Instituto de Geociências, Universidade de São Paulo, São Paulo. https://doi.org/10.11606/T.44.2009.tde-04032009-091537

Moreno, P.I. & León A.L. 2003. Abrupt vegetation changes during the last glacial to Holocene transition in mid-latitude South America. *Journal of Quaternary Science,* 18 (18), 787–800. https://doi.org/10.1002/jqs.801

Moreno, PI. 2004. Millennial-scale climate variability in northwest Patagonia over the last 15 000 yr. *Journal of Quaternary Science,* 19 (1), 35–47. https://doi.org/10.1002/jqs.813

Moreno, P.I., Alloway, B.V., Villarosa, G., Outes, V., Henríquez, W.I., De Pol-Holz, R. & Pearce, N.J.G., 2015. A past-millennium maximum in postglacial activity from Volcán Chaitén, southern Chile. *Geology,* 43 (1), 47–50. https://doi.org/10.1130/G36248.1

Moreno, P. I., Vilanova, I., Villa-Martínez, R., Dunbar, R.B., Mucciarone, D.A., Kaplan, M.R., Garreaud, R. D.M., Rojas, M., Moy, C. M., De Pol-Holz, C.M. & Lambert, F. 2018. Onset and Evolution of Southern Annular Mode-Like Changes at Centennial Timescale. *Scientific Reports*, 8: 3458. https://doi.org/10.1038/s41598-018-21836-6

Moreno, H., Lara, L. & Orozco, G. 2010. *Geología del Volcán Osorno, Región de Los Lagos*. Servicio Nacional de Geología y Minería, Carta Geológica de Chile, Serie Geológica Básica, N° 126. *Servicio Nacional de Geología y Minería – Chile* (SERNAGEOMIN)*, Subdirección Nacional de Geología*.

Morgado, E., Parada, M.A., Contreras, C., Castruccio, A., Gutiérrez, F. & McGee, L.E. 2015. Contrasting records from mantle to surface of Holocene lavas of two nearby arc volcanic complexes: Caburgua-Huelemolle Small Eruptive Centers and Villarrica Volcano, Southern Chile. *Journal of*

Volcanology and Geothermal Research. 306, 1–16. https://doi.org/10.1016/j.jvolgeores.2015.09.023

Morgado, E., Morgan, D.J., Harvey, J., Parada, M.Á., Castruccio, A., Brahm, R., Gutiérrez, F., Georgiev, B. & Hammond, S.J. 2019. Localised heating and intensive magmatic conditions prior to the 22–23 April 2015 Calbuco volcano eruption (Southern Chile). *Bulletin of Volcanology*. 81: 24. https://doi.org/10.1007/s00445-019-1280-2

Motoki, A., Orihashi, Y., Naranjo, J.A., Hirata, D., Skvarca, P. & Anma, R. 2006. Geologic reconnaissance of Lautaro Volcano, Chilean Patagonia. *Revista Geológica de Chile*, 33 (1), 177–187. http://doi.org/10.4067/S0716-02082006000100008

Naranjo, J.A. & Moreno, H. 1991. Actividad explosiva postglacial en el volcan Llaima, Andes del Sur (38° 45'S). *Revista Geológica de Chile*, 18, 69–80.

Naranjo, J.A., Moreno, H. & Banks, N.G. 1993. *La erupción del volcán Hudson en 1991 (46°S), Región XI, Aisén, Chile,* Boletín 44. *Servicio Nacional Geología y Minería -Chile* (SERNAGEOMIN), *Subdirección Nacional de Geología*.

Naranjo, J.A. & Moreno, H. 2005. *Geología del Volcán Llaima, Región de la Araucanía. Carta Geológica de Chile, Serie Geología Básica*, N° 88. *Servicio Nacional de Geología y Minería – Chile* (SERNAGEOMIN), *Subdirección Nacional de Geología*.

Naranjo, J.A., Singer, B.S., Jicha, B.R., Moreno, H. & Lara, L.E. 2017. Holocene tephra succession of Puyehue-Cordón Caulle and Antillanca/Casablanca volcanic complexes, southern Andes (40–41°S). *Journal of Volcanology and Geothermal Research,* 332, 109–128. https://doi.org/10.1016/j.jvolgeores.2016.11.017

Naranjo, J.A. & Stern, C.R. 2004. Holocene tephrochronology of the southernmost part (42°30'-45°S) of the Andean Southern Volcanic Zone. *Revista Geológica de Chile,* 31 (2), 225–240. https://doi.org/10.4067/S0716-02082004000200003

Naranjo, J.A., Stern, C.R. 1998. Holocene explosive activity of Hudson Volcano, southern Andes. *Bulletin of Volcanology* 59, 291–306. https://doi.org/10.1007/s004450050193

Scikit-learn: Machine Learning in Python, Pedregosa et al., JMLR 12, pp. 2825-2830, 2011.

Pearce, N.J.G, Bendall, C.A., and Westgate, J.A. 2008. Comment on "Some numerical considerations in the geochemical analysis of distal microtephra" by A.M. Pollard, S.P.E. Blockley and C.S. Lane. Aplied Geochemistry, 23 (5), 1353-1364. https://doi.org/10.1016/j.apgeochem.2008.01.002

Pesce, O.H. & Moreno, P.I. 2014. Vegetation, fire and climate change in central-east Isla Grande de Chiloé (43°S) since the Last Glacial Maximum, northwestern Patagonia. *Quaternary Science Reviews,* 90, 143–157. https://doi.org/10.1016/j.quascirev.2014.02.021

Pignatelli, A. & Piochi, M. 2021. Machine learning applied to rock geochemistry for predictive outcomes: The Neapolitan volcanic history case. *Journal of Volcanology and Geothermal Research*, 415: 107254. https://doi.org/10.1016/j.jvolgeores.2021.107254

Prieto, A., Stern, C.R. & Estévez, J.E. 2013. The peopling of the Fuego-Patagonian fjords by littoral hunter-gatherers after the mid-Holocene H1 eruption of Hudson Volcano. *Quaternary International*, 317, 3–13. https://doi.org/10.1016/j.quaint.2013.06.024

Rawson, H., Naranjo, J. A., Smith, V. C., Fontijn, K., Pyle, D. M., Mather, T. A. & Moreno, H. 2015. The frequency and magnitude of post-glacial explosive eruptions at Volcán Mocho-Choshuenco, southern Chile. *Journal of Volcanology and Geothermal Research*, 299, 103–129. https://doi.org/10.1016/j.jvolgeores.2015.04.003

Rawson, H., Keller, T., Fontijn, K., Pyle, D. M., Mather, T.A., Smith, V.C. & Naranjo, J. A. 2016. Compositional variability in mafic arc magmas over short spatial and temporal scales: Evidence for the signature of mantle reactive melt channels. *Earth and Planetary Science Letters*. 456, 66–77. https://doi.org/10.1016/j.epsl.2016.09.056

Reubi, O., Bourdon, B., Dungan, M.A., Koornneef, J.M., Sellés, D., Langmuir, C.H. & Aciego, S. 2011. Assimilation of the plutonic roots of the Andean arc controls variations in U-series disequilibria at Volcán Llaima, Chile. *Earth and Planetary Science Letters*, 303 (1–2), 37–47. https://doi.org/10.1016/j.epsl.2010.12.018

Romero, J.E., Morgavi, D., Arzilli, F., Daga, R., Caselli, A., Reckziegel, F., Viramonte, J., Díaz-Alvarado, J., Polacci, M., Burton, M. & Perugini, D. 2016. Eruption dynamics of the 22–23 April 2015 Calbuco Volcano (Southern Chile): Analyses of tephra fall deposits. *Journal of Volcanology and Geothermal Research*, 317, 15-29. https://doi.org/10.1016/j.jvolgeores.2016.02.027

Schindlbeck, J.C., Freundt, A. & Kutterolf, S. 2014. Major changes in the post-glacial evolution of magmatic compositions and pre-eruptive conditions of Llaima Volcano, Andean Southern Volcanic Zone, Chile. *Bulletin of Volcanology*. 76: 830. https://doi.org/10.1007/s00445-014-0830-x

Sellés, D. & Moreno, H. 2011. *Geología del Volcán Calbuco, Región de Los Lagos*. *Carta Geológica de Chile, Serie Geología Básica*, N° 130. *Servicio Nacional de Geología y Minería – Chile* (SERNAGEOMIN), *Subdirección Nacional de Geología*.

Siani, G., Michel, E., De Pol-Holz, R., DeVries, T., Lamy, F., Carel, M., Isguder, G., Dewilde, F. & Lourantou, A. 2013. Carbon isotope records reveal precise timing of enhanced Southern Ocean upwelling during the last deglaciation. *Nature Communications*, 4: 2758. https://doi.org/10.1038/ncomms3758

Simmons, I.C., McGarvie, D., Cortés, J.A., Calder, E.S. & Pavez, A. 2020. Holocene volcanism at the Quetrupillán Volcanic Complex (39°30' S, 71°43' W), southern Chile. *Volcanica*, 3 (1), 115–137. https://doi.org/10.30909/vol.03.01.115137

Singer, B.S., Jicha, B.R., Harper, M.A., Naranjo, J.A., Lara, L.E. & Moreno-Roa, H. 2008. Eruptive history, geochronology, and magmatic evolution of the Puyehue-Cordón Caulle volcanic complex, Chile. *GSA Bulletin,* 120 (5–6), 599–618. https://doi.org/10.1130/B26276.1

Smith, R.E., Smith, V.C., Fontijn, K., Gebhardt, A.C., Wastegård, S., Zolitschka, B., Ohlendorf, C., Stern, C. & Mayr, C. 2019. Refining the Late Quaternary tephrochronology for southern South America using the Laguna Potrok Aike sedimentary record. *Quaternary Science Reviews*. 218, 137–156. https://doi.org/10.1016/j.quascirev.2019.06.001

Stern, C. R., Moreno, P. I., Villa-Martínez, R., Sagredo, E. A., Prieto, A. & Labarca, R. 2011. Evolution of ice-dammed proglacial lakes in Última Esperanza, Chile: implications from the late-glacial R1 eruption of Reclús volcano, Andean Austral Volcanic Zone. *Andean Geology*, 38 (1), 82–97.

Stern, C.R., de Porras, M.E. & Maldonado, A. 2015. Tephrochronology of the upper Río Cisnes valley (44°S), southern Chile. *Andean Geology*, 42 (2), 173–189. https://doi.org/10.5027/andgeoV42n2-a02

Stern, C.R. 1992. Tephrochronology of Magallanes: new data and implications. *Anales del Instituto de la Patagonia* 21: 129-139.

Stern, C.R., 2008. Holocene tephrochronology record of large explosive eruptions in the southernmost Patagonian Andes. *Bulletin of Volcanology,* 70, 435–454. https://doi.org/10.1007/s00445-007-0148-z

Uslular, G., Kıyıkçı, F., Karaarslan, E. & Gençalioğlu Kuşcu, G. 2022. Application of machine-learning algorithms for tephrochronology: a case study of Plio-Quaternary volcanic fields in the South Aegean Active Volcanic Arc. *Earth Science Informatics* 15, 1167–1182. https://doi.org/10.1007/s12145-022-00797-5

Tagiri, M., Moreno, H., López-Escobar, L. & Notsu, K. 1993. Two magma types of the high-alumina basalt series of Osorno, Southern Andes (41°06'S)-plagioclase dilution effect. *Journal of Mineralogy, Petrology and Economic Geology*, 88 (7), 359–371. https://doi.org/10.2465/ganko.88.359

van Buuren, S. 2018. *Flexible Imputation of Missing Data* (Second edition). CRC/Chapman & Hall. https://doi.org/10.1201/9780429492259

Villarosa, G., Outes, V., Hajduk, A., Crivelli Montero, E., Sellés, D., Fernández, M. & Crivelli, E. 2006. Explosive volcanism during the Holocene in the Upper Limay River Basin: The effects of ashfalls on human societies, Northern Patagonia, Argentina. *Quaternary International*, 158 (1), 44–57. https://doi.org/10.1016/j.quaint.2006.05.016

Watt, S.F.L., Pyle, D.M., Mather, T.A., Martin, R.S. & Matthews, N.E. 2009. Fallout and distribution of volcanic ash over Argentina following the May 2008 explosive eruption of Chaitén, Chile. *Journal of Geophysical Research*, 114: B04207. https://doi.org/10.1029/2008JB006219

Watt, S.F.L., Pyle, D.M. & Mather, T.A. 2011a. Geology, petrology and geochemistry of the dome complex of Huequi volcano, southern Chile. *Andean Geology,* 38 (2), 335–348. https://doi.org/10.5027/andgeov38n2-a05

Watt, S.F.L., Pyle, D.M., Naranjo, J.A., Rosqvist, G., Mella, M., Mather, T.A. & Moreno, H., 2011b. Holocene tephrochronology of the Hualaihue region (Andean southern volcanic zone, ~42° S), southern Chile. *Quaternary International*, 246 (1–2), 324–343. https://doi.org/10.1016/j.quaint.2011.05.029

1397 Watt, S.F.L., Pyle, D.M. & Mather, T.A. 2013. Evidence of mid-to late-Holocene explosive rhyolitic eruptions
1398 from Chaitén Volcano, Chile. *Andean Geology,* 40 (2), 216–226.
1399 https://doi.org/10.5027/andgeoV40n2-a02

1400 Weller, D., Miranda, C.G., Moreno, P.I., Villa-Martínez, R. & Stern, C.R. 2014. The large late-glacial Ho
1401 eruption of the Hudson volcano, southern Chile. *Bulletin of Volcanology*, 76, Article: 831.
1402 https://doi.org/10.1007/s00445-014-0831-9

1403 Weller, D.J., Miranda, C.G., Moreno, P.I., Villa-Martínez, R. & Stern, C.R. 2015. Tephrochronology of the
1404 southernmost Andean Southern Volcanic Zone, Chile. *Bulletin of Volcanology*, 77, Article: 107.
1405 https://doi.org/10.1007/s00445-015-0991-2

1406 Weller, D.J., de Porras, M.E., Maldonado, A., Méndez, C. & Stern, C.R. 2017. Holocene tephrochronology of
1407 the lower Río Cisnes valley, souhern Chile. *Andean Geology*, 44 (3), 229–248.
1408 https://doi.org/10.5027/andgeov44n3-a01

1409 Weller, D.J., de Porras, M.E., Maldonado, A., Méndez C. & Stern, C.R. 2019. Petrology, geochemistry, and
1410 correlation of tephra deposits from a large early-Holocene eruption of Mentolat volcano, southern
1411 Chile. *Journal of South American Earth Sciences*, 90, 282–295.
1412 https://doi.org/10.1016/j.jsames.2018.12.020

1413 Wilkinson, M., Dumontier, M., Aalbersberg, I., G., Appleton, M., Axton, A., Baak, N., Blomberg, J-W., Boiten,
1414 L.B., da Silva Santos, P.E., Bourne, J., Bouwman, A.J., Brookes, T., Clark, M., Crosas, I., Dillo, O.,
1415 Dumon, S., Edmunds, C.T., Evelo, R., Finkers, et al. 2016. The FAIR Guiding Principles for scientific
1416 data management and stewardship. *Scientific Data*, 3: 160018 (2016).
1417 https://doi.org/10.1038/sdata.2016.18

1418 Yu, S.-Y., Shen, J., Colman, S.M. 2007. Modeling the radiocarbon reservoir effect in lacustrine systems.
1419 *Radiocarbon,* 49 (3), 1241–1254. http://doi.org/10.1017/S0033822200043150

1420