Insights into Magma Storage Beneath a Frequently Erupting Arc Volcano (Villarrica, Chile) from Unsupervised Machine Learning Analysis of Mineral Compositions

Felix O. Boschetty[1], David J. Ferguson[1], Joaquín A. Cortés[2], Eduardo Morgado[3], Susanna K. Ebmeier[1], Daniel J. Morgan[1], Jorge E. Romero[4], and Carolina Silva Parejas[5]

[1]Institute of Geophysics and Tectonics, School of Earth and Environment, University of Leeds, Leeds, UK
[2]Department of Geography & Geology, Edge Hill University, Ormskirk, UK
[3]Escuela de Geología, Universidad Mayor, Manuel Montt 367, Providencia, Santiago, Chile
[4]Department of Earth and Environmental Sciences, University of Manchester, Manchester, UK
[5]Escuela de Geología, Facultad de Ingeniería, Universidad Santo Tomás, Ejército Libertador 146, Santiago, Chile

Key Points:

- Unsupervised machine learning reveals previously undetected compositional clusters in Villarrica's crystal cargoes
- Thermodynamic models demonstrate magma storage occurs in a vertically extensive system containing variably evolved mushy reservoirs
- Temporal trends in crystal cargo contents and mineral zoning suggest magma mixing triggers unusual large mafic ignimbrites

Key words: Unsupervised Machine Learning, Crystal Cargoes, Thermodynamic Modeling, Magma Mixing, Large Mafic Ignimbrites, Villarrica

Corresponding author: Felix O. Boschetty, `eefob@leeds.ac.uk`

Abstract

A key method to investigate magma dynamics is the analysis of the crystal cargoes carried by erupted magmas. These cargoes may comprise crystals that crystallize in different parts of the magmatic system (throughout the crust) and/or at different times. While an individual eruption likely provides a partial view of the sub-volcanic plumbing system, compiling data from multiple eruptions can build a picture of a whole magmatic system. In this study we use machine learning techniques to analyze a large (>2000) compilation of mineral compositions from a highly active arc volcano: Villarrica, Chile. Villarrica's post-glacial eruptive activity (14 ka–present) displays large variation in eruptive style (mafic ignimbrites to Hawaiian style effusive eruptions) yet its eruptive products have a near constant basalt-basaltic andesite bulk-rock composition. What therefore, is driving explosive eruptions at Villarrica and can differences in storage dynamics be related to eruptive style? Here we use hierarchical cluster analysis to detect previously unseen structure in the composition of olivine, plagioclase and clinopyroxene crystals erupted at Villarrica, revealing the presence of compositionally distinct clusters within each crystal population. Using rhyolite-MELTS thermodynamic modeling we related these clusters to intensive magmatic variables: temperature, pressure, water content and oxygen fugacity. Our results provide evidence for the existence of multiple discrete (spatial and temporal) magma reservoirs beneath Villarrica where melts differentiate and mix with incoming more primitive magma. The compositional diversity within an erupted crystal cargo strongly correlates with eruptive intensity, and we postulate that mixing between primitive and differentiated magma drives explosive activity at Villarrica.

Plain Language Summary

Studies of volcanoes often focus on a single eruption. However, the magmatic systems beneath volcanoes are complex: magmas crystallize throughout the crust and the minerals erupted at the surface can be formed just prior to eruption, or thousands of years earlier. Here we use machine-learning methods to group mineral compositions from many eruptions of an active Chilean volcano (Villarrica) to build a picture of the magmatic system. By using the appropriate mathematical treatment, we find that there are distinct groups of mineral compositions that were not identified by past studies. These different compositions are used to demonstrate that different batches of magma have mixed throughout Villarrica's post-glacial history. This suggests that mixing of different magmas drives explosive eruptions at Villarrica volcano.

1 Introduction

Arc volcanoes produce most of the Earth's subaerial volcanic activity and are responsible for some of the largest historical eruptions (Siebert et al., 2015). However the structure of the magmatic systems that feed these volcanoes is still largely unknown. The traditional view that the magmas erupted from arc volcanoes reside within a melt-dominated sub-volcanic 'magma chamber' has been superseded in recent years by a more nuanced view of magma storage and supply, whereby melts ascend through a vertically-extensive series of melt-rich zones, termed a 'transcrustal magma system' (TCMS, Cashman et al., 2017). This conceptual model describes the complex processing of primary magmas throughout the entire crust by crystallization, assimilation, and mixing (e.g., Annen et al., 2006, 2015). This combination of processes provides a theoretical framework to investigate the origins of the variety of mineral compositions and textures found in a single eruption's crystal cargo.

Another recent advance in understanding magmatic systems is that crystal cargoes can be rapidly assembled from different parts of a magmatic system after protracted storage (e.g., Bergantz et al., 2015; Cooper & Kent, 2014; Mutch et al., 2019). Together with the TCMS model, they explain how a mineral assemblage erupted during a single event may contain: crystals formed from the carrier melt (autocrysts), crystals remobilized from other parts of the magmatic system (antecrysts), those from outside the magmatic system (xenocrysts), and crystals that form due to undercooling upon eruption (microlites) (Jerram & Martin, 2008). Therefore the crystal cargo of an eruptive deposit can be thought of as a snapshot of the underlying magmatic system. While an individual eruption likely provides a partial view of the sub-volcanic plumbing, compiling data from multiple eruptions can be used to build up a picture of the whole system. Building this complete picture requires well-characterized magmatic products from as many closely-spaced eruptions as possible.

Here we use established unsupervised machine learning techniques to analyze a large (>2000) compilation of mineral compositions from a highly active arc volcano: Villarrica, Southern Andes, Chile. We reveal previously unidentified structure in erupted mineral compositions and identify trends throughout Villarrica's eruptive history related to crystal zoning and eruptive style. We utilize thermodynamic modeling to constrain the physical and chemical characteristics of Villarrica's magmatic system to assess the suitability of the TCMS model. Finally we discuss the role of magma mixing in driving explosive behavior at Villarrica volcano, and the implications for future eruptions.

2 Unsupervised Machine Learning Applied to Geochemical Data

The term 'unsupervised machine learning' describes a class of algorithms that are designed to find patterns in multidimensional data. They are termed 'unsupervised' as they do not require any prior knowledge of the relationships between data. Unsupervised machine learning techniques can be used to gain insight into the structure of multivariate compositional data (data that describe quantity relative to a whole, i.e. close to a constant value such as 100%), which is common in geochemistry (e.g., Chiasera & Cortés, 2011; X. Liu et al., 2020; Templ et al., 2008). The main advantage of these methods is that they allow the data to be interrogated in a multivariate sense and can highlight trends that otherwise would be difficult to identify using traditional Harker-style bi-variate plots (Cortés, 2009). Cluster analysis is an unsupervised machine learning technique that attempts to group data by some measure of similarity. The most similar data points are iteratively combined, until all data points belong to a single cluster. Hierarchical cluster analysis describes the relationship of all the data points to each other during this process, producing a hierarchy of data clusters organized by their similarity.

The majority of past studies that utilize cluster analysis have used it to characterize the composition of volcanic products via analysis of whole-rock data sets, both for individual volcanoes (e.g., Mt Etna, Italy (Corsaro et al., 2013); Izu-Oshima, Japan (Kuritani et al., 2018)) and regional volcanism (e.g., the Virunga Volcanic Province (Barette et al., 2017)). Others have applied cluster analysis to individual phases from a single volcano e.g., silicate melt inclusions (Hamada et al., 2020); tephra glass (E. J. Liu et al., 2020); and minerals (Caricchi et al., 2020; Cortés et al., 2007; Gleeson et al., 2021). Some, but not all, of these studies have recognized that compositional data require special mathematical treatment prior to analysis. In this study we use a log-ratio approach (i.e., Aitchison, 1986) to transform the compositional data into a compatible Euclidean geometry.

3 Geological Background

Villarrica (39.5°S, 71.9°W) is a Quaternary stratovolcano in the Andean volcanic arc, and part of the Chilean Central Southern Volcanic Zone (Figure 1a). The volcano is Chile's most active, with more than 50 recorded eruptions since 1558 (Petit-Breuilh, 2004), and is considered by the Chilean geological survey (SERNAGEOMIN), as the most hazardous of the 92 geologically active Chilean volcanoes (SERNAGEOMIN, 2019). Villarrica's post-glacial (14 ka–present) eruptive activity displays a wide range in eruptive intensity and magnitude. This includes two major eruptive events, that generated the Licán (ca. 13.9 ka BP) and Pucón (ca. 3.7 ka BP) mafic ignimbrites, with estimated volumes of 10 and 5 km^3 (non-DRE), respectively (Lohmar et al., 2012; Silva Parejas et al., 2010). In contrast, historic eruptions (1900–present) have ranged from Hawaiian to violent Strombolian and are dominated by effusive lava flows (Pizarro et al., 2019). However, the paroxysmal March 2015 eruption, which lasted just 30 minutes, was characterized by a 1.5 km high fire fountain (Romero et al., 2018) and is the most recent demonstration of Villarrica's explosive potential. Despite this variety in post-glacial eruptive style, Villarrica's volcanic products have a limited compositional range: 98% of the juvenile whole-rock compositions collated in this study have 52–57 wt% SiO_2 (Figure 1). However, whole-rock data may not fully reflect the compositional variety of a magmatic system as heterogeneity may be present at a scale smaller than that of the whole rock sample (Pichavant et al., 2007). Therefore in this study, we focus on the main mineral phases erupted at Villarrica, whose compositions give insight into magma dynamics and the physical conditions of the magmatic system.

Typical Villarrica lavas and tephras are porphyritic with 10–15% modal crystals of mainly plagioclase feldspar and olivine, subordinate clinopyroxene, and small amounts of chromian spinel (Lohmar, 2008; Lohmar et al., 2012; Morgado et al., 2015; Pioli et al., 2015; Pizarro et al., 2019). Olivine crystals are usually euhedral to subhedral, often with resorbed rims. Clinopyroxene usually occurs as an unzoned subhedral phase, while plagioclase typically occurs as subhedral, reverse-zoned crystals with oxide inclusions, or as subhedral and unzoned crystals lacking inclusions. Groundmasses range from highly vesicular to highly crystalline, and are typically formed of plagioclase microlites and glass.

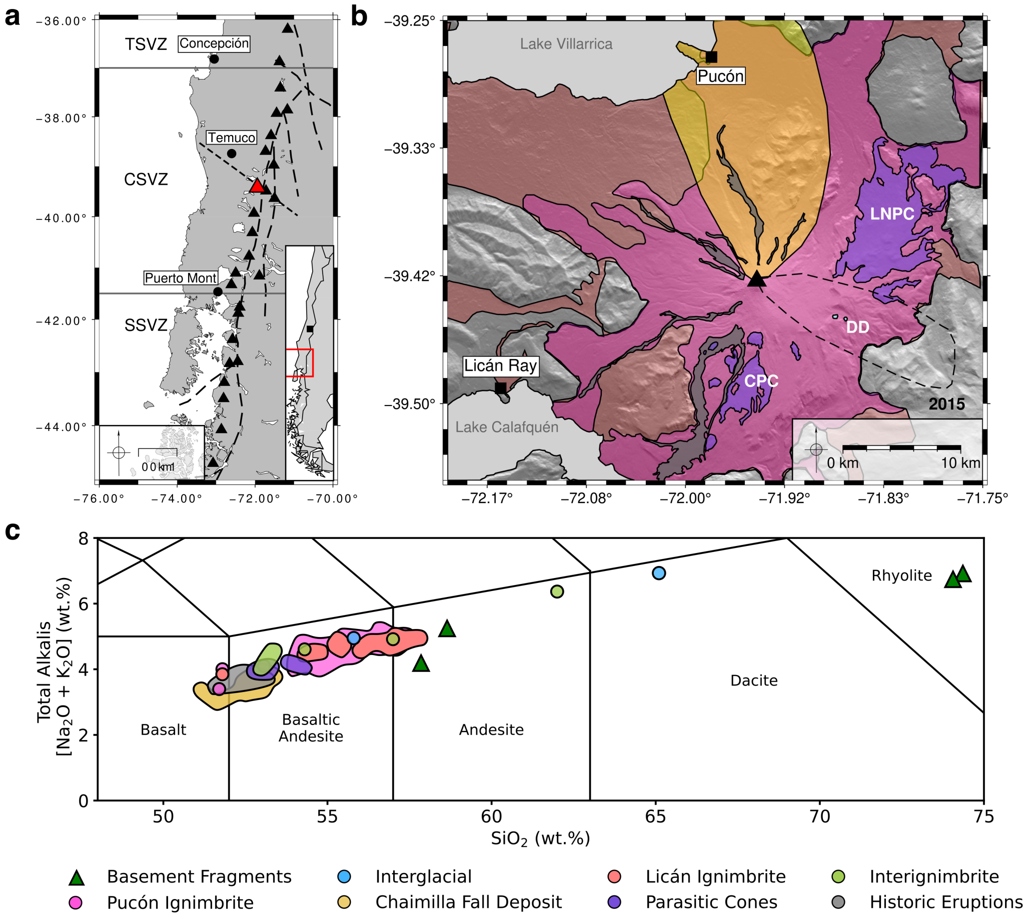

Figure 1. **a.** Map showing the location of Villarrica (red triangle) and other nearby active volcanoes (black triangles) in Chile's Southern Volcanic Zone (SVZ). The long-dashed black line is the Liquiñe-Ofqui Fault Zone, short-dashed black line is the Mocho-Villarrica Fault Zone (Cembrano & Lara, 2009). TSVZ, Transitional South Volcanic Zone. CSVZ, Central South Volanic Zone. SSVZ, Southern South Volcanic Zone (Stern, 2004). Inset shows the location of the CSVZ (red box) and the capital Santiago (black square). **b.** Simplified geological map (modified from Moreno and Clavero (2006)) showing the extent of key Villarrica eruptive deposits. The color of each deposit corresponds to the legend in 1c. LNPC, Los Nevados Parasitic Cones. CPC, Chaillupén Parasitic Cones. DD, Dacitic Domes. The location of the March 2015 tephra is shown by the black dashed line (Romero et al., 2018). **c.** Total-Alkali Silica (TAS, Le Bas et al., 1986) diagram showing the bulk rock compositions of Villarrica eruptions and basement rocks from past studies (Clavero-Ribes, 1996; Hickey-Vargas et al., 1989, 2016; Lohmar, 2008; McGee et al., 2017; Morgado et al., 2015; Pioli et al., 2015; Pizarro et al., 2019; Silva Parejas, 2008; Wehrmann et al., 2014; Witter et al., 2004; Zajacz & Halter, 2007). This shows the homogeneous nature of Villarrica eruptive products: the vast majority of the juvenile products (98%) have 52–57 wt% SiO_2.

4 Materials and Methods

4.1 Database Compilation

We compiled a database of existing electron microprobe analyses of the most abundant mineral phases that are present in almost all of Villarrica's erupted products: olivine, plagioclase and clinopyroxene. These were sourced from published studies (Costantini et al., 2011; Morgado et al., 2015; Pioli et al., 2015; Pizarro et al., 2019; Wehrmann et al., 2014; Witter et al., 2004; Zajacz & Halter, 2009), theses (Clavero-Ribes, 1996; Lohmar, 2008), and unpublished analyses. All data and their sources can be found in the Supplementary Material.

To ensure that only phases relevant to the magmatic system were considered, any analyses labeled as xenolith or microlite by the original authors were removed from the data sets. The analyses were then manually screened for errors e.g., misclassified, mixed-phase and poor-quality analyses. Any analysis without an analytical total between 98 wt% and 102 wt% was removed. The cations per formula unit (cfu) were calculated for each mineral analysis based on stoichiometry, using 4, 32 and 6 oxygens for olivine, plagioclase and clinopyroxene, respectively. Total iron was assumed to be entirely FeO for plagioclase and olivine analyses. The method of Droop (1987) was used to calculate the proportion of Fe^{2+} to Fe^{3+} in clinopyroxene analyses. Clinopyroxene analyses with a total cfu outside 4.00 ± 0.02 were removed. After screening, 2267 analyses (out of an initial 2611) were deemed suitable, these are broken down by eruption in Table 1.

4.2 Data Quality and Limitations

The database used in this study contains a large (>2200) total number of analyses, but the number of analyses per eruption is much smaller: only the 2015 eruption has nearly 100 analyses for either of the two most modally abundant minerals, olivine and plagioclase (Table 1). This has implications for relating compositions to textural features e.g., crystal zoning. Cheng et al. (2017) suggests that 100 or more analyses are required to characterize the compositional and textural features of a complexly zoned mineral population. However, we find no relationship between zoning and compositional clustering.

All eruptive deposits have not been sampled and analyzed equally. Table 1 shows that there are a higher number of analyses of high-volume eruptions (the Licán and Pucón

Table 1. Breakdown of Villarrica microprobe analyses included in the compiled database by eruption. Eruption stages are shown in italics where relevant. Eruptions are ordered from oldest to youngest. Radiometric ages and their source are shown where known.

Deposit	**Age (years BP)**	**Ol**	**Plag**	**Cpx**
Dacitic Dome	95,000 ± 15,000[a]	11	20	8
Intraglacial Pyroclastic Deposit	40,000–14,000[a]	3	9	8
Licán Ignimbrite	14,500–13,500[b,c]			
Initial Fall Deposit		2	6	1
Main Eruption		19	85	48
Pucura Lava	>10,600[d]	8	10	5
Afunalhue Pyroclastic Flow	4,090[d]	5	2	8
Pre-Pucón Surge		3	14	10
Pre-Pucón Lava		6	12	3
Pucón Ignimbrite	3,510-3,710[e]			
Initial Fall Deposit		6	8	4
Unit 1		27	31	32
Unit 2		27	56	33
Unspecified		25	56	40
Post-Pucón Lava		30	64	22
Chaimilla Fall Deposit	3,180 ± 40[f]			
Lower		58	110	20
Upper		64	71	19
Unspecified		6	12	10
Los Nevados Cones	<2,600 & >2,600[a]	25	42	26
Chaillupén Cones	<3,700 & >3,700[a]	10	11	17
1921		42	91	54
1948		19	67	31
1963		4	0	0
1971		57	71	17
1984		16	54	0
1999		49	48	9
2000		43	81	0
March 2015		142	64	1
post-March 2015		11	25	1
Unknown		6	0	0
Totals		724	1120	423

BP, Before Present. Ol, Olivine. Plag, Plagioclase. Cpx, Clinopyroxene. [a]Moreno and Clavero (2006). [b]Moreno (1993). [c]Clavero-Ribes (1996). [d]Lara and Clavero (2004). [e]Silva Parejas et al. (2010). [f]Costantini et al. (2011).

ignimbrites), historic eruptions, and those with easily accessible deposits. The main route up Villarrica, the Pucón Ski Center Road, cuts both the Pucón ignimbrite and the Chaimilla Fall Deposit and provides easy access to historic eruptions. Conversely, the Los Nevados and Chaillupén parasitic cones, and the Dacitic Domes are further from established roads and tracks. While this bias prevents us from commenting on individual eruptions that are poorly studied, a first order understanding of Villarrica's magmatic system is still possible.

Finally, there are likely systematic errors related to the different analytical equipment in different labs, which are in turn calibrated with different standards. This results in different analytical uncertainties for each of the past studies complicating direct comparison. There is also little overlap between the eruptions sampled by different studies. Without this, there is no way to quantify these potential inter-lab biases. Despite this, over half the data was obtained from a single lab, which reduces the chance of this having an effect (Lohmar, 2008). Furthermore, as discussed later in the paper, we find no systematic bias in the clustering that results from the source of the data.

4.3 Compositional Data Constraints and the Log-Ratio Transformation

Compositional data carry only relative information, subject to non-negative and constant-sum constraints (100 wt% etc.). These constraints mean that often used statistical methods that assume that data are normally distributed (i.e. unconstrained to a constant value and varying from $-\infty$ to $+\infty$), are not directly applicable. In his seminal work, Chayes (1971) established that it is the ratios between compositional data that measure variability rather than absolute differences. To circumvent this problem, Aitchison (1986) introduced several log-ratio transformations: linear transformations that map compositional data into an unconstrained Euclidean space.

There are several proposed log-ratio transformations: the additive log-ratio transformation (Aitchison, 1986), the centered log-ratio transformation (Aitchison, 1986) and the isometric log-ratio transformation (Egozcue et al., 2003). In this study, the mineral data sets were transformed using the isometric log-ratio (ilr) transformation (equation 1), implemented using the Pyrolite python library (M. Williams et al., 2020). This transformation was chosen as it ensures the transformed data have a non-singular covariance matrix and preserves the geometric properties of the raw compositional data (Egozcue et al., 2003):

$$ilr(x) = \sqrt{\frac{i}{i+1}} \ln\left[\frac{g(x_1,\ldots,x_i)}{x_{i+1}}\right], \;\; i = 1, 2, \ldots, D-1, \qquad (1)$$

where, x is a compositional analysis, i is a specific part, D is the number of parts (elements analyzed), and $g(x_i)$ is the geometric mean of the parts of x:

$$g(x_i) = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}}. \qquad (2)$$

A requirement of using any log-ratio transformation is that the data cannot contain zeros (Cortés et al., 2007; Fry et al., 2000). Zeros in compositional data can be structural (e.g., K in clinopyroxene), below detection of the analytical technique, or simply not analyzed (Fry et al., 2000). We removed structural zeros by only considering elements that reasonably exist in a mineral's structure. To deal with zeros resulting from detection limits, a detection limit of 0.05 wt% was assumed for all elements. The detection limits were generalized because detection limits were not reported in all cases. Zeros resulting from detection limits were replaced using the Multiplicative Replacement method of Martín-Fernández et al. (2000) and Fry et al. (2000), which is equivalent to distributing a detection limit threshold evenly among the below detection zeros. Not all elements (especially minor elements) were analyzed in every study. Elements that were measured for less than half of each of the mineral data sets were not used in our analysis. A table containing the elements used in cluster analysis for each of the mineral data sets can be found in the Supplementary Material.

Variables with low abundances but high relative variances (often minor and trace elements) have high log-ratio variances and therefore dominate any analysis of the complete log-ratio transformed data set (Baxter et al., 2005; Greenacre & Lewi, 2009; Greenacre, 2019). To prevent this, we normalized the log-transformed data set using the column (part) medians and standard deviations:

$$X' = x'_{ij} = \frac{x_{ij} - \tilde{x}_i}{\sigma_i} \qquad (3)$$

Where X' is the normalized data set, x_{ij} the i^{th} part of the j^{th} analysis, $\tilde{x}_i$ is the column median, and σ_i is the column standard deviation. The median was chosen over the arithmetic mean as the transformed data sets were non-normal: all four *ilr*-transformed data sets failed the Henze-Zirkler multivariate normality test with a specificity of 0.05

(Henze & Zirkler, 1990). Relationships between the transformed compositional data were then explored using hierarchical clustering methods over a Euclidean space.

4.4 Hierarchical Cluster Analysis

Hierarchical clustering algorithms attempt to group data into clusters by some measure of similarity. Agglomerative clustering methods start by grouping the two most similar data points into a cluster, then treating them as a single data point. The most similar data points or clusters are then iteratively combined until only a single cluster containing all the data remains. To cluster the data, a measure of dissimilarity (often called distance) must be chosen. A popular measure is the Euclidean distance which is the equivalent of Pythagoras's Theorem but over more than two dimensions:

$$d_E = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \tag{4}$$

Where d_E is the Euclidean distance matrix, i is the number of parts, and x and y are the points considered. Next, a linkage method must be chosen, i.e. a method describing how the distances between points and clusters are used to relate them. We used ward clustering (Ward, 1963), which minimizes the in-cluster variance, over alternatives such as single-linkage or average-linkage as it doesn't suffer from chaining (W. T. Williams & Lambert, 1966; Wishart, 1969). The hierarchical clustering algorithm was implemented using the scikit-learn python library (Pedregosa et al., 2011). Dendrograms depicting the resultant hierarchy were used to determine a suitable number of clusters independently for each mineral data set. Choosing the number of clusters that best represent the data set is a balance between showing global versus local variability: a small number of clusters will highlight the largest differences in the data at the cost of potentially obscuring more subtle differences, and vice versa for a large number of clusters. Figure 2 shows that the olivine and plagioclase data sets are well described by two clusters until almost half of their respective maximum distances. However, to maximize our ability to detect more subtle compositional changes between clusters we chose the next lowest number of clusters that well describe them. This was chosen by finding the maximum separation (second highest for olivine and plagioclase) between branches on the dendrogram. This results in three, five and four clusters for olivine, plagioclase and clinopyroxene, respectively.

We assessed the robustness of the identified clusters by repeatedly (1000x) resampling half of each mineral's data set and performing hierarchical clustering. the clusters from subsampled data were consistent with those identified in the complete data set, demonstrating that they are not strongly dependent on the size of our complete database. Further details can be found in the Supplementary Material.

4.5 Multivariate Outlier Detection

Each cluster contained points normally distributed about its center, therefore outliers are, according to the empirical rule for normal distributions, those values located at distances larger than three standard deviation from its center. To identify potential outliers, the Mahalanobis distance, i.e. the distance of a given data point x and a distribution (Mahalanobis, 1936), was calculated for each cluster:

$$D_M(\vec{x}) = \sqrt{(\vec{x} - \vec{\mu})^T C^{-1} (\vec{x} - \vec{\mu})} \tag{5}$$

where D_M is the Mahalanobis distance, $\vec{x}$ is a matrix of data points in each cluster, and $\vec{\mu}$ and C are the location and covariance estimators. The location and covariance estimators were robustly calculated using the Minimum Covariance Determinant estimator (MCD) (as in Filzmoser & Hron, 2008), implemented using the FastMCD algorithm (Rousseeuw & Driessen, 1999) available in the scikit-learn python package (Pedregosa et al., 2011). The Mahalanobis distances for each cluster can be approximated by a χ^2 distribution (Rousseeuw & Zomeren, 1990). A critical Mahalanobis distance was determined for each cluster, above which a point is considered an outlier:

$$D_M > (\chi^2_{p,1-\alpha})^{1/2}, \tag{6}$$

which is the square root of the upper-α quartile of the χ^2 distribution with p degrees of freedom, which were 5, 7, and 10 for olivine, plagioclase and clinopyroxene, respectively. The typical choice for (1-α) 0.975 (Rousseeuw & Zomeren, 1990) was used i.e., the outliers will be contained in the upper 2.5% of the χ^2 distribution. Any identified outliers were removed from the data sets.

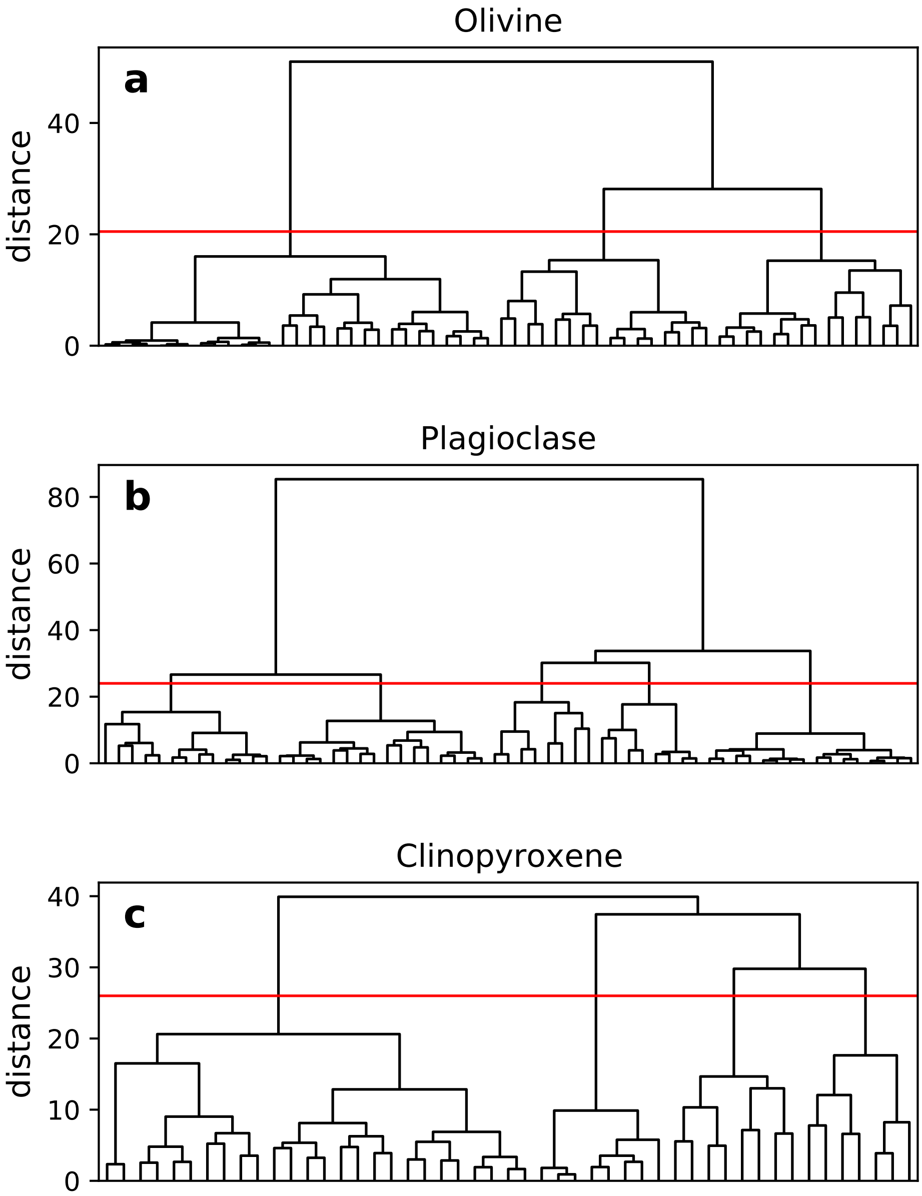

Figure 2. Dendrogram for each mineral data set, only the top five levels of each dendrogram are shown. The cutoff line which determines the number of clusters is shown by the red horizontal line.

4.6 Estimating Cluster Centers

To characterize each of the identified clusters, a measure of central tendency is required. For compositional data, the best unbiased estimator of the expected value is defined by the closed geometric mean (Cortés et al., 2007, and references within). To ensure that the center of the cluster is a valid mineral composition, the closed geometric mean of each cluster was calculated, and the nearest (according to a Euclidean distance matrix) data point to each mean was used to represent each cluster.

4.7 Thermodynamic Modeling

The current consensus, especially in arc settings, is that the majority of erupted minerals are remobilized antecrysts from a mushy magmatic system (Edmonds et al., 2019). This means it is unlikely an erupted mineral be in equilibrium with its carrier liquid, making traditional mineral-liquid thermobarometry impossible. For the same reason, mineral-mineral thermobarometry (e.g., olivine-augite) may not be accurate, even if both minerals appear to be in chemical equilibrium, unless textural evidence also suggests they are in equilibrium. To circumvent these issues we compare measured mineral compositions to many thermodynamic simulations of a simple Villarrica-like system, with a range of intensive variables. This avoids the need for mineral-liquid and mineral-mineral equilibrium.

Thermodynamic modeling was carried out to investigate the likely provenance of the identified mineral clusters with respect to the physical and chemical conditions of magma evolution at Villarrica. We used the rhyolite-MELTS v1.2.0. algorithm (Ghiorso & Gualda, 2015; Gualda et al., 2012), via the alphaMELTS 2.0 (originally Adiabat_1ph (Smith & Asimow, 2005)) front end, to model crystal fractionation from a primitive input melt. The initial bulk composition of each model was estimated using the most primitive glass composition reported at Villarrica: a post-entrapment-crystallization corrected, olivine-hosted melt inclusion (VL15B-ol5-inc1, Mg# 65 at QFM+1) from an upper unit of the Chaimilla Fall Deposit (Pioli et al., 2015). Isobaric fractional crystallization models were performed with temperature decreasing in 1°C increments from above the liquidus, at a range of pressures (25, 50, 75 and 100 MPa and then increments of 100 up to 700 MPa), variable initial water contents (0.0–6.0 wt% in increments of 0.5), and oxygen fugacity buffers (fO_2) QFM-0.5 (0.5 units below the Quartz-Fayalite-Magnetite buffer) to QFM+2.25, in increments of 0.25. A total of 1560 rhyolite-MELTS simulations were performed, one for each pressure, H_2O and fO_2 permutation.

,, ,

We identified the best-fit intensive variables by comparing the composition of minerals crystallized in each rhyolite-MELTS simulation to those in each identified cluster. We calculated the cfu, and then *ilr*-transformed the composition of olivine, plagioclase and clinopyroxene crystallized at each temperature step for all 1560 rhyolite-MELTS simulation by the same procedure as for the measured compositions. As detection limits are not defined sensu stricto in MELTS, any zeros in the simulated compositions were replaced with 0.0001 cfu. With both the simulated and measured compositions in the same Euclidean space, we used the Euclidean distance between each as a measure of similarity (equation 4). The best fitting intensive variables were those that produce the 'closest' compositions to those measured: those that have the smallest distance.

To back out the best-fit intensive variables for each cluster, we calculate the weighted arithmetic mean and weighted standard deviation of the best-fitting T, P, fO_2, and H_2O from all measured compositions that comprise the cluster. The weights used were the inverse of the Euclidean distance: the measured compositions that were best reproduced by a simulation are given a higher weight. The weighted average and standard deviation for each mineral cluster are given in Table 3 and plotted in Figure 7.

We used a Monte-Carlo approach to estimate the effects of the analytical uncertainty in the initial bulk-composition inputted into the Rhyolite-MELTS simulations and the analytical uncertainty in the clustered mineral compositions. The maximum resultant uncertainties are comparable to those of mineral-liquid or mineral-mineral thermobarometric methods: ca. ±50 °C, ±200 MPa, ±0.5 log units, and ±1.5 wt% H_2O. This is to be expected as both MELTS and thermobarometers are calibrated on experimental data sets from the Library of Experimental Phase Relations (LEPR). Further details can be found in the Supplementary Material.

5 Results

The compositions of minerals erupted in Villarrica's crystal cargoes are diverse, especially in comparison to whole-rock compositions (Figure 1c). Olivine compositions range from Fo_{50-87}, plagioclase An_{50-96}, and clinopyroxene Mg_{57-92}. Despite this range in mineral compositions, hierarchical cluster analysis and outlier detection were performed successfully on the three mineral data sets. We identified 4, 94 and 15 outliers for olivine, plagioclase and clinopyroxene, which corresponds to 0.6, 8.4 and 3.5% of each respective data set. The representative compositions (centers) of each identified mineral cluster are shown in Table 2. The clusters are ordered from most primitive to most evolved (e.g., Ol_1 is the most primitive olivine). Violin and box-plot diagrams showing the compositional differences for each mineral's clusters are included in the Supplementary Material. We compared the cluster designations with each composition's source study and found no obvious dependence. Cluster designations appear to be purely compositional.

To visualize the distribution of the identified clusters and verify that they were compositionally distinct, two techniques were chosen to reduce the dimensionality of the *ilr*-transformed data. First, principal component analysis (PCA) was performed and the resulting first and second principal components plotted (Figure 3a-c). Then a t-distributed Stochastic Neighbor Embedding (t-SNE) algorithm (Van der Maaten & Hinton, 2008) was implemented with a perplexity value of 40. 10,000 iterations were used to ensure that the projection was stable (Figure 3d-f). Both the PCA and t-SNE projections show that the clusters identified by hierarchical cluster analysis are well-defined and have good separation. The distribution of outliers implies that outlier detection was successful: outliers are evenly distributed among the identified clusters and do not form clusters of their own.

A major advantage of cluster analysis and the multivariate methods used in this study is that they allow patterns to be identified in multivariate space and reveal trends that are hard to detect using traditional bivariate plots. A good example of this is shown by the olivine data set. Figure 4 shows probability density functions (PDFs) for each of the mineral data sets and the identified clusters generated using Kernel Density Estimation (Silverman, 1986). The molar forsterite content (X_{Fo}=100·Mg/(Mg+Fe^{2+}+Mn)), molar anorthite content (X_{An}=100·Ca/(Ca+Na+K) in moles), and molar magnesium number

($Mg_{\#}$=100·Mg/(Mg+Fe^{2+})) were calculated, for olivine, plagioclase and clinopyroxene, respectively. The resulting PDFs were multiplied by the number of data points in each cluster, making them analogous to non-normalized histograms. The distribution of molar Fo for the complete olivine data set might suggest the existence of two clusters with

Table 2. Representative compositions of each identified mineral cluster, given in wt% oxide and cations per formula unit. The concentration of FeO and Fe_2O_3 in clinopyroxene was calculated with the method of Droop (1987). Clusters are ordered from most to least primitive.

Mineral	Olivine			Plagioclase					Clinopyroxene			
Cluster	Ol_1	Ol_2	Ol_3	Pl_1	Pl_2	Pl_3	Pl_5	Pl_5	Cpx_1	Cpx_2	Cpx_3	Cpx_4
SiO_2 (*wt%*)	39.67	38.30	37.21	45.56	48.08	51.30	53.67	57.61	51.51	50.15	51.29	51.52
TiO_2									0.46	0.78	0.53	0.67
Al_2O_3				33.27	32.95	30.23	28.61	27.22	2.29	3.59	2.52	1.86
Cr_2O_3									0.49	0.09	0.10	b.d.
Fe_2O_3	-	-	-	-	-	-	-	-	1.57	2.09	0.00	1.86
FeO	15.57	21.40	26.56	0.49	0.64	0.79	0.87	0.43	6.16	6.97	8.91	9.85
MnO	0.24	0.35	0.44						0.20	0.23	0.22	0.45
MgO	45.21	39.80	34.86	0.08	0.12	0.12	0.18	b.d.	16.88	14.51	14.59	14.52
CaO	0.18	0.25	0.24	17.59	16.93	13.79	12.66	9.58	19.04	20.55	19.93	19.32
Na_2O				1.26	2.03	3.76	4.31	5.82	0.27	0.29	0.23	0.27
K_2O				b.d.	b.d.	0.10	0.18	0.19				
Si (*mol*)	0.991	0.992	0.997	10.375	10.601	11.184	11.500	12.060	1.958	1.934	1.968	1.958
Ti									0.013	0.023	0.015	0.019
Al				4.464	4.281	3.883	3.612	3.358	0.051	0.082	0.057	0.091
Cr									0.007	0.001	0.002	b.d.
Fe^{3+}	-	-	-	-	-	-	-	-	0.022	0.030	0.000	0.027
Fe^{2+}	0.325	0.464	0.595	0.093	0.118	0.144	0.156	0.075	0.193	0.222	0.281	0.314
Mn	0.012	0.008	0.010						0.007	0.008	0.007	0.014
Mg	1.683	1.537	1.393	0.027	0.039	0.039	0.058	0.016	0.956	0.834	0.835	0.823
Ca	0.005	0.007	0.007	4.291	3.999	3.221	2.907	2.149	0.775	0.849	0.819	0.787
Na				0.278	0.434	0.795	8.94	1.181	0.010	0.011	0.009	0.010
K					b.d.	b.d.	0.014	0.025	0.025			
Fo (*mol%*)	83.60	76.54	69.71									
An				93.76	90.07	79.93	75.98	64.04				
$Mg_{\#}$									83.19	78.94	74.83	72.40

Blank cells denote elements not used in cluster analysis. b.d., below detection limit. -, not calculated. Fo, molar forsterite content. An, molar anorthite content. $Mg_{\#}$, magnesium number.

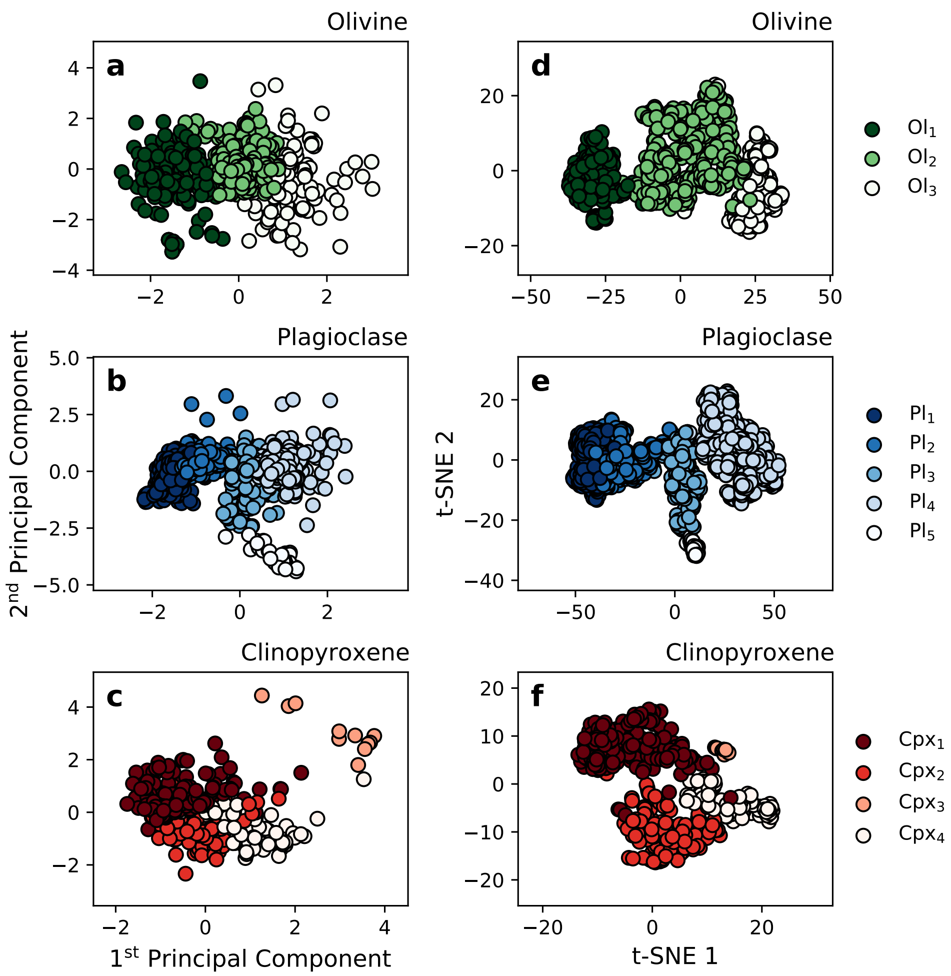

Figure 3. **(a-c)** Clustered data projected using the first and second principal components. **(d-f)** Clustered data projected using t-distributed stochastic neighborhood embedding (t-SNE) axes. Both projections show that the clusters identified by hierarchical clustering are well-defined.

Fo_{76} and Fo_{84}. However our clustering approach, which considers the concentration of multiple elements (including minor elements), detects three clusters which could not be identified looking at the PDF of the complete data set alone. This demonstrates how cluster analysis can identify and reveal otherwise hidden structure in mineral composition data sets by detecting similarities in all elements analyzed.

Olivine compositions were grouped into three clusters, whose cfu contents show the expected proportionality with Fo (Table 2). Fe^{2+}, Si and Mn are all inversely proportional to Fo. However, Ca varies independently of Fo and the other elements, and therefore has potential to give insight into differing parental magma minor-element compositions and/or differing H_2O-contents (Gavrilenko et al., 2016; Kamenetsky et al., 2006).

Plagioclase compositions were grouped into five clusters. These show expected trends with increasing An-content: increasing Ca and Al; and decreasing Si, Na and K (Table 2). However, Mg and Fe covary independently of the other elements and An-content. That they covary implies that the variations reflect one or more physical processes and are not due to the potentially high analytical uncertainties associated with Fe and Mg in plagioclase (Ginibre & Wörner, 2007). Variations in Fe and Mg in plagioclase independent of other elements, have been attributed to changes in Fe and Mg in the parental melt (Ginibre & Wörner, 2007; Singer et al., 1995), but might reflect melt-plagioclase disequilibrium caused by rapid growth (Ginibre & Wörner, 2007; Mollo et al., 2011; Singer et al., 1995).

Clinopyroxene compositions were grouped into four clusters. However, the representative compositions show more complex trends than for plagioclase and olivine (Table 2). The identified clusters were characterized as follows: (1) high-Cr and high-$Mg_{\#}$, (2) high-Al and high-Ti, (3) low-Fe^{3+}, and (4) low-Al and low-$Mg_{\#}$. This might be the result of the higher-complexity of the clinopyroxene mineral structure, compared to feldspar and olivine. However, the identified clusters still appear to show some expected correlations with $Mg_{\#}$, with the most primitive Cpx_1 having the lowest incompatible elements, e.g., Mn and Ti, and Cpx_4 having higher incompatible elements and Ca. Cpx_3 has much lower calculated Fe^{3+} content compared to the other three clusters which could reflect significantly different oxidation conditions during crystallization.

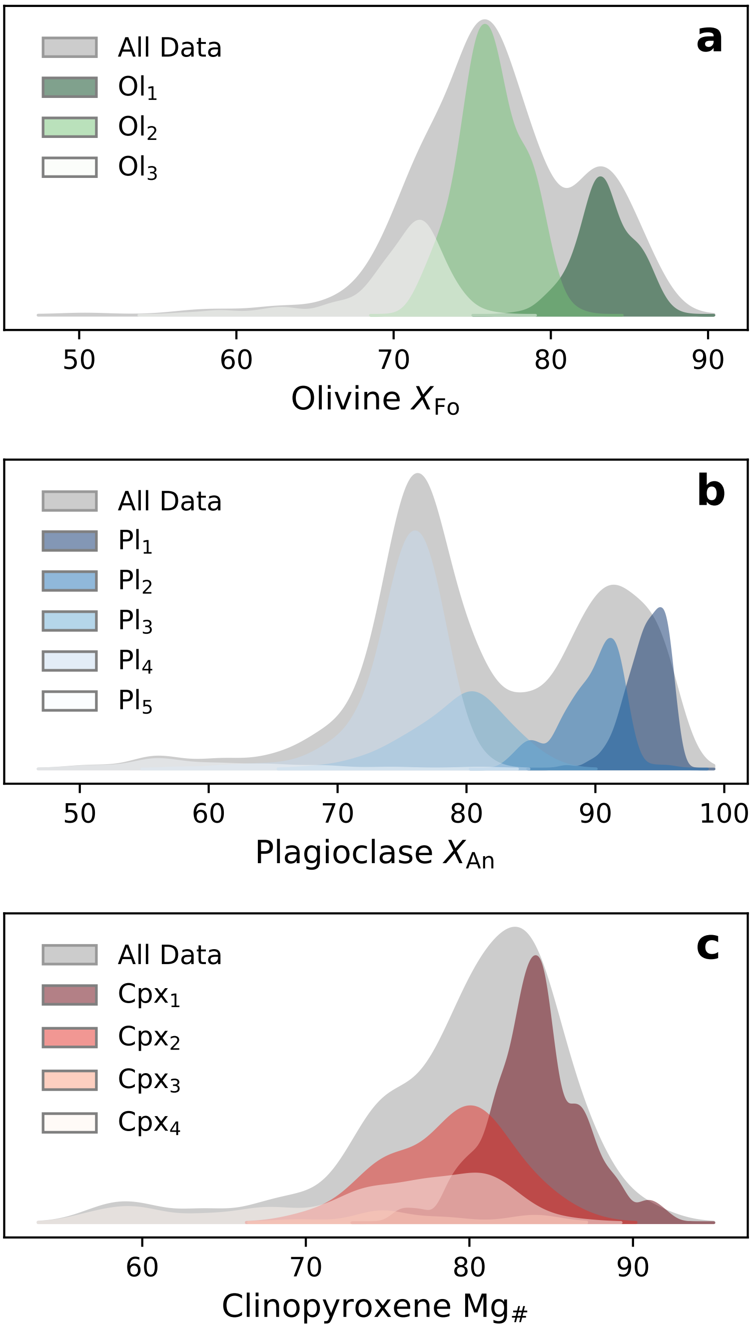

Figure 4. Plots of probability density functions (PDF) generated using Kernel Density Estimators (KDE) of each mineral subset and its clusters. Each PDF is scaled by the number of data points it was generated from, analogous to a histogram. Comparisons of the distribution of identified clusters with each respective complete data set highlights the ability of cluster analysis to identify otherwise hidden structure in mineral compositions.

5.1 Controls On Cluster Membership

5.1.1 Crystal Zoning

Original studies classified each mineral analysis as unzoned, zoned – cores, intermediates, or rims. However, this information was not reported for a significant number of analyses (olivine: 20%, plagioclase: 32%, clinopyroxene: 35%), and there is no way to judge the accuracy of the designations. Therefore it was assumed that each of the source studies categorized their analyses in the same manner. Figure 5 shows the proportion of crystal zones in each cluster for each category. There is no clear relationship between the cluster distribution and zoning for any of the three minerals.

5.1.2 Eruption Age

Figure 6 shows the proportion of phenocrysts from each cluster within individual eruptive deposits sorted by age. There is a stark contrast between the composition of crystal cargoes of historic eruptions (erupted in the last 100 years) and prehistoric eruptions (ca. 95–1.6 ka). Generally, the crystal cargoes of historic eruptions contain a smaller variety of crystal compositions. This is especially pronounced for feldspar and clinopyroxene compositions (Figure 6b,c). Both feldspar and clinopyroxene compositions become markedly more homogeneous through to the present day: erupted plagioclase compositions are dominated by Pl_4 and clinopyroxene compositions are dominated by Cpx_1. Generally, historic eruptions contain a higher proportion of Ol_2 than prehistoric eruptions, but olivine compositions are more variable than both plagioclase and clinopyroxene (Figure 6).

Of the three minerals considered in this study, olivine cluster proportions vary the most from eruption to eruption. For example, the Licán ignimbrite has a large proportion of primitive olivine (Ol_1) whereas the Pucón ignimbrite is made up almost entirely of more evolved olivine (Ol_3). The historic eruptions are dominated by more evolved olivine (Ol_2). However, the March 2015 eruption is distinct from the other historic eruptions, as it has large proportion of all three olivine clusters. Therefore the olivine portion of the March 2015 eruption's crystal cargo is more similar to prehistoric crystal cargoes, than the other historic eruptions.

The proportion of plagioclase clusters varies widely between the historic and prehistoric eruptions. Generally, the prehistoric eruptions contain a greater variety of pla-

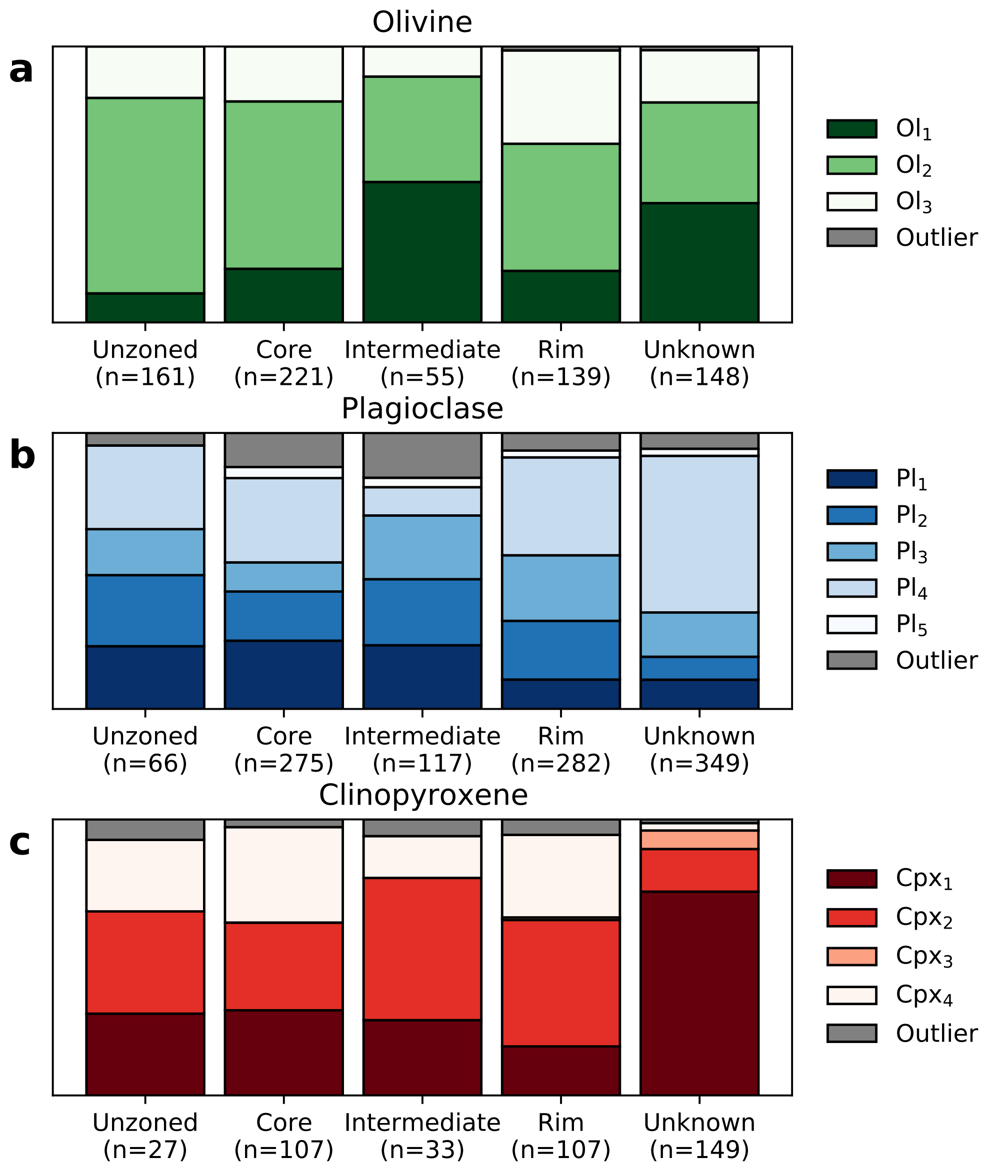

Figure 5. Stacked bar charts showing the effect of crystal zoning on cluster membership. The number of analyses in each category is shown in brackets. There is no clear pattern between cluster membership and crystal zoning.

gioclase compositions: the Dacitic Dome, Licán and Pucón ignimbrites are the only to contain all five plagioclase clusters. The most evolved plagioclase cluster (Pl_5) only occurs in the three oldest eruptions: the Dacitic Dome, and the Licán and Pucón ignimbrites. The most primitive plagioclase (Pl_1) is restricted almost entirely to prehistoric eruptions. In contrast, the crystal cargoes of the historic eruptions are dominated by the intermediate Pl_4, whose proportion increases through time from the 1921 through to the March 2015 eruption.

Clinopyroxene cluster membership broadly corresponds to that of olivine and plagioclase: the greatest amount of variation is present in the oldest eruptions. Clinopyroxene becomes both more scarce and more compositionally homogeneous in later eruptions: cluster Cpx_2, Cpx_3 and Cpx_4 are most abundant in the prehistoric eruptions while Cpx_1 is noticeably more abundant in the historic eruptions. Very little was erupted after the 1948 eruption.

5.1.3 Eruptive Center Location

Villarrica's flanks host over 30 parasitic or adventitious cones which form two groups: the Los Nevados group to the northeast and the Chaillupén group to the south (Moreno & Clavero, 2006) (Figure 1b). Other than the scoria and lavas of these two groups of cones, all other eruptive deposits are thought to originate from Villarrica's main vent, although the location of this main vent has changed through time (Moreno & Clavero, 2006). The cluster membership for the two groups of parasitic cones are shown in Figure 6 (labeled LNPC and CPC, respectively). The proportion of phenocrysts in each cluster for the two groups of cones are distinct. For example, the Chaillupén cones contain the most primitive olivine (Ol_1), and none of the most evolved clinopyroxene (Cpx_4), unlike the Los Nevados cones. This could be because the two sets of cones tap different parts of Villarrica's magmatic system, or they sample the system at different points in time. Compared to the eruptions from the central vent, the two sets of parasitic cones are more diverse than the historic eruptions, and most similar to the prehistoric eruptions. This is especially apparent when comparing the olivine and clinopyroxene cluster memberships. More detailed studies of the two groups of cones are needed to constrain their ages and their petrolgical differences to each other and products from the main vent (e.g., Robidoux et al., 2021).

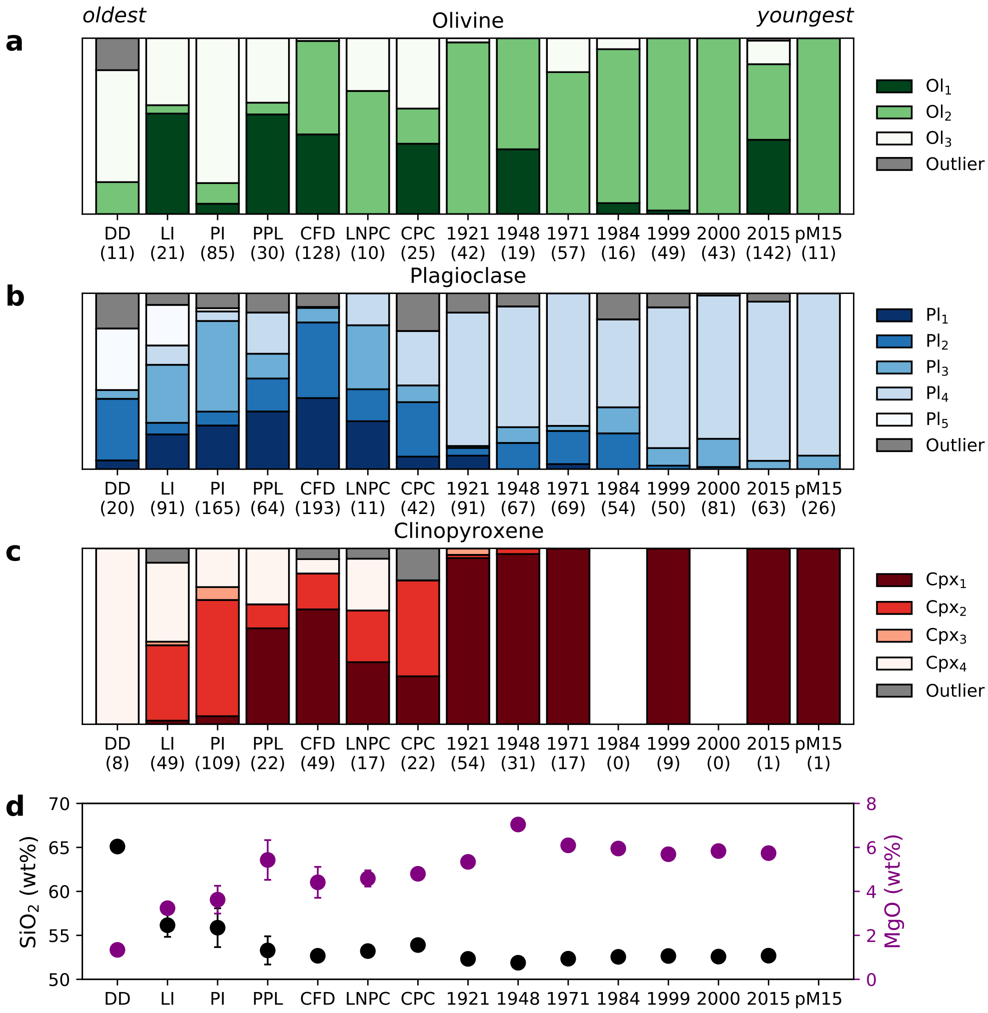

Figure 6. **(a-c)** Stacked bar charts showing the cluster membership of the different eruptions sorted by age. The number of analyses for each eruption are shown in brackets. Only those eruptions with more than 10 phenocryst compositions are shown. DD, Dacitic Dome. LI, Licán Ignimbrite. PI, Pucón Ignimbrite. PPL, Post-Pucón Lava. CFD, Chaimilla Fall Deposit. LNPC, Los Nevados Parasitic Cones. CPC, Chillupén Parasitic Cones pM15, post-March 2015. Historic eruptions are labeled by eruption year. **(d)** Plot of mean whole-rock SiO_2 and MgO contents. Error bars show ± 1 standard deviation. There are no whole-rock analyses for the post-March 2015 samples. Whole-rock data sources are the same as in Figure 1c.

5.1.4 Eruption Volume and Intensity

Villarrica's post-glacial eruptions show large variation in volume and intensity, most notable are the differences between the explosive high-volume mafic ignimbrites (Licán and Pucón) with the smaller-volume, comparatively-effusive, historic eruptions. The large volume of the ignimbrites and their association with caldera collapse suggests that a larger portion of the magmatic system was evacuated during those eruptions. This might be expected to produce products with a higher compositional variety, e.g., both evolved and primitive compositions, especially if the magmatic system were to contain multiple somewhat isolated bodies that differentiated independently. This is indeed reflected in the cluster memberships of the historic versus older eruptions. Generally, the more explosive, older eruptions have a higher diversity in mineral compositions, especially the Licán and Pucón ignimbrites. Of the historic eruptions the intense March 2015 eruption, that produced a 1.5 km high fire fountain, erupted more diverse olivine compositions than any other historic eruption.

5.2 Thermodynamic Modeling

The representative compositions identified by hierarchical clustering were well reproduced by the relatively simple model set up, shown by the low minimum distances (Table 3). However, the most primitive clusters (Ol_1 and Pl_1) have higher minimum distances than the other olivine and plagioclase clusters. This suggests that the initial bulk composition used, a primitive melt inclusion found at Villarrica, was not sufficiently primitive to reproduce exactly these compositions. However the calculated distances are not so different as to suggest that the models do not provide a reasonable indication of the conditions of crystallization.

The best fitting pressure and temperature conditions from the rhyolite-MELTS simulations agree with the broad estimates available from previous thermobarometric studies (Figure 7). Pressures from both our simulations and past thermobarometry imply that polybaric crystallization, extending to at least the mid to lower crust (ca. 600 MPa), is required to produce the variety of erupted mineral compositions at Villarrica. The predicted temperatures from rhyolite-MELTS (800-1100°C) are less than those calculated by thermobarometry (1050–1250°C). This discrepancy is likely due to the high (>2 wt%) water contents of majority of the best fitting rhyolite-MELTS simulations as the olivine-

Table 3. The weighted average of best fitting conditions for all compositions in each identified mineral cluster compared to rhyolite-MELTS thermodynamic models. One weighted standard deviation of the best fitting conditions per cluster is shown in brackets. Oxygen fugacity is shown in log units relative to the QMF buffer.

Mineral	Cluster	T (°C)		P (MPa)		$\Delta f O_2$		H_2O (wt%)		Min Dist	
Olivine	Ol_1	1090	(33)	382	(147)	0.73	(0.47)	4.65	(0.84)	0.042	(0.068)
	Ol_2	1048	(18)	245	(89)	0.05	(0.35)	3.47	(0.70)	0.024	(0.018)
	Ol_3	1018	(29)	267	(109)	-0.19	(0.29)	3.54	(0.97)	0.033	(0.031)
Plagioclase	Pl_1	1010	(16)	216	(46)	-0.23	(0.56)	5.24	(0.43)	0.133	(0.183)
	Pl_2	1031	(42)	185	(86)	0.16	(0.70)	4.14	(1.32)	0.035	(0.024)
	Pl_3	1050	(70)	232	(190)	0.70	(0.90)	2.67	(1.65)	0.031	(0.012)
	Pl_4	1055	(81)	311	(223)	0.79	(0.86)	2.36	(1.71)	0.044	(0.015)
	Pl_5	978	(74)	336	(245)	0.98	(1.01)	3.36	(1.69)	0.025	(0.012)
Clinopyroxene	Cpx_1	1006	(68)	57	(35)	1.91	(0.41)	0.79	(1.43)	0.029	(0.035)
	Cpx_2	1010	(64)	86	(74)	0.92	(0.61)	1.58	(1.72)	0.017	(0.018)
	Cpx_3	926	(103)	105	(108)	1.10	(0.54)	1.89	(2.49)	0.083	(0.074)
	Cpx_4	985	(73)	49	(27)	1.40	(0.44)	0.41	(1.12)	0.037	(0.065)

T, temperature. P, pressure. Min Dist, the minimum distance betwen *ilr*-transformed simulated and measured compositions.

augite geothermometer used by past studies was not calibrated on experiments with high water contents, and therefore likely overestimated temperatures (Loucks, 1996).

The most primitive olivine (Ol_1) formed at relatively high temperatures and pressures (1093 °C, 382 MPa). Furthermore, Ol_1 has both the highest fO_2 and water content (QFM+0.73, 4.65 wt%). This is in agreement with experimental and natural studies of olivine crystallization from basaltic melts (Feig et al., 2010; Gavrilenko et al., 2016). This suggests that it formed from relatively primitive undegassed melts with a relatively low residence time in the crust prior to crystallization. In contrast, the most evolved olivine cluster (Ol_4) formed at lower temperature, pressure, fO_2, and water content. The two most primitive plagioclase clusters (Pl_1 and Pl_2) have high water contents (¿4.0 wt%) and relatively high temperatures (°C), as expected from experimental studies (Panjasawatwong et al., 1995; Takagi et al., 2005; Waters & Lange, 2015). In contrast, the most evolved plagioclase (Pl_5) has a significantly lower crystallization temperature (978 °C), and a notably higher pressure and fO_2 (336 MPa, QFM0.98). All representative clinopyroxene compositions are best fit at lower pressures (–100 MPa), and at a variety of temperatures (970–1041°C) and water contents (0–2.0 wt%).

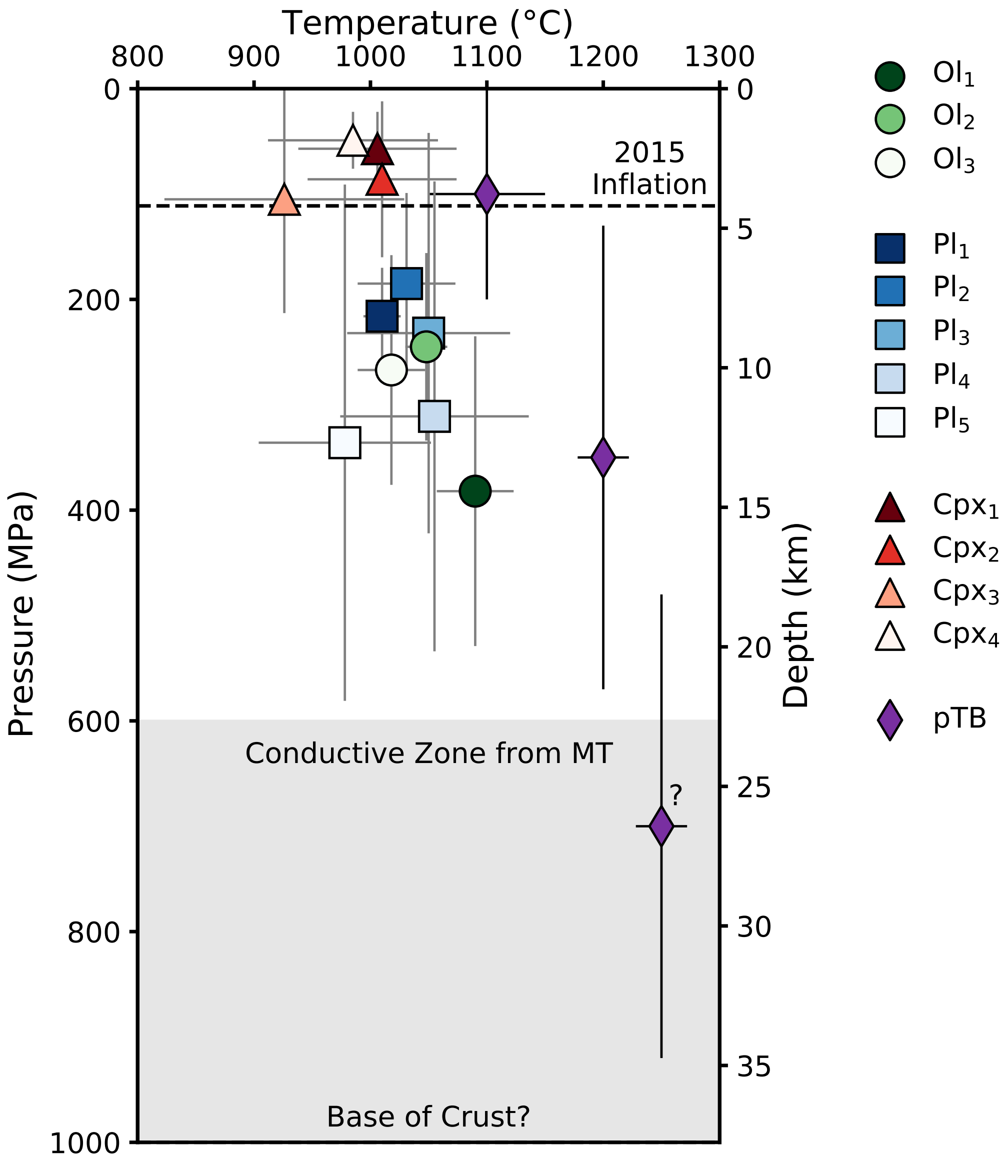

Figure 7. Weighted average of best fitting temperature and pressures from rhyolite-MELTS simulations for each of the identified mineral clusters. Gray error bars show ± one weighted standard deviation of best fit pressures and temperatures per cluster. pTb, compilation of past thermobarometry from past studies that correspond to a shallow, mid and deep-crustal storage (Lohmar, 2008; Lohmar et al., 2012; Morgado et al., 2015; Pioli et al., 2015; Pizarro et al., 2019; Witter et al., 2004). The error bars correspond to the range of uncertainty for each result. The deep region is based off a single olivine-augite pair and therefore is less certain than the shallow and mid-crustal results. The corresponding depth is calculated assuming a constant density of 2.7 gcm^{-3}. The black dashed line is the calculated depth (ca. 4.2 km) of reinflation after the March 2015 eruption (Delgado et al., 2017). The gray shading corresponds to an area of high conductivity (ca. 19-50 km) (Kapinos et al., 2016).

6 Discussion

6.1 Vertically and Laterally Extensive Magma Processing

Several models of volcanic systems that integrate geophysical, geochemical, and petrological information have concluded that magma processing beneath arc volcanoes likely takes place throughout the crust in trans-crustal magma systems (Annen et al., 2006, 2015; Cashman et al., 2017; Hildreth & Moorbath, 1988). There are multiple lines of evidence that suggest that this is the case at Villarrica volcano. We have compared representative mineral compositions identified via clustering, with thermodynamic fractional crystallization models of a simple Villarrica-like system. The comparisons suggest that polybaric crystallization, at pressures up to ca. 600 MPa, are required to produce the mineral compositions erupted at Villarrica (Figure 7). This is supported by thermobarometry from past studies that focused on individual eruptions (Figure 7). The majority of pressures calculated from thermobarometry are in the shallow crust (0–150 MPa), but pressures calculated from olivine-augite pairs erupted during historic eruptions extend to 700 MPa. Additionally, Kapinos et al. (2016) detected a low conductivity zone beneath Villarrica volcano that extends from 19–50 km, beyond the base of the crust (deeper than any of the crystallization pressures identified in our study).

As well as vertical connectivity, there is also evidence of lateral connectivity within Villarrica's magmatic system (e.g., Ebmeier et al., 2018; Lerner et al., 2020) (Figure S1). Nested calderas, kilometers in diameter, were produced during large eruptions at Villarrica and surround the current central vent: Caldera 1 (ca. 100ka), Caldera 2 (ca. 14ka), and Caldera 3 (ca. 3.7ka) (Moreno & Clavero, 2006). Calderas 1 and 2 cover an area of 6.5 by 4.2 km (long axes), Caldera 3 is roughly circular and 2 km in diameter. Domes and sills erupted on caldera walls demonstrate lateral transport of magma during large eruptions (Moreno & Clavero, 2006). After the March 2015 eruption, Delgado et al. (2017) detected reinflation at depths of 4.2 km in part of Villarrica's magmatic system, ca. 5 km SE of the central vent near the edge of Calderas 1 and 2. Further afield still, are the two groups of parasitic cones on Villarrica's flanks. The Los Nevados group contain fissures and cones that extend to ca. 10 km to the NE and the Chaillupén group extend ca. 12 km to the south of the present central vent. Recently, Pavez et al. (2020) detected a low conductivity zone (ca. 4 km deep) associated with the Los Nevados group. Combined, all these lines of evidence demonstrate that Villarrica's magmatic system is spatially extensive and imply that different parts of the system are tapped to accumulate the crystal cargo of each eruption.

6.2 The Role of Magma Mixing

The mixing of compositionally distinct magmas has been suggested as the trigger for eruptions at many arc volcanoes globally (e.g., Cassidy et al., 2015; Bouvet de Maisonneuve et al., 2012; Kahl et al., 2011; Ruprecht et al., 2012) , to the extent that it is considered ubiquitous (Cashman & Edmonds, 2019). Here we define magma mixing as the physical interaction of magmas that have different chemical composition and/or intensive variables. It doesn't imply total homogenization of the two magmas.

The compositional variety observed within erupted crystal cargoes , combined with evidence for spatially extensive magma processing, suggest that Villarrica's magmatic system comprises multiple, variably-evolved reservoirs distributed throughout the crust. This is supported by widely varying best fitting intensive variables from rhyolite-MELTS thermodynamic simulations (Table 3). In addition to the range of best-fit pressures, there is a large range of best fitting crystallization temperatures that reproduce crystal compositions (850–1100°C). There is also a large variation in best-fitting water contents from anhydrous to 5.5 wt%. This implies that degassing plays a large role in driving crystallization in Villarrica's magmatic system, as has been suggested for other arc volcanoes (e.g., Bouvet de Maisonneuve et al., 2012; Blundy et al., 2006). The large range of best fitting fO_2 values (-0.5–2.00 ΔQFM) suggests that crystallization is occurring from melts that have undergone different amounts of degassing and fractional crystallization (Carmichael, 1991; Lindsley & Frost, 1992; Sato, 1978). The presence of minerals produced by different intensive variables in a single crystal cargo means there must be physical interaction between reservoirs in Villarrica's magmatic system Reservoirs that are infrequently disturbed by ascending primitive magma, are able to cool and differentiate via fractional crystallization to produce evolved crystal compositions e.g., Pl_5 (An_{64}). Prior to eruption, ascending primitive melt interacts with multiple reservoirs, accumulating differing minerals. This mixing of magmas with different compositions produces new compositions which may be recorded in zoned crystals (Figure 5). These antecrysts are accumulated as melts ascend resulting in the variety of erupted mineral compositions in a single crystal cargo (Figure 6).

The important role of magma mixing at Villarrica is supported by our observation that the proportion of identified clusters do not show clear trends when they are compared to crystal zoning (Figure 5). If fractional crystallization is the dominant process during magma processing in Villarrica's magmatic system, we would expect mineral cores to be dominated by primitive compositions and rims by more evolved compositions. However, both primitive and evolved clusters are present as cores, intermediate, and rim zones of all three of the commonly erupted minerals. This implies that magma mixing plays a role in assembling crystal cargoes at Villarrica (e.g., Ruprecht et al., 2012; Streck, 2008)

Furthermore, textural evidence for magma mixing is present in almost all Villarrica's eruptive deposits. The Dacitic Dome contains reverse and oscillatory-zoned plagioclase, olivine with fayalitic rims (Fo_{58}) and reverse-zoned clinopyroxene (Lohmar, 2008). The Licán ignimbrite has both reverse and oscillatory-zoned plagioclase, reverse-zoned clinopyroxene, and orthopyroxene rims surrounding olivine crystals (Lohmar et al., 2012). A pre-Pucón surge deposit contains banded pumice, the pale bands have a bulk SiO_2 of 63 wt% (Moreno et al., 1994; Lohmar, 2008). The Pucón ignimbrite contains plagioclase crystals with resorbed cores and rims, sieve textures, and low-An microlites. Clinopyroxene is often reverse-zoned (Lohmar, 2008). Additionally, what is thought to be a dacitic enclave in a basaltic scoria bomb of the Pucón ignimbrite has been observed (McCurry & Schmidt, 2001; McCurry et al., 2004). The Chaimilla Fall Deposit contains complexly zoned plagioclase crystals with both reverse rims and widespread evidence of resorption (Pioli et al., 2015). Multiple historic eruptions also contain similar indicators: An-poor plagioclase is present as both rims of oscillatory-zoned crystals and crystals with resorption textures. Both olivine and clinopyroxene crystals display resorption textures (Morgado et al., 2015; Pizarro et al., 2019).

The ubiquity of evidence for magma mixing in Villarrica's eruptive products implies complex magma dynamics involving multiple variably-evolved reservoirs. These may be intermittently connected and tapped prior to eruption resulting in the variable crystal cargoes of eruptive deposits.

6.3 Linking Crystal Cargo and Whole-Rock Composition to Eruptive Behavior

The complex temporal trends in the compositions of erupted crystal cargoes show strong correlations with both whole-rock composition and eruptive style. There is a stark contrast between the composition of crystal cargoes of historic eruptions and prehistoric eruptions (Figure 6). Generally, the crystal cargoes of historic eruptions contain a smaller variety of crystal compositions, this is especially pronounced for feldspar and clinopyroxene compositions. This disparity strongly correlates with whole-rock compositions (Figure 6d): historic eruptions have average bulk SiO_2 and MgO contents of 54 and 5 wt%, respectively, whereas prehistoric eruptions range from 54–65 wt% SiO_2 and 1–5 wt% MgO. Both the composition of erupted crystal cargoes and whole-rock compositions strongly correlate with eruptive style: the most differentiated eruptive deposits contain the more variable crystal cargoes and are the most explosive and high-volume.

6.3.1 Dacitic Dome

The Dacitic Dome (c.a. 95 ka) is the most differentiated of Villarrica's products analyzed to date, with substantially higher bulk SiO_2 content than any subsequent eruption (65 wt% vs. 55 wt%, Figure 1). However the magma that formed the Dacitic Dome is almost certainly a mix of evolved and more primitive magmas (e.g., Eichelberger et al., 2006; Reubi & Blundy, 2009), supported by textural observations described above (Lohmar, 2008). The dacitic composition may have formed from mafic magmas ascending through evolved mushes (e.g., Reubi & Blundy, 2009) or from differentiated melts ascending through mafic mushes (e.g., Kent et al., 2010). Of the three eruptions that contain it, the crystal cargo of the Dacitic Dome contains the highest proportion of An-poor plagioclase (Pl_5, Figure 6b), along with its high bulk-SiO_2 content.

6.3.2 The Licán and Pucón Mafic Ignimbrites

The Licán (ca. 14 ka) and Pucón (ca. 3.7 ka) mafic ignimbrites are deposits from the two most explosive, high-volume eruptions at Villarrica (Lohmar et al., 2012; Silva Parejas, 2008). Both have high, but variable, bulk-SiO_2 contents (Figure 6d) compared to historic eruptions. They are also the last two eruptions whose crystal cargoes contain the evolved Pl_5. Their whole-rock chemistry correlates with the proportion of Pl_5 in their crystal cargoes: the Licán has lower MgO and slightly higher SiO_2 than the Pucón, and has a higher proportion of Pl_5 (Figure 6b,d).

Rhyolite-MELTS models show that Pl_5 likely crystallized from an evolved melt (ca. 60 wt% SiO_2), after substantial differentiation via crystal fractionation, resulting in a bulk crystal fraction of at least 50%. The eruption of a reservoir with such a high crystal fraction requires liberation via mush disaggregation to reduce its viscosity (Sparks & Marshall, 1986). This may result from a combination of chemical or thermal mixing, and/or volatile fluxing from an incoming more primitive basaltic melt (e.g Bachmann & Bergantz, 2006; Bergantz et al., 2015; Bouvet de Maisonneuve et al., 2012; Pistone et al., 2017; Zellmer et al., 2016). Upon mixing with the primitive magma, plagioclase crystals from the evolved reservoir would become reversely-zoned, with low-An cores and higher-An rims. In turn, primitive plagioclase carried by this primitive melt would grow low-An rims upon mixing (as suggested in Lohmar, 2008; Lohmar et al., 2012). Therefore mixing between an evolved and primitive magma explains how Pl_5 can exist as cores, intermediates, and rims (Figure 5).

The correlation of whole-rock compositions with the proportion of low-An Pl_5 in erupted crystal cargoes, combined with textural data and thermodynamic modeling, supports the triggering of these mafic-ignimbrite generating eruptions via destabilization of a differentiated (dacitic) mush by an influx of primitive magma.

6.3.3 The March 2015 Fire Fountain

In contrast to the ignimbrite-forming eruptions, historic eruptions have been dominated by effusive activity, punctuated by fire fountaining which appears to be related to lahar generation (e.g., 1908, 1948-49, 1963, 1964 1971 and 2015, (Lara & Clavero, 2004)). The paroxysmal eruption of March 2015 is significant in that it was the most intense historic eruption at Villarrica, producing a fire fountain that was 1.5 km in height but lasted just thirty minutes (Romero et al., 2018). The composition of olivine erupted in 2015 are more varied than both all preceding historic eruptions, and in spatter erupted later that year (Figure 6). This is in stark contrast to the March 2015's otherwise homogeneous cargo, which consists of only intermediate plagioclase clusters Pl_3 and Pl_4 and few clinopyroxene crystals.This implies that the mafic portion of the March 2015's crystal cargo was accumulated by primitive melt as it ascended through Villarrica's magmatic

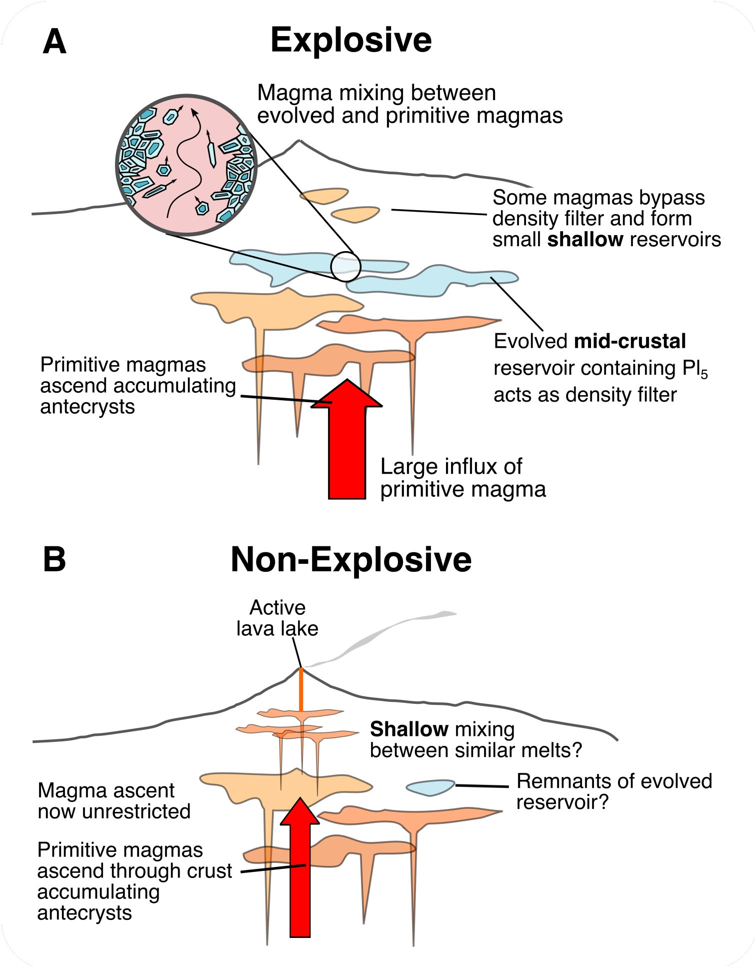

Figure 8. Conceptual model for Villarrica's magmatic system prior to **A.** an explosive, ignimbrite-forming eruption and **B.** a non-explosive historic eruption.

system. Plagioclase may have formed in response to degassing upon ascent (as in Blundy et al., 2006) as with the other historic eruptions. The low-SiO_2 and high-MgO bulk composition of erupted spatter suggests that the intensity of the eruption is unlikely to be caused by the same mechanism as the mafic ignimbrites. Instead high primary volatile contents likely drove fast magma ascent, resulting in the vigorous fountaining behavior (Allison et al., 2021; Barth et al., 2019; La Spina et al., 2021) and allowing it to punch through existing reservoirs and assemble its more varied crystal cargo.

6.4 A Model for Generating Explosive Eruptions at Villarrica Volcano

The eruption of the Dacitic Dome (ca. 95 ka) demonstrates that evolved portions of Villarrica's magmatic system existed prior to the most explosive post-glacial eruptions known at Villarrica, the Licán ignimbrite. Therefore portions of Villarrica's magmatic system were differentiated beyond the typical whole rock compositions of erupted historic products (Figure 1). The existence of this evolved (dacitic) reservoir has implications for magma dynamics. It has been shown that felsic mushes are sufficiently viscous that they are non eruptible (Marsh, 2002; Sparks & Marshall, 1986). In combination with their low relative density, they can act as a density filter, preventing more dense mafic melts from ascending to the surface (Kent et al., 2010). Therefore the establishment of a significant volume of evolved mush within the magmatic system could dampen the eruption of mafic magma and increase the overall volume of evolved melt until a sufficiently large volume of ascending primitive melt is able to destabilize it (Sparks & Marshall, 1986).

The eruption of large-volume eruptive deposits, such as Licán and Pucón ignimbrites, requires significant volumes of mobile magma to exist within the magmatic system (Druitt & Sparks, 1984). If the majority of Villarrica's magmatic system is composed of near-solidus mush (as implied by the TCMS model), this would require a large volume of magma to destabilize it (Marsh, 2002; Sparks & Marshall, 1986). The composition of minerals (Ol_1 and Pl_1, Table 2) and melt inclusions at Villarrica suggest primitive magmas were present within its system (Pioli et al., 2015). The best fitting conditions for the most evolved plagioclase (Pl_5) are in the middle crust, at low temperatures and after significant crystal fractionation (Figure 7). Therefore we propose that the trigger of the Licán and Pucón ignimbrites was a large influx of primitive magma, which mixed with an evolved reservoir (Figure 8a). The remnants of the mush that this primitive magma interacted with is shown by the presence of evolved plagioclase, Pl_5 in both ingimbrites' crystal cargo. In the case of the Licán ignimbrite, the trigger may have been facilitated by widespread deglaciation in the Southern Andes around 14 ka ago, altering the stress state of the crust to facilitate magma ascent (e.g., Jellinek et al., 2004; Rawson et al., 2016; Watt et al., 2013; Wilson & Russell, 2020). After the eruption of the Licán ignimbrite, another evolved reservoir of similar composition likely accumulated over a period of ca. 10 ka. The existence of an evolved reservoir prior to the Pucón Ignimbrite (ca. 3.7 ka) is suggested by the mineralogy of a Pre-Pucón pyroclastic surge deposit Lohmar (2008), which contains a large proportion of evolved (63 wt% SiO_2) pumice. The Pucón ignimbrite forming eruption may then have been triggered in a similar fashion to the Licán eruption, i.e. destabilization of an evolved reservoir by an influx of primitive magma.

These evolved reservoirs were likely depleted by mixing with primitive magma prior to the ignimbrite-forming eruptions, as subsequently erupted crystal cargoes (3.7 ka to present) contain no trace of Pl_5 (Figure 6). Without the density filter provided by the evolved reservoirs, the most recently erupted magmas were able to ascend through Villarrica's magmatic system faster, perhaps only mixing in the shallowest parts, resulting in less differentiated erupted compositions (Figure 8b). The absence of mixing between highly differentiated and primitive compositions results in substantially lower explosivity. Instead recent lava fountaining is likely driven by rapid ascent, enabled by high volatile-contents (e.g., the March 2015 eruption). The near continuous activity at Villarrica's lava lake over the last ca. 40 years demonstrates the ability of magma to ascend relatively unhindered through the magmatic system. Absence of activity at the summit and/or the eruption of crystals in equilibrium with differentiated melt compositions (e.g., Pl_5) may signal the onset of evolved reservoir development, and signal an increased likelihood of high-volume explosive eruption occurring.

7 Conclusions

Use of multivariate cluster analysis has allowed us to identify previously unidentified structure in the composition of minerals erupted throughout Villarrica's eruptive history. Comparisons of identified representative compositions with >1500 Rhyolite-MELTS thermodynamic simulations show that magma processing at Villarrica takes place at a range of pressures, temperatures, water-contents, and oxygen fugacities. These thermodynamic models strongly suggest that magma storage beneath the volcano is characterized by a series of ephemerally-connected, variably-evolved mush dominated sills. Prior to eruption, magma ascends through this trans-crustal magmatic system, accumulating different antecrysts which are erupted as variable crystal cargoes.

We have identified temporal trends in the composition of erupted crystal cargoes that correlate with trends in whole-rock composition and eruption style. We propose that prior to high-volume, ignimbrite-forming explosive mafic eruptions (the Licán and Pucón ignimbrites) much of Villarrica's magmatic system had differentiated by fractional crystallization. These mid-crustalevolved reservoirs acted as density filters, suppressing the eruption of small volumes of primitive magma. Only when a sufficiently-large influx of primitive magma entered the system was this evolved portion destabilized. Magma mixing ensued, producing the complex textures observed by past studies, and resulting in the relatively primitive, basaltic-andesite compositions of the two ignimbrites. The only remnants of the evolved reservoir are low-anorthite plagioclase which is present as both cores and rims of more primitive crystals.

The lack of evidence for an evolved reservoir after these high-volume explosive eruptions explains the homogeneous, more primitive, whole-rock compositions of subsequent eruptions. Without the density filter of the evolved reservoirs, more primitive melts ascend relatively uninhibited, mixing and degassing in the shallow subsurface before eruption. Further petrological work utilizing melt inclusion compositions, trace-element data, and diffusion chronometry will allow further investigation of magma dynamics at Villarrica.

Data Availability Statement

All compositional data used in the clustering analysis for this study are available in the Supplementary Material and from the National Geoscience Data Center (NGDC) at https://doi.org/10.5285/703acf75-8996-45a4-b4b3-42afca269a1c.

Acknowledgments

FOB is funded by the National Environmental Research Council (NERC) Panorama Doctoral Training Partnership (DTP) NE/S007458/1, SKE is funded by a NERC Independent Research Fellowship (NE/R015546/1). JR is supported through a Dean's Doctoral Scholar Awards of the University of Manchester and NSFGEO-NERC-funded project DisEqm (NE/N018575/1) and V-PLUS projects. All authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. The authors would like to thank Silke Lohmar for facilitating access to compositional data. We are grateful to Luca Carrichi and an anonymous reviewer for critical reviews that helped to greatly improve this paper, and Paul Asimow for editorial suggestions.

References

Aitchison, J. (1986). *The statistical analysis of compositional data.* London ; New York: Chapman and Hall.

Allison, C. M., Roggensack, K., & Clarke, A. B. (2021, December). Highly explosive basaltic eruptions driven by CO2 exsolution. *Nature Communications*, *12*(1), 217. doi: 10.1038/s41467-020-20354-2

Annen, C., Blundy, J. D., Leuthold, J., & Sparks, R. S. J. (2015). Construction and evolution of igneous bodies: Towards an integrated perspective of crustal magmatism. *Lithos*, *230*, 206–221. doi: 10.1016/j.lithos.2015.05.008

Annen, C., Blundy, J. D., & Sparks, R. S. J. (2006). The Genesis of Intermediate and Silicic Magmas in Deep Crustal Hot Zones. *Journal of Petrology*, *47*(3), 505–539. doi: 10.1093/petrology/egi084

Bachmann, O., & Bergantz, G. W. (2006, January). Gas percolation in upper-crustal silicic crystal mushes as a mechanism for upward heat advection and rejuvenation of near-solidus magma bodies. *Journal of Volcanology and Geothermal Research*, *149*(1-2), 85–102. doi: 10.1016/j.jvolgeores.2005.06.002

Barette, F., Poppe, S., Smets, B., Benbakkar, M., & Kervyn, M. (2017). Spatial variation of volcanic rock geochemistry in the Virunga Volcanic Province: Statistical analysis of an integrated database. *Journal of African Earth Sciences*, *134*, 888–903. doi: 10.1016/j.jafrearsci.2016.09.018

Barth, A., Newcombe, M., Plank, T., Gonnermann, H., Hajimirza, S., Soto, G. J., … Hauri, E. (2019, December). Magma decompression rate correlates with explosivity at basaltic volcanoes — Constraints from water diffusion in olivine. *Journal of Volcanology and Geothermal Research*, *387*, 106664. doi: 10.1016/j.jvolgeores.2019.106664

Baxter, M. J., Beardah, C. C., Cool, H. E. M., & Jackson, C. M. (2005). Compositional Data Analysis of Some Alkaline Glasses. *Mathematical Geology*, *37*(2), 183–196. doi: 10.1007/s11004-005-1308-3

Bergantz, G. W., Schleicher, J. M., & Burgisser, A. (2015). Open-system dynamics and mixing in magma mushes. *Nature Geoscience*, *8*(10), 793–796. doi: 10.1038/ngeo2534

Blundy, J., Cashman, K., & Humphreys, M. (2006, September). Magma heating by decompression-driven crystallization beneath andesite volcanoes. *Nature*, *443*(7107), 76–80. doi: 10.1038/nature05100

Bouvet de Maisonneuve, C., Dungan, M., 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. doi: 10.1016/j.jvolgeores.2011.09.010

Caricchi, L., Petrelli, M., Bali, E., Sheldrake, T., Pioli, L., & Simpson, G. (2020). A Data Driven Approach to Investigate the Chemical Variability of Clinopyroxenes From the 2014–2015 Holuhraun–Bárdarbunga Eruption (Iceland). *Frontiers in Earth Science*, *8*, 18. doi: 10.3389/feart.2020.00018

Carmichael, I. S. E. (1991, January). The redox states of basic and silicic magmas: A reflection of their source regions? *Contributions to Mineralogy and Petrology*, *106*(2), 129–141. doi: 10.1007/BF00306429

Cashman, K. V., & Edmonds, M. (2019, February). Mafic glass compositions: A record of magma storage conditions, mixing and ascent. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *377*(2139), 20180004. doi: 10.1098/rsta.2018.0004

Cashman, K. V., Sparks, R. S. J., & Blundy, J. D. (2017). Vertically extensive and unstable magmatic systems: A unified view of igneous processes. *Science*, *355*(6331), eaag3055. doi: 10.1126/science.aag3055

Cassidy, M., Edmonds, M., Watt, S. F. L., Palmer, M. R., & Gernon, T. M. (2015, February). Origin of Basalts by Hybridization in Andesite-dominated Arcs. *Journal of Petrology*, *56*(2), 325–346. doi: 10.1093/petrology/egv002

Cembrano, J., & Lara, L. (2009, June). The link between volcanism and tectonics in the southern volcanic zone of the Chilean Andes: A review. *Tectonophysics*, *471*(1-2), 96–113. doi: 10.1016/j.tecto.2009.02.038

Chayes, F. (1971). *Ratio correlation: a manual for students of petrology and geochemistry.* University of Chicago Press.

Cheng, L., Costa, F., & Carniel, R. (2017, September). Unraveling the presence of multiple plagioclase populations and identification of representative two-dimensional sections using a statistical and numerical approach. *American Mineralogist*, *102*(9), 1894–1905. doi: 10.2138/am-2017-5929CCBYNCND

Chiasera, B., & Cortés, J. A. (2011). Predictive regions for geochemical compositional data of volcanic systems. *Journal of Volcanology and Geothermal Re-*

search, *207*(3-4), 83–92. doi: 10.1016/j.jvolgeores.2011.07.009

Clavero-Ribes, J. (1996). *Ignimbritas andesitico-basalticas postglaciales del volcán villarrica, andes del sur (39° 25's)* (Unpublished master's thesis). Universidad de Chile, Chile.

Cooper, K. M., & Kent, A. J. R. (2014). Rapid remobilization of magmatic crystals kept in cold storage. *Nature*, *506*(7489), 480–483. doi: 10.1038/nature12991

Corsaro, R. A., Falsaperla, S., & Langer, H. (2013). Geochemical pattern classification of recent volcanic products from Mt. Etna, Italy, based on Kohonen maps and fuzzy clustering. *International Journal of Earth Sciences*, *102*(4), 1151–1164. doi: 10.1007/s00531-012-0851-7

Cortés, J. A. (2009). On the Harker Variation Diagrams; A Comment on "The Statistical Analysis of Compositional Data. Where Are We and Where Should We Be Heading?" by Aitchison and Egozcue (2005). *Mathematical Geosciences*, *41*(7), 817–828. doi: 10.1007/s11004-009-9222-8

Cortés, J. A., Palma, J. L., & Wilson, M. (2007). Deciphering magma mixing: The application of cluster analysis to the mineral chemistry of crystal populations. *Journal of Volcanology and Geothermal Research*, *165*(3-4), 163–188. doi: 10.1016/j.jvolgeores.2007.05.018

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. doi: 10.1016/j.jvolgeores.2010.12.010

Delgado, F., Pritchard, M. E., Ebmeier, S., González, P., & Lara, L. (2017). Recent unrest (2002–2015) imaged by space geodesy at the highest risk Chilean volcanoes: Villarrica, Llaima, and Calbuco (Southern Andes). *Journal of Volcanology and Geothermal Research*, *344*, 270–288. doi: 10.1016/j.jvolgeores.2017.05.020

Droop, G. T. R. (1987). A general equation for estimating Fe3+ concentrations in ferromagnesian silicates and oxides from microprobe analyses, using stoichiometric criteria. *Mineralogical Magazine*, *51*(361), 431–435. doi: 10.1180/minmag.1987.051.361.10

Druitt, T. H., & Sparks, R. S. J. (1984, August). On the formation of calderas during ignimbrite eruptions. *Nature*, *310*(5979), 679–681. doi: 10.1038/310679a0

Ebmeier, S. K., Andrews, B. J., Araya, M. C., Arnold, D. W. D., Biggs, J., Cooper, C., ... Williamson, J. L. (2018, December). Synthesis of global satellite observations of magmatic and volcanic deformation: Implications for volcano monitoring & the lateral extent of magmatic domains. *Journal of Applied Volcanology*, *7*(1), 2. doi: 10.1186/s13617-018-0071-3

Edmonds, M., Cashman, K. V., Holness, M., & Jackson, M. (2019, February). Architecture and dynamics of magma reservoirs. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *377*(2139), 20180298. doi: 10.1098/rsta.2018.0298

Egozcue, J. J., Pawlowsky-Glahn, V., Mateu-Figueras, G., & Barcelo-Vidal, C. (2003). Isometric Logratio Transformations for Compositional Data Analysis. *Mathematical Geology*, 22.

Eichelberger, J. C., Izbekov, P. E., & Browne, B. L. (2006, March). Bulk chemical trends at arc volcanoes are not liquid lines of descent. *Lithos*, *87*(1-2), 135–154. doi: 10.1016/j.lithos.2005.05.006

Feig, S. T., Koepke, J., & Snow, J. E. (2010). Effect of oxygen fugacity and water on phase equilibria of a hydrous tholeiitic basalt. *Contributions to Mineralogy and Petrology*, *160*(4), 551–568. doi: 10.1007/s00410-010-0493-3

Filzmoser, P., & Hron, K. (2008). Outlier Detection for Compositional Data Using Robust Methods. *Mathematical Geosciences*, *40*(3), 233–248. doi: 10.1007/s11004-007-9141-5

Fry, J. M., Fry, T. R. L., & McLaren, K. R. (2000). Compositional data analysis and zeros in micro data. *Applied Economics*, *32*(8), 953–959. doi: 10.1080/000368400322002

Gavrilenko, M., Herzberg, C., Vidito, C., Carr, M. J., Tenner, T., & Ozerov, A. (2016). A Calcium-in-Olivine Geohygrometer and its Application to Subduction Zone Magmatism. *Journal of Petrology*, *57*(9), 1811–1832. doi: 10.1093/petrology/egw062

Ghiorso, M. S., & Gualda, G. A. R. (2015). An H2O–CO2 mixed fluid saturation model compatible with rhyolite-MELTS. *Contributions to Mineralogy and Petrology*, *169*(6), 53. doi: 10.1007/s00410-015-1141-8

Ginibre, C., & Wörner, G. (2007). Variable parent magmas and recharge regimes of the Parinacota magma system (N. Chile) revealed by Fe, Mg and Sr zoning in plagioclase. *Lithos*, *98*(1-4), 118–140. doi: 10.1016/j.lithos.2007.03.004

Gleeson, M. L. M., Gibson, S. A., & Stock, M. J. (2021, March). *Journal of Petrology*, *61*(11-12), egaa094. doi: 10.1093/petrology/egaa094

Greenacre, M. (2019). Variable Selection in Compositional Data Analysis Using Pairwise Logratios. *Mathematical Geosciences*, *51*(5), 649–682. doi: 10.1007/s11004-018-9754-x

Greenacre, M., & Lewi, P. (2009). Distributional Equivalence and Subcompositional Coherence in the Analysis of Compositional Data, Contingency Tables and Ratio-Scale Measurements. *Journal of Classification*, *26*(1), 29–54. doi: 10.1007/s00357-009-9027-y

Gualda, G. A. R., Ghiorso, M. S., Lemons, R. V., & Carley, T. L. (2012). Rhyolite-MELTS: A Modified Calibration of MELTS Optimized for Silica-rich, Fluid-bearing Magmatic Systems. *Journal of Petrology*, *53*(5), 875–890. doi: 10.1093/petrology/egr080

Hamada, M., Iwamori, H., Brandl, P. A., Ushikubo, T., Shimizu, K., Ito, M., ... Savov, I. P. (2020). Temporal Evolution of Proto-Izu–Bonin–Mariana Arc Volcanism over 10 Myr: Constraints from Statistical Analysis of Melt Inclusion Compositions. *Journal of Petrology*, *61*(1), egaa022. doi: 10.1093/petrology/egaa022

Henze, N., & Zirkler, B. (1990). A class of invariant consistent tests for multivariate normality. *Communications in statistics-Theory and Methods*, *19*(10), 3595–3617.

Hickey-Vargas, R., Roa, H. M., Escobar, L. L., & Frey, F. A. (1989). Geochemical variations in Andean basaltic and silicic lavas from the Villarrica-Lanin volcanic chain (39.5° S): An evaluation of source heterogeneity, fractional crystallization and crustal assimilation. *Contributions to Mineralogy and Petrology*, *103*(3), 361–386. doi: 10.1007/BF00402922

Hickey-Vargas, R., Sun, M., & Holbik, S. (2016). Geochemistry of basalts from small eruptive centers near Villarrica stratovolcano, Chile: Evidence for lithospheric mantle components in continental arc magmas. *Geochimica et Cosmochimica Acta*, *185*, 358–382. doi: 10.1016/j.gca.2016.03.033

Hildreth, W., & Moorbath, S. (1988, April). Crustal contributions to arc magmatism in the Andes of Central Chile. *Contributions to Mineralogy and Petrology*, *98*(4), 455–489. doi: 10.1007/BF00372365

Jellinek, A. M., Manga, M., & Saar, M. O. (2004, September). Did melting glaciers cause volcanic eruptions in eastern California? Probing the mechanics of dike formation: VOLCANISM AND DEGLACIATION. *Journal of Geophysical Research: Solid Earth*, *109*(B9), n/a-n/a. doi: 10.1029/2004JB002978

Jerram, D. A., & Martin, V. M. (2008). Understanding crystal populations and their significance through the magma plumbing system. *Geological Society, London, Special Publications*, *304*(1), 133–148. doi: 10.1144/SP304.7

Kahl, M., Chakraborty, S., Costa, F., & Pompilio, M. (2011, August). Dynamic plumbing system beneath volcanoes revealed by kinetic modeling, and the connection to monitoring data: An example from Mt. Etna. *Earth and Planetary Science Letters*, *308*(1-2), 11–22. doi: 10.1016/j.epsl.2011.05.008

Kamenetsky, V. S., Elburg, M., Arculus, R., & Thomas, R. (2006). Magmatic origin of low-Ca olivine in subduction-related magmas: Co-existence of contrasting magmas. *Chemical Geology*, *233*(3-4), 346–357. doi: 10.1016/j.chemgeo.2006.03.010

Kapinos, G., Montahaei, M., Meqbel, N., & Brasse, H. (2016). Three-dimensional electrical resistivity image of the South-Central Chilean subduction zone. *Tectonophysics*, *666*, 76–89. doi: 10.1016/j.tecto.2015.10.016

Kent, A. J. R., Darr, C., Koleszar, A. M., Salisbury, M. J., & Cooper, K. M. (2010). Preferential eruption of andesitic magmas through recharge filtering. *Nature Geoscience*, *3*(9), 631–636. doi: 10.1038/ngeo924

Kuritani, T., Yamaguchi, A., Fukumitsu, S., Nakagawa, M., Matsumoto, A., & Yokoyama, T. (2018). Magma Plumbing System at Izu-Oshima Volcano, Japan: Constraints From Petrological and Geochemical Analyses. *Frontiers in Earth Science*, *6*, 178. doi: 10.3389/feart.2018.00178

Lara, L. E., & Clavero, J. (2004). *Villarrica Volcano (39.5 S), Southern Andes, Chile* (Vol. Bulletin 61).

La Spina, G., Arzilli, F., Llewellin, E. W., Burton, M. R., Clarke, A. B., de' Michieli Vitturi, M., … Mader, H. M. (2021, January). Explosivity of basaltic lava fountains is controlled by magma rheology, ascent rate and outgassing. *Earth and Planetary Science Letters*, *553*, 116658. doi: 10.1016/j.epsl.2020.116658

Le Bas, M. J., Maitre, R. W. L., Streckeisen, A., Zanettin, B., & IUGS Subcommission on the Systematics of Igneous Rocks. (1986). A Chemical Classification of Volcanic Rocks Based on the Total Alkali-Silica Diagram. *Journal of Petrology*, *27*(3), 745–750. doi: 10.1093/petrology/27.3.745

Lerner, A. H., O'Hara, D., Karlstrom, L., Ebmeier, S. K., Anderson, K. R., & Hurwitz, S. (2020, July). The Prevalence and Significance of Offset Magma Reservoirs at Arc Volcanoes. *Geophysical Research Letters*, *47*(14). doi: 10.1029/2020GL087856

Lindsley, D. H., & Frost, B. R. (1992, October). Equilibria among Fe-Ti oxides, pyroxenes, olivine, and quartz: Part I. Theory. *American Mineralogist*, *77*(9-10), 987–1003.

Liu, E. J., Cashman, K. V., Miller, E., Moore, H., Edmonds, M., Kunz, B. E., ... Chigna, G. (2020). Petrologic monitoring at Volcán de Fuego, Guatemala. *Journal of Volcanology and Geothermal Research*, *405*, 107044. doi: 10.1016/j.jvolgeores.2020.107044

Liu, X., Wang, W., Pei, Y., & Yu, P. (2020). A knowledge-driven way to interpret the isometric log-ratio transformation and mixture distributions of geochemical data. *Journal of Geochemical Exploration*, *210*, 106417. doi: 10.1016/j.gexplo.2019.106417

Lohmar, S. (2008). *Petrologia de las Ignimbritas Licán y Pucón (Volcan Villarrica) y Curacautin (Volcan LLaima) en los Andes del sur de Chile* (Unpublished doctoral dissertation). Universidad de Chile, Chile.

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

Loucks, R. R. (1996, October). A precise olivine-augite Mg-Fe-exchange geothermometer. *Contributions to Mineralogy and Petrology*, *125*(2-3), 140–150. doi: 10.1007/s004100050211

Mahalanobis, P. C. (1936). On the generalized distance in statistics. *National Institute of Science of India*, *2*, 49–55.

Marsh, B. D. (2002, June). On bimodal differentiation by solidification front instability in basaltic magmas, part 1: Basic mechanics. *Geochimica et Cosmochimica Acta*, *66*(12), 2211–2229. doi: 10.1016/S0016-7037(02)00905-5

Martín-Fernández, J. A., Barceló-Vidal, C., & Pawlowsky-Glahn, V. (2000). Zero Replacement in Compositional Data Sets. In H.-H. Bock et al. (Eds.), *Data Analysis, Classification, and Related Methods* (pp. 155–160). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-59789-3_25

McCurry, M., Chadwick, J., Wright, K., Smith, R., & Ford, M. (2004). Preliminary la/icp-ms and epma examination of dacite enclaves and melt inclusions in phenocrysts from basaltic andesite pyroclasts from the pucón ignimbrite, volcán villarrica, southern andean volcanic zone: Implications for mafic ignimbrite volcanism..

McCurry, M., & Schmidt, K. (2001). Petrology and oxygen isotope geochemistry of the pucon ignimbrite-southern andean volcanic zone, chile: Implications for genesis of mafic ignimbrites..

McGee, L. E., Brahm, R., Rowe, M. C., Handley, H. K., Morgado, E., Lara, L. E., … Valdivia, P. (2017). A geochemical approach to distinguishing competing tectono-magmatic processes preserved in small eruptive centres. *Contributions to Mineralogy and Petrology*, *172*(6), 44. doi: 10.1007/s00410-017-1360-2

Mollo, S., Putirka, K., Iezzi, G., Del Gaudio, P., & Scarlato, P. (2011). Plagioclase–melt (dis)equilibrium due to cooling dynamics: Implications for thermometry, barometry and hygrometry. *Lithos*, *125*(1-2), 221–235. doi: 10.1016/j.lithos.2011.02.008

Moreno, H. (1993). Volcán villarrica, geología y evaluación del riesgo volcánico, regiones IX y x, 39 25 s. *Informe Final Proyecto FONDECYT*, *1247*, 1–112.

Moreno, H., & Clavero, J. (2006). Geología del volcán Villarrica, Regiones de La Araucanía y de Los Lagos. Carta Geológica de Chile, Serie Geología Básica, No. 98. Servicio Nacional de Geología y Minería. Mapa Escala 1:50,000. Santiago, Chile.

Moreno, H., Clavero, J., & Lara, L. (1994). Actividad Explosiva Postglacial del Volcan Villarrica, Andes del Sur (39o25'S). In *7th Congreso Geológico Chileno* (Vol. 1, pp. 329–333). Universidad de Concepcion.

Morgado, E., Parada, M., Contreras, C., Castruccio, A., Gutiérrez, F., & McGee, L. (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. doi: 10.1016/j.jvolgeores.2015.09.023

Mutch, E. J. F., Maclennan, J., Shorttle, O., Edmonds, M., & Rudge, J. F. (2019). Rapid transcrustal magma movement under Iceland. *Nature Geoscience*, *12*(7), 569–574. doi: 10.1038/s41561-019-0376-9

Panjasawatwong, Y., Danyushevsky, L. V., Crawford, A. J., & Harris, K. L. (1995). An experimental study of the effects of melt composition on plagioclase-melt equilibria at 5 and 10 kbar: Implications for the origin of magmatic high-An plagioclase. *Contributions to Mineralogy and Petrology*, *118*(4), 420–432. doi: 10.1007/s004100050024

Pavez, M., Schill, E., Held, S., Díaz, D., & Kohl, T. (2020, August). Visualizing preferential magmatic and geothermal fluid pathways via electric conductivity at Villarrica Volcano, S-Chile. *Journal of Volcanology and Geothermal Research*, *400*, 106913. doi: 10.1016/j.jvolgeores.2020.106913

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., ... Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830.

Petit-Breuilh, M. (2004). *La historia eruptiva de los volcanes hispanoamericanos (siglos xvi al xx).* Servicio de Publicaciones Exmo. Cabildo Insular de Lanzarote. Huelva, Spain.

Pichavant, M., Costa, F., Burgisser, A., Scaillet, B., Martel, C., & Poussineau, S. (2007, September). Equilibration Scales in Silicic to Intermediate Magmas Implications for Experimental Studies. *Journal of Petrology*, *48*(10), 1955–1972. doi: 10.1093/petrology/egm045

Pioli, L., Scalisi, L., Costantini, L., Di Muro, A., Bonadonna, C., & Clavero, J. (2015). Explosive style, magma degassing and evolution in the Chaimilla eruption, Villarrica volcano, Southern Andes. *Bulletin of Volcanology*, *77*(11), 93. doi: 10.1007/s00445-015-0976-1

Pistone, M., Blundy, J., & Brooker, R. A. (2017, April). Water transfer during magma mixing events: Insights into crystal mush rejuvenation and melt extraction processes. *American Mineralogist*, *102*(4), 766–776. doi: 10.2138/am-2017-5793

Pizarro, C., Parada, M. A., Contreras, C., & Morgado, E. (2019). Cryptic magma recharge associated with the most voluminous 20th century eruptions (1921, 1948 and 1971) at Villarrica Volcano. *Journal of Volcanology and Geothermal Research*, *384*, 48–63. doi: 10.1016/j.jvolgeores.2019.07.001

Rawson, H., Pyle, D. M., Mather, T. A., Smith, V. C., Fontijn, K., Lachowycz, S. M., & Naranjo, J. A. (2016, April). The magmatic and eruptive response of arc volcanoes to deglaciation: Insights from southern Chile. *Geology*, *44*(4), 251–254. doi: 10.1130/G37504.1

Reubi, O., & Blundy, J. (2009). A dearth of intermediate melts at subduction zone volcanoes and the petrogenesis of arc andesites. *Nature*, *461*(7268), 1269–1273. doi: 10.1038/nature08510

Robidoux, P., Pastén, D., Levresse, G., Diaz, G., & Paredes, D. (2021, July). Volatile Content Implications of Increasing Explosivity of the Strombolian Eruptive Style along the Fracture Opening on the NE Villarrica Flank: Minor Eruptive Centers in the Los Nevados Group 2. *Geosciences*, *11*(8). doi: 10.3390/geosciences11080309

Romero, J. E., Vera, F., Polacci, M., Morgavi, D., Arzilli, F., Alam, M. A., ... Keller, W. (2018). Tephra From the 3 March 2015 Sustained Column Related to Explosive Lava Fountain Activity at Volcán Villarrica (Chile). *Frontiers in Earth Science*, *6*, 98. doi: 10.3389/feart.2018.00098

Rousseeuw, P. J., & Driessen, K. V. (1999). A Fast Algorithm for the Minimum Covariance Determinant Estimator. *Technometrics*, *41*(3), 212–223. doi: 10.1080/00401706.1999.10485670

Rousseeuw, P. J., & Zomeren, B. C. V. (1990). Unmasking Multivariate Outliers and Leverage Points. *Journal of The American Statistical Association*, *85*(411), 633–639.

Ruprecht, P., Bergantz, G. W., Cooper, K. M., & Hildreth, W. (2012). The crustal magma storage system of volcán quizapu, chile, and the effects of magma mixing on magma diversity. *Journal of Petrology*, *53*(4), 801–840.

Sato, M. (1978, June). Oxygen fugacity of basaltic magmas and the role of gas-forming elements. *Geophysical Research Letters*, *5*(6), 447–449. doi: 10.1029/

GL005i006p00447

SERNAGEOMIN. (2019). *Ranking de riesgo de volcanes activos de Chile 2019* (Tech. Rep.). SERNAGEOMIN. Retrieved from https://www.sernageomin.cl/geologia

Siebert, L., Cottrell, E., Venzke, E., & Andrews, B. (2015, January). Chapter 12 - Earth's Volcanoes and Their Eruptions: An Overview. In H. Sigurdsson (Ed.), *The Encyclopedia of Volcanoes (Second Edition)* (pp. 239–255). Amsterdam: Academic Press. doi: 10.1016/B978-0-12-385938-9.00012-2

Silva Parejas, C. (2008). *Evolution and Dynamics of the 3.6 Ka BP Pucón Eruption of Villarrica Volcano, Chile* (Unpublished doctoral dissertation). Universidad de Chile, Santiago, Chile.

Silva Parejas, C., Druitt, T. H., Robin, C., Moreno, H., & Naranjo, J.-A. (2010). The Holocene Pucón eruption of Volcán Villarrica, Chile: Deposit architecture and eruption chronology. *Bulletin of Volcanology*, *72*(6), 677–692. doi: 10.1007/s00445-010-0348-9

Silverman, B. W. (1986). *Density estimation for statistics and data analysis* (No. 26). Boca Raton: Chapman & Hall/CRC.

Singer, B. S., Dungan, M. A., & Layne, G. D. (1995). Textures and Sr, Ba, Mg, Fe, K and Ti compositional profiles in volcanic plagioclase clues to the dynamics of calc-alkaline magma chambers. *American Mineralogist*, *80*(7-8), 776–798. doi: 10.2138/am-1995-7-815

Smith, P. M., & Asimow, P. D. (2005). Adiabat_1ph: A new public front-end to the MELTS, pMELTS, and pHMELTS models: ADIABAT_1PH FRONT-END. *Geochemistry, Geophysics, Geosystems*, *6*(2). doi: 10.1029/2004GC000816

Sparks, R., & Marshall, L. (1986, September). Thermal and mechanical constraints on mixing between mafic and silicic magmas. *Journal of Volcanology and Geothermal Research*, *29*(1-4), 99–124. doi: 10.1016/0377-0273(86)90041-7

Stern, C. R. (2004, December). Active Andean volcanism: Its geologic and tectonic setting. *Revista geológica de Chile*, *31*(2). doi: 10.4067/S0716-02082004000200001

Streck, M. J. (2008, January). Mineral Textures and Zoning as Evidence for Open System Processes. *Reviews in Mineralogy and Geochemistry*, *69*(1), 595–622. doi: 10.2138/rmg.2008.69.15

Takagi, D., Sato, H., & Nakagawa, M. (2005). Experimental study of a low-alkali tholeiite at 1–5 kbar: Optimal condition for the crystallization of high-An plagioclase in hydrous arc tholeiite. *Contributions to Mineralogy and Petrology*, *149*(5), 527–540. doi: 10.1007/s00410-005-0666-7

Templ, M., Filzmoser, P., & Reimann, C. (2008). Cluster analysis applied to regional geochemical data: Problems and possibilities. *Applied Geochemistry*, *23*(8), 2198–2213. doi: 10.1016/j.apgeochem.2008.03.004

Van der Maaten, L., & Hinton, G. (2008). Visualizing data using t-SNE. *Journal of machine learning research*, *9*(11).

Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. *Journal of The American Statistical Association*, *58*(301), 236–244.

Waters, L. E., & Lange, R. A. (2015). An updated calibration of the plagioclase-liquid hygrometer-thermometer applicable to basalts through rhyolites. *American Mineralogist*, *100*(10), 2172–2184. doi: 10.2138/am-2015-5232

Watt, S. F., Pyle, D. M., & Mather, T. A. (2013, July). The volcanic response to deglaciation: Evidence from glaciated arcs and a reassessment of global eruption records. *Earth-Science Reviews*, *122*, 77–102. doi: 10.1016/j.earscirev.2013.03.007

Wehrmann, H., Hoernle, K., Jacques, G., Garbe-Schönberg, D., Schumann, K., Mahlke, J., & Lara, L. E. (2014). Volatile (sulphur and chlorine), major, and trace element geochemistry of mafic to intermediate tephras from the Chilean Southern Volcanic Zone (33–43°S). *International Journal of Earth Sciences*, *103*(7), 1945–1962. doi: 10.1007/s00531-014-1006-9

Williams, M., Schoneveld, L., Mao, Y., Klump, J., Gosses, J., Dalton, H., ... Barnes, S. (2020). Pyrolite: Python for geochemistry. *Journal of Open Source Software*, *5*(50), 2314. doi: 10.21105/joss.02314

Williams, W. T., & Lambert, J. M. (1966). Multivariate Methods in Plant Ecology: V. Similarity Analyses and Information-Analysis. *The Journal of Ecology*, *54*(2), 427. doi: 10.2307/2257960

Wilson, A., & Russell, J. (2020, October). Glacial pumping of a magma-charged lithosphere: A model for glaciovolcanic causality in magmatic arcs. *Earth and Planetary Science Letters*, *548*, 116500. doi: 10.1016/j.epsl.2020.116500

Wishart, D. (1969). Numerical Classification Method for deriving Natural Classes. *Nature*, *221*(5175), 97–98. doi: 10.1038/221097a0

Witter, J. B., Kress, V. C., Delmelle, P., & Stix, J. (2004). Volatile degassing, petrology, and magma dynamics of the Villarrica Lava Lake, Southern Chile. *Journal of Volcanology and Geothermal Research*, *134*(4), 303–337. doi: 10.1016/j.jvolgeores.2004.03.002

Zajacz, Z., & Halter, W. (2007). LA-ICPMS analyses of silicate melt inclusions in co-precipitated minerals: Quantification, data analysis and mineral/melt partitioning. *Geochimica et Cosmochimica Acta*, *71*(4), 1021–1040. doi: 10.1016/j.gca.2006.11.001

Zajacz, Z., & Halter, W. (2009). Copper transport by high temperature, sulfur-rich magmatic vapor: Evidence from silicate melt and vapor inclusions in a basaltic andesite from the Villarrica volcano (Chile). *Earth and Planetary Science Letters*, *282*(1-4), 115–121. doi: 10.1016/j.epsl.2009.03.006

Zellmer, G. F., Pistone, M., Iizuka, Y., Andrews, B. J., Gómez-Tuena, A., Straub, S. M., & Cottrell, E. (2016, November). Petrogenesis of antecryst-bearing arc basalts from the Trans-Mexican Volcanic Belt: Insights into along-arc variations in magma-mush ponding depths, H_2O contents, and surface heat flux. *American Mineralogist*, *101*(11), 2405–2422. doi: 10.2138/am-2016-5701