- -

The magmatic evolution of South-East Crater (Mt. Etna) during the February–April 2021 sequence of lava fountains from a mineral chemistry perspective

- -

Musu, Alessandro; Corsaro, Rosa Anna; Higgins, Oliver; Jorgenson, Corin Christen; Petrelli, Maurizio; Caricchi, Luca

RESEARCH ARTICLE

The magmatic evolution of South-East Crater (Mt. Etna) during the February–April 2021 sequence of lava fountains from a mineral chemistry perspective

Alessandro Musu[1] · Rosa Anna Corsaro[2] · Oliver Higgins[3] · Corin Jorgenson[1] · Maurizio Petrelli[4] · Luca Caricchi[1]

Abstract

The South-East Crater (SEC) at Mt. Etna started a period of lava fountaining in December 2020, producing over 60 paroxysms until February 2022. The activity had an intense sequence from February 16 to April 1, 2021, totaling 17 paroxysmal events separated by repose times varying from 1 to 7 days. The eruptive sequence was extensively monitored, providing a unique opportunity to relate the chemistry and texture of the erupted products to eruption dynamics. We investigate the temporal evolution of the magmatic system through this eruptive sequence by quantifying variations in the composition and texture of clinopyroxene. Clinopyroxene major element transects across crystals from five representative lava fountains allow us to determine the relative proportions of deep versus shallow-stored magmas that fed these events. We use hierarchical clustering (HC), an unsupervised machine learning technique, to objectively identify clinopyroxene compositional clusters and their variations during this intense eruptive phase. Our results show that variations of monitoring parameters and eruption intensity are expressed in the mineral record both as changes in cluster proportions and the chemical complexity of single crystals. We also apply random forest thermobarometry to relate each cluster to P-T conditions of formation. We suggest that the February–April 2021 eruptive sequence was sustained by the injection of a hotter and deeper magma into a storage area at 1–3 kbar, where it mixed with a slightly more evolved magma. The February 28 episode emitted the most mafic magma, in association with the highest mean lava fountain height and highest time–averaged discharge rate, which make it the peak of the analyzed eruptive interval. Our results show that after this episode, the deep magma supply decreased and the erupted magma become gradually more chemically evolved, with a lower time–average discharge rate and fountain height. We propose this approach as a means to rapidly, objectively, and effectively link petrological and geophysical/geochemical monitoring during ongoing eruptions. We anticipate that the systematic application of this approach will serve to shed light on the magmatic processes controlling the evolution of ongoing eruptions.

Keywords Clinopyroxene · Hierarchical clustering · Mineral stratigraphy · Thermobarometry · Random forest machine learning

Introduction

Trans-crustal magmatic processes and chemico-physical variations within volcanic plumbing systems profoundly influence the style and size of eruptions (Bachmann and Bergantz

✉ Alessandro Musu
    alessandro.musu@unige.ch

1   Department of Earth Sciences, University of Geneva, Rue des Maraîchers 13, CH-1205 Geneva, Switzerland

2   Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Etneo-Sezione di Catania, Piazza Roma 2, 95123 Catania, Italy

3   Geology, School of Natural Sciences, Trinity College Dublin, Dublin, Ireland

4   Department of Physics and Geology, University of Perugia, Piazza dell'Università, 1, 06123 Perugia, Italy

2008; Cassidy et al. 2018). Understanding these parameters and their temporal evolution represents a key step in improving our understanding of volcanic activity and the ability to interpret pre-eruptive monitoring signals (Alparone et al. 2003; Pichavant et al. 2009; Andronico and Corsaro 2011). Mineral chemistry represents an essential archive recording pre and syn-eruptive processes, as minerals record their growth conditions by changing chemistry (Ginibre et al. 2007; Blundy and Cashman 2008; Costa 2008; Ruprecht et al. 2012; Cashman and Blundy 2013; Mollo et al. 2013; Barboni and Schoene 2014; Zellmer et al. 2016; Petrone et al. 2016; Cheng et al. 2017; Morgavi et al. 2017; Probst et al. 2018; Ubide et al. 2019; Caricchi et al. 2020; Weber et al. 2020; Sheldrake and Higgins 2021; Higgins et al. 2021; Boschetty et al. 2022). However, to quantitatively link monitoring parameters with processes occurring at depth, it is necessary to objectively identify chemically distinct zones in minerals. Such objective classification also allows for the intercomparison between eruptions, which is otherwise complicated by a degree of subjectivity of analyses based exclusively on human judgment. Without further delay, we state here that this approach is in no way a substitute for petrographic analysis and expert judgment. Machine learning is just a tool to enhance our interpretative capacity.

We focus on clinopyroxene because it is commonly found in mafic to intermediate magmas and crystallizes over a wide range of temperatures (T) and pressures (P; Hirschmann et al. 2008; Putirka 2008; Petrone et al. 2016; Ubide et al. 2019). Clinopyroxene chemistry is also sensitive to changes in magma composition, water content, pressure, and temperature (Putirka 2008; Mollo et al. 2018; Ubide et al. 2019). Secondly, the overall low diffusion rate of chemical constituents in its crystal lattice effectively preserves a record of the growth conditions (Müller et al. 2013; Petrone et al. 2016; Ubide and Kamber 2018; Ubide et al. 2019). Indeed, it has been demonstrated how Fe and Mg diffuse slower in clinopyroxene compared with other mafic minerals (Müller et al. 2013, Ubide and Kamber 2018). Additionally, of the elements in a clinopyroxene Fe and Mg are some of the fastest diffusing elements, meaning that other elements have even lower diffusivities (Müller et al. 2013; Ubide and Kamber 2018; Lierenfeld et al. 2019). These properties make clinopyroxene not only an excellent tracer of the chemical evolution of magmatic reservoirs (Winpenny and Maclennan 2011; Ubide and Kamber 2018; Caricchi et al. 2020; Boschetty et al. 2022), but also a valuable thermobarometer (Putirka et al. 1996, Nazzareni et al. 2001; Putirka 2008; Mollo et al. 2015; Petrelli et al. 2020; Higgins et al. 2022; Nazzareni et al. 2022; Jorgenson et al. 2022).

Mt. Etna volcano is intensively monitored by the Istituto Nazionale di Geofisica e Vulcanologia-Osservatorio Etneo (INGV-OE) via the collection of geochemical, volcanological, geophysical, and satellite monitoring data, together with a regular sampling of all eruptive events (Corsaro and Miraglia 2022). Thus, the paroxysmal sequence of the South-East summit crater of Etna from February to April 2021 provides an attractive opportunity to link petrology, monitoring signals, and eruptive parameters (such as, the repose time before the eruptive event, the mean fountain height, the erupted volume during the fountaining activity, the time-averaged discharge rate and the cumulative reduced displacement). To this aim, we analyzed the lava fountain episodes of February 16, 19, and 28 and March 2 and 10, when the eruptive frequency was relatively high (one episode every 1–3 days). Additionally, the presence of abundant literature data permits the comparison with past eruptions (Behncke and Neri 2003; Di Renzo et al. 2019). In general, continuous monitoring and existing data make Mt. Etna a uniquely suitable candidate for studying the temporal evolution of plumbing system processes and their influence on the eruptive behavior.

Geological setting

Mt. Etna is located in eastern Sicily (Italy; Fig.1a). Its >3300 m elevation and 1250 km^2 areal extent make it the largest active European volcano (Branca et al. 2004). It is also one of the most active volcanoes in the world (Cappello et al. 2013; Corsaro and Miraglia 2022). Mt. Etna volcano presents a wide range of eruptive behaviors, from purely effusive to more explosive eruptions, including strong strombolian and violent lava-fountaining events (Branca and Del Carlo 2004; Ferlito et al. 2014; Corsaro and Miraglia 2022). It produces eruptions from fissure vents along its flanks and its summit craters (Branca and Del Carlo 2004; Di Renzo et al. 2019). For both types of eruptions, a significant increase in volcanic activity has been observed over the last decades (Behncke and Neri 2003; Branca and Del Carlo 2005), in particular since 1971, with the formation of the South-East Crater (Cappello et al. 2013). The summit area of the volcano is composed of four active vents (Fig. 1a): Voragine (VOR), Bocca Nuova (BN), North-East Crater (NEC), and South-East Crater (SEC). Among these, SEC represents the youngest but also the most active crater (Andronico and Corsaro 2011; Di Renzo et al. 2019; Corsaro and Miraglia 2022). SEC has often produced cyclical or episodic eruptive activity (Parfitt and Wilson 1994; Spina et al. 2019), characterized by a series of paroxysmal lava fountaining events with very short repose times (Andronico and Corsaro 2011; Spina et al. 2019; Corsaro and Miraglia 2022). This is exemplified by 23 lava fountain episodes in 1998 (Neri et al. 2011), the 64 paroxysms in 2000 (Andronico and Corsaro 2011), and the 44 fountaining events between January 2011 and December 2013 (Calvari et al. 2018; Bonaccorso et al. 2021).
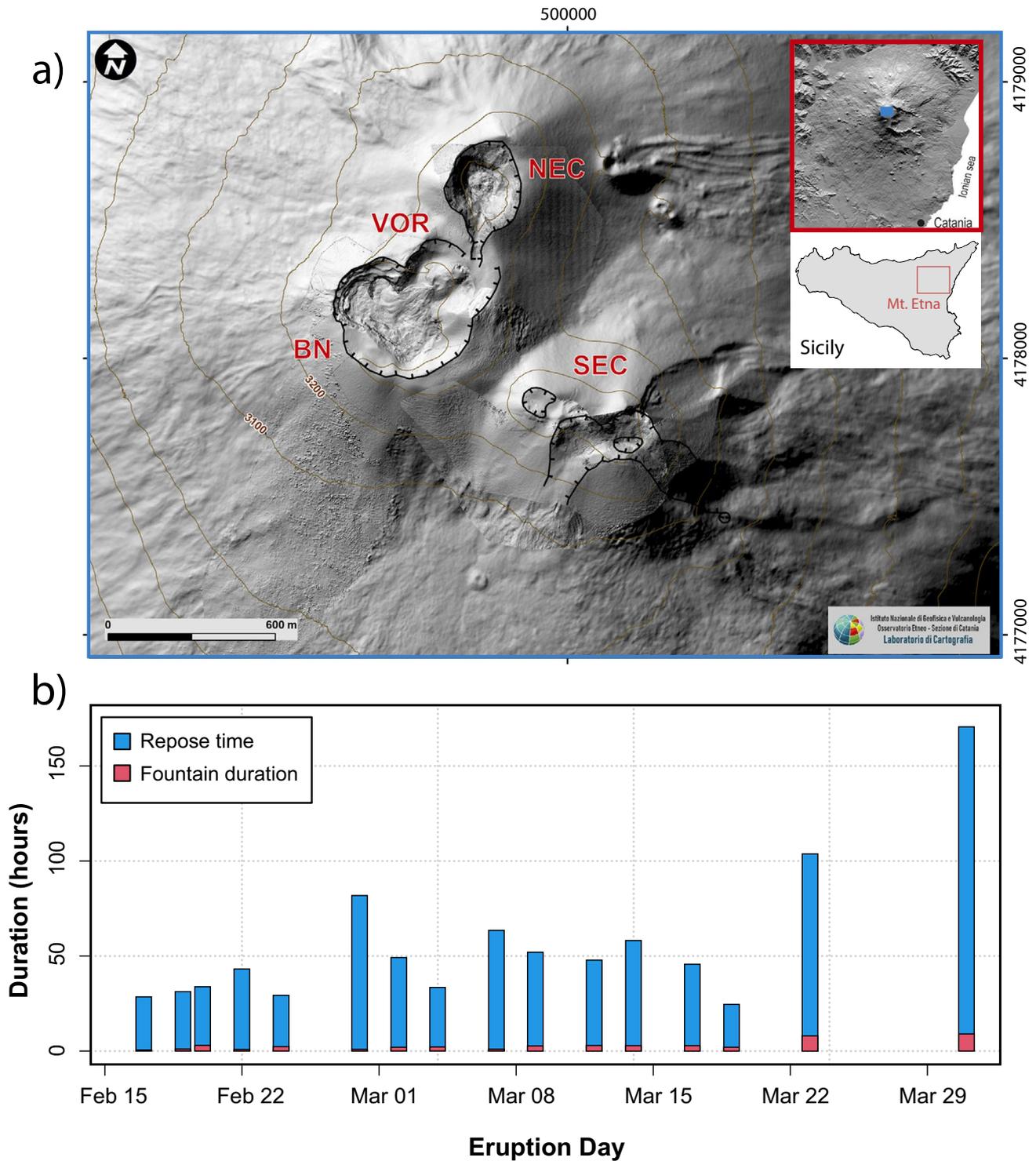
Fig. 1 **a** Digital elevation model (DEM) of the Mt. Etna summit area, image modified from Corsaro and Miraglia, 2022. The red rectangle indicates the position of Mt. Etna in Sicily, the blue rectangle the location of the summit area on Etna; **b** bar plot of the paroxysmal events occurred at SEC between February 16 and April 1. Blue represents the repose time before the eruptive event and red represents the duration of the lava fountain activity (Calvari and Nunnari 2022)

A new cyclical eruptive phase started at the SEC on December 13, 2020, producing over 60 paroxysms up to February 21, 2022 (Andronico et al. 2021; Bonaccorso et al. 2021; Marchese et al. 2021; Calvari and Nunnari 2022). The activity went through a relatively intense sequence from February 16 to April 1, 2021, with a paroxysm occurring every 1.85 days (± 0.67 days) until March 19 and longer repose times for the last two events (March 23–24 and March 31–April 1, Fig. 1b) for a total of 17 paroxysmal events (De Gori et al. 2021; Marchese et al. 2021; Calvari and Nunnari 2022; Corsaro and Miraglia 2022). Each paroxysmal event followed the typical pattern observed at the SEC in recent decades (Alparone et al. 2003; Andronico and Corsaro 2011; Calvari et al. 2011, 2021; Calvari and Nunnari 2022; Corsaro and Miraglia 2022), with an intensification of strombolian activity which has culminated in fountaining event (Calvari and Nunnari 2022; Corsaro and Miraglia 2022). During the fountaining activity, a sustained eruption column was generated, which could extend to several kilometers in height (Andronico et al. 2021; Calvari et al. 2021; Calvari and Nunnari 2022). As shown in Fig. 1, the sustained fountaining events between February and April 2021 lasted from about 51 min to a maximum of ~ 13 h, with a median duration of about 2.7 h (lava fountain duration data from Calvari and Nunnari 2022; Fig. 1b).

Materials and methods

Small lapilli (maximum long-axis length 2 cm) were sampled from the tephra fallout deposits of the 16, 19, and 28 February and 2 and 10 March lava fountains. From each eruption, randomly selected lapilli were prepared for chemical analysis.

Electron probe micro analyzer (EPMA)

Major element analyses of clinopyroxene crystals and glass were collected using a JEOL 8200 Superprobe at the University of Geneva and a JEOL JXA-8530F at the University of Lausanne. Clinopyroxene analyses were collected with a focused beam (diameter between 1 and 2 µm) at an accelerating voltage of 15 keV and a beam current of 20 nA. The glass analyses were conducted with either a 10-µm-diameter beam (15 keV, 6 nA) or a 5 µm diameter (15 keV and 4 nA) depending on glass pool size.

Calibration for clinopyroxene quantitative analyses was performed using the following standards: forsterite for Mg, fayalite for Fe, Mn-Ti oxide for Mn and Ti, albite for Na, wollastonite for Si and Ca, orthoclase for Al, NiO for Ni, and chromium oxide for Cr. Chemical analyses were collected along rim-to-core transects with a point spacing of ~ 2 µm for a total of 1250 analyses. Additional 29 analyses were collected for glass at various locations within the groundmass of each sample.

Analyses of clinopyroxene were filtered by removing all data with a total oxide content below 98% and above 101% and cations per formula unit higher than 4.04 and lower than 3.98 (recalculated assuming all iron in clinopyroxene to be FeO). Chemical analyses of clinopyroxenes and glass are available in the supplementary materials table (Supplementary Tables 1 and 3).

Data transformation

A geochemical dataset (like the EPMA data used in this study) can be defined as a "closed" dataset. A dataset is considered "closed" when the variables it contains, in our case the oxide concentration, are not independent but are related in some way to each other. In this context, the oxides are expressed as wt%, so their nominal sum is 100% (Templ et al. 2008). Conducting statistical analysis directly on "closed datasets" can lead to issues (Butler 1976; Aitchison 1986; Templ et al. 2008) as some statistical approaches require the data to be normally distributed and not constrained to a constant total value (Boschetty et al. 2022). Various data transformations have been proposed to solve this issue such as the additive log-ratio transformation, the centered log-ratio transformation (Aitchison 1986), and the isometric log-ratio (ilr) transformation (Egozcue et al. 2003). Ilr has been shown to work well with geochemical data (Aitchison and Egozcue 2005; Templ et al. 2008; Filzmoser et al. 2009; Carranza 2011; Reimann et al. 2011; Boschetty et al. 2022). Ilr is a data transformation method that can be easily applied to a multivariate environment and has the particular property of representing the variables of our closed system in a real Euclidean space (Templ et al. 2008; Filzmoser et al. 2009; Carranza 2011). This method is useful to transform skewed data distributions into distributions closer to normal (Templ et al. 2008; Filzmoser et al. 2009). This property is useful for improving clustering results as shown by Boschetty et al. (2022). We apply the ilr transformation to our filtered data, using the formula proposed by (Egozcue et al. 2003), where the i components, transformed with the ilr function (y=ilr(x)), have been calculated as follows:

$$y_i = \sqrt{\frac{i}{i+1}} \ln \left[\frac{g(x_1, \ldots, x_i)}{x_{i+1}} \right], i = 1, 2, \ldots, D-1. \quad (1)$$

In Eq. 1, D are the parts of the compositional analysis (the number of analyzed elements for each spot analysis, x), i is one particular part of D (one of the ith analyzed elements, e.g., the wt% of SiO_2) and g(x_i) is the geometric mean of the ith parts of x, which in our case corresponds to the geometric

mean of the analyzed elements, for example, from $x_1 = SiO_2$ to $X_i = CaO$ (Egozcue et al. 2003).

Working with the ilr transformation requires non-zero values in the starting dataset (Cortés et al. 2007; Templ et al. 2008; Carranza 2011; Boschetty et al. 2022). The ilr transformation has been performed using the library "compositions" included in the open-source software R (R Core Team 2021). Zeros can occur in the investigated dataset for different reasons (Pawlowsky-Glahn and Olea 2004; Cortés et al. 2007; Boschetty et al. 2022). In the case of below-detection-limit zeros, replacement methods are available (Fry et al. 2000; Martín-Fernández et al. 2000; Boschetty et al. 2022), but from a general perspective, if zeros represent more than half of the studied dataset, the element in question should not be used for cluster analysis (Boschetty et al. 2022). In our case, most of the Cr_2O_3 analyses are below the detection limit; thus, this oxide was not considered for cluster analysis.

Data normalization

An additional issue with geochemical data is the different ranges and absolute values over which the different oxides vary. Variability and absolute abundance may result in the least concentrated oxides overwhelmingly controlling cluster analysis (Templ et al. 2008; Caricchi et al. 2020; Boschetty et al. 2022). To circumvent this issue, the data were normalized with the median-mad method (Templ et al. 2008; Rousseeuw and Hubert 2011; Eesa and Arabo 2017). This method transforms the dataset by subtracting the median and dividing by the median absolute deviation (MAD; Huber 2004; Rousseeuw and Hubert 2011). The MAD is the median of the absolute deviations from the median and can be calculated as follows:

$$MAD = median\left\{\left|x_i - M_n\right|\right\}, \;\; with \; M_n = median\left\{x_i\right\} \quad (2)$$

Calculating MAD using Eq. (2) the final normalization formula will be:

$$X' = \frac{x_i - M_n}{MAD} \quad (3)$$

where X' is the normalized dataset.

The median-mad normalization can be easily applied to the transformed dataset using the "stats" package in R (R Core Team 2021).

Hierarchical clustering

Clustering methods are useful to subdivide multivariate observations into representative and homogeneous groups (Templ et al. 2008). Hierarchical clustering (HC) is an unsupervised learning method to identify clusters of similar chemical composition within and between crystals, by assigning each analysis to a cluster. This approach can

recognize similar chemical sequences from the core to the rim of disparate crystals and help identify their chemico-physical growth conditions (Caricchi et al. 2020). This method has been already successfully applied to address geochemical and petrological problems (Templ et al. 2008; Caricchi et al. 2020; Boschetty et al. 2022) and can be implemented using the open-source software R, using the library "cluster" (R Core Team 2021). In detail, the standard implementation of the HC uses the Euclidean distance (d_{ij}) of the filtered, transformed, and normalized dataset to estimate the similarity (or dissimilarity) among different analyses in the Euclidean space:

$$d_{ij} = \sqrt{\left(x_i - x_j\right)^2 + \left(y_i - y_j\right)^2 + \cdots + \left(z_i - z_j\right)^2}, \quad (4)$$

where the variables from x to z represent the ilr components and i and j are the indexes of two different analyses. To cluster similar occurrences, different linkage criteria can be used. One of the most widely used criteria is the Ward minimum variance (Ward 1963). This method starts considering each analysis as an individual cluster. Then, it pairs similar occurrences into clusters, considering the smallest values of d_{ij} between them. It continues by progressively joining occurrences and incorporating within the clusters only the points which minimize average variance within clusters (Caricchi et al. 2020). Results can be visualized in a dendrogram. In the dendrogram, the branch length corresponds to the d_{ij} distance necessary to merge two distinct clusters into one.

Outlier detection

In natural datasets, outliers (i.e., observations which deviate from the majority) are commonplace. In our case, this could result from the mixed analysis of two zones belonging to different compositional clusters (Rousseeuw and Hubert 2011). The identification of outliers is based on the difference (or distance) between a specific analysis and a cluster of analyses. The Mahalanobis distance provides such a measure (Mahalanobis 1936; Boschetty et al. 2022) and was used in this study to identify and remove outliers. The Mahalanobis distance for each observation is calculated as:

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

where $\vec{x}$ represent a matrix of observation in a defined cluster, C is the center location estimator of the distribution, and $\vec{\mu}$ the covariance estimator. Following the procedure reported by Boschetty et al. (2022), we marked as outliers all observations that lie outside of the 97.5 percentile range of the distribution.

Cluster visualization

To visualize and validate the results of the cluster analysis, we apply principal component analysis (PCA), a dimensional reduction method (Hotelling 1933; Abdi and Williams 2010; Jolliffe and Cadima 2016; Boschetty et al. 2022), on the ilr-transformed dataset. PCA is a dimensional reduction method which allows for the visualization and interpretation of complex multivariate dataset while minimizing the loss of information (Hotelling 1933; Abdi and Williams 2010; Jolliffe and Cadima 2016; Boschetty et al. 2022). PCA allows us to visualize our multidimensional dataset (e.g., the chemical composition of the crystals, their oxide content, or in our case the six ILR-transformed vectors) in a smaller number of components representative of the multidimensional variance of our dataset (Hsieh 2009; Unglert et al. 2016; Bisciotti et al. 2022). In our case PCA1 and PCA2 explain alone up to the 92% of the total chemical variance (loadings, scores, and contributions of PCA analysis are reported in the Supplementary Table 4). If the clustering correctly identifies compositional clusters, they should be distinguishable when plotted in PCA1-PCA2 space, which is the case as shown in Fig. 2a. Additionally, we test the validity of the cluster analysis by visually inspecting if specific clusters correspond to textural features within crystals (Fig. 2b).

## Identifying the representative number of clusters

Cluster analysis is an unsupervised learning method and there is no a priori correct number of clusters. Therefore, different strategies must be adopted to identify the relevant number of clusters for the Etna dataset. Several quantitative clustering validation techniques are available (Halkidi et al. 2001; Templ et al. 2008; Charrad et al. 2014; Sheldrake and Higgins 2021), although when dealing with geochemical datasets, they are insufficient to robustly determine the exact number of clusters (Templ et al. 2008, Supplementary document 1). The result of multivariate cluster analysis is a distance matrix in Euclidean space (Templ et al. 2008; Caricchi et al. 2020; Sheldrake and Higgins 2021; Boschetty et al. 2022). A good method to visualize the distance matrix is represented by a dendrogram, where the *y*-axis represents the distances between observations (Temple et al., 2008; Caricchi et al., 2020; Sheldrake and Higgins, 2021; Boschetty et al., 2022). A qualitative examination of the dendrogram can inform about the number of clusters within a dataset, whereby the dendrogram is cut at different heights to separate the dataset into different numbers of clusters (Fig. 2c, d). Templ et al. (2008) demonstrate that the number of clusters in a geochemical dataset should be assessed with respect to the known properties of the object of investigation. In this regard, we generated a series of results in an iterative fashion, increasing the number of clusters from two to eight. For each configuration, we first tested the chemical validity of the clustering by visualizing the data in PCA1-PCA2 space through color coding the points according to the relative clusters, allowing us to verify that each cluster was chemically distinct (Fig. 2a, Hotelling 1933; Jolliffe 2002; Abdi and Williams 2010; Jolliffe and Cadima 2016). Secondly, we ensured that each cluster corresponded to a texturally defined area of the crystal. A visual textural analysis was conducted by overlaying the analyzed points on the BSE images, assigning each point the color of the relevant cluster (Fig. 2b, Fig. 3a), and checking that each cluster corresponded to a texturally distinguishable area that was large enough not to be the result of mixing between two chemical zones. The match between the zoning pattern and clusters for selected crystals was an additional confirmation that the clustering identified geochemically meaningful groups of analyses (Fig. 3a–d).

Random forest thermobarometry

We applied random forest (RF) thermobarometry (Petrelli et al. 2020; Higgins et al. 2022; Jorgenson et al. 2022) to the clinopyroxene chemical dataset to estimate the pressure (P) and temperature (T) at which crystals grew. RF is a machine learning approach capable of building a robust prediction that extends the capability of the decision trees algorithm (Breiman 2001; Simm et al. 2014; Jorgenson et al. 2022). In the decision trees algorithm (Breiman 2001), each tree is a hierarchical flowchart composed of numerous branches. Their structure is determined using training data, in our case the chemical composition of experimentally derived clinopyroxene (Hirschmann et al. 2008; Jorgenson et al. 2022). Each branch ends with a prediction of the investigated parameter (i.e., P or T). The decision tree algorithm can be applied to natural data by following the branch structure to arrive at P or T estimate. However, predictions based on single decision tree models are inaccurate due to so-called overfitting (Zhou et al. 2020). If overfitting occurs, the model can only accurately predict the experimentally derived clinopyroxene (training dataset) but not the natural samples (unknowns). Thus, a more robust model, like the RF, is required (Petrelli et al. 2020; Zhou et al. 2020; Jorgenson et al. 2022). Following the suggestions reported by Jorgenson et al. (2022), we develop a RF model characterized by 200 trees, resulting in 200 different predictions of the investigated parameter (i.e., P and T) for each unknown sample. For natural data prediction, we used the median of the voting distribution as the best P–T estimate. In our specific case, the RF algorithm was applied using the open-source R script available on GitHub (https://github.com/corinjorgenson/RandomForest-cpx-thermobarometer) provided by Jorgenson et al. (2022). The results of the hierarchical clustering and random forest thermobarometry analyses can be found in the supplementary material table (Supplementary Table 2)

Fig. 2 **a** Plot of the first and second principal components (PCA1 and PCA2) from the principal component analysis performed on the geochemical dataset and color contoured according to the number of clusters. **b** BSE image on a sector zoned cpx with the analyzed points colored according to the number of clusters. **c** and **d** dendrogram divided in 2 and 6 clusters, respectively
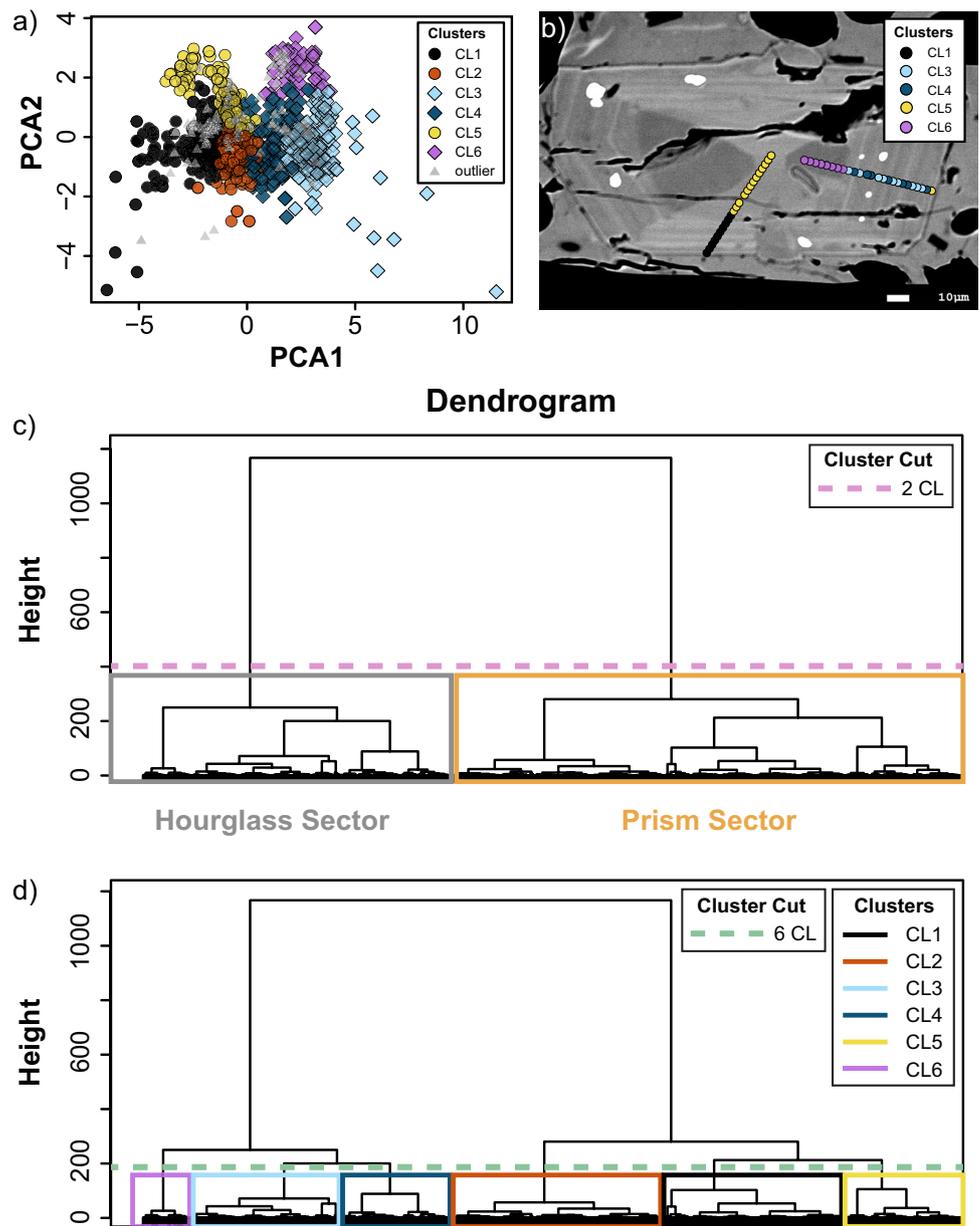

Random forest classifier

We finally test a RF classifier assigning clinopyroxene analyses from volcanic products that erupted from the SEC in the last two decades to the clusters we identified for the period between February and April 2021. We trained the algorithm using the clinopyroxene dataset of this work and applied it to clinopyroxene analyses for the eruptions of 2002–2003; 2006; 2007; 2008; 2011–2012, available on GEOROC (DIGIS Team 2022). This approach is useful to compare eruptions and eventually identify patterns between the fraction of different clusters and eruptive dynamics. The RF algorithm was trained on 200 trees. The performance of the method was tested by randomly removing 10% of the

observations from the starting train dataset, for which the chemical clusters were known, and applying the model to this 10% by treating it as an unknown. The classifier correctly classified over 95% of the observations in the test dataset.

Textural and chemical complexity of minerals

To quantify compositional and textural variations among and within different paroxysms, we define a new parameter, the textural complexity (tx_{comp}), calculated as the number of times a change of cluster occurs from the rim to the core of a crystal (C_{cl}) divided by the length of the analyzed profile (L_a):

Fig. 3 **a**, **b**, **c**, and **d** back scat-
tered images of clinopyroxene
from the February 16, 19, and
28 and March 02 in order. The
analyzed points on the clinopy-
roxene are colored according to
the number of clusters. **e** Violin
plots of cluster composition for
all the analyzed elements
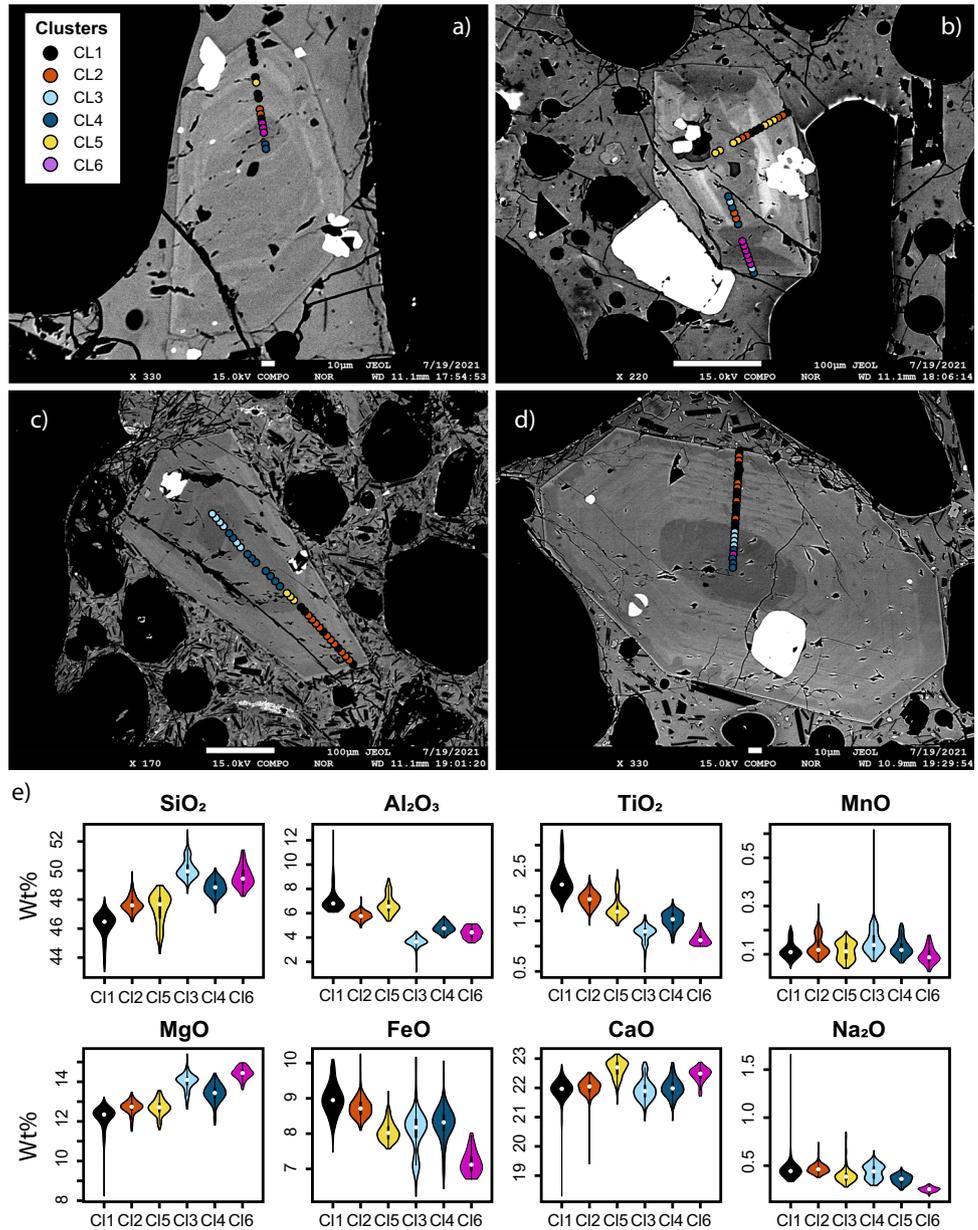
$$tx_{comp} = {}^{C_{cl}}\!/_{L_a}. \tag{6}$$

We couple this parameter with the variance from rim to
core for each oxide normalized for the length of the analyzed
transect. These two parameters provide a quantitative meas-
ure of the textural and chemical complexity of each crystal.

To ensure the representativeness of these parameters,
special attention must be paid to the crystal selection. The
sectioning effect, i.e. how the 3D crystal was cut to produce
the analyzed 2D section, is not negligible (Shea et al. 2015;
Cheng et al. 2017). The random cuts of a crystal can be
responsible for the presence of a wide variety of artifactual
zonation (e.g., different cuts of crystals with the same num-
ber of zones — Cheng et al. 2017). To limit this effect, only

the largest crystals (major axis > 80 µm) showing similar
section cuts were used in the calculation of chemical and
textural parameters. In this process, the presence of sector
zoning helped to identify similar sections.

Results

The analyzed lapilli have trachy-basaltic composition, and
their mineral assembly is constituted of plagioclase+clino
pyroxene+olivine+oxide phenocrysts (microcrysts having
a length < 100 µm and mesocrysts with a length between
100 and 500 µm, Zellmer 2021; Mollo et al. 2022). The
groundmass has differing proportions of microlites for each

eruption, and the microlites are mostly composed of plagioclase and clinopyroxene. All the analyzed clinopyroxene phenocrysts show concentric oscillatory zonation, and almost all also show a clear hourglass zonation. The glass composition shows an increase in Al_2O_3 and MgO from February 16 until February 28, which then decrease in the 2 and 10 of March eruptive products (Figs. S1 and S4; Supplementary Table 3), in agreement with data from Corsaro and Miraglia (2022).

Cluster identification

The dendrogram shows two clear groups of clusters (Fig. 2c, d). As shown in Fig. 2c, the two clusters texturally distinguish between hourglass sectors, and prism sectors as their chemical composition are the most compositionally contrasted. However, two clusters are insufficient to separate discrete cores and concentric zones. Six clusters provide the best partitioning of data from both chemical and textural perspectives as shown by the plot of the two principal components of PCA and the correspondence between the cluster and the zones identified in the BSE images (Figs. 2a, b and Figs. 3a–d). Attempting to partition the data in a larger number of clusters results in a lack of correspondence between cluster and textural features and in a lower quality of the chemical split in the PCA1–PCA2 space. This approach reinforces the importance of combining textural and chemical observations when describing zoned magmatic crystals (Sheldrake and Higgins 2021).

Cluster chemistry

Three of the six clusters (CL1, 2, and 5) are found in some cores, in concentric zones within the prism sectors, and within outer growth rims. The other three clusters (CL3, 4, and 6) are only found in hourglass sectors, either in cores or mantles (Fig. 3a). Many of the analyzed crystals are cut perpendicular to the c-axis, where the hourglass sector occupies the center of the crystals (Leung 1974; Ubide et al. 2019). However, other crystals are cut so that only small portions of the hourglass zoning are visible. This results in an uneven number of points analyzed in the hourglass zone, which can vary from sample to sample. This non-uniform sampling could lead to misleading variations in the fraction of the hourglass clusters (CL3, CL4, and CL6) between different paroxysms. Thus, we only consider the analyses collected within the prism sector.

In general, the hourglass sectors present lower concentrations of Al_2O_3 and TiO_2 (4.17 and 1.34 wt.% with respect to 6.42 and 2.01 wt.% in the prism sectors, for Al_2O_3 and TiO_2, respectively) and an enrichment in SiO_2 and MgO (with a mean composition of SiO_2 and MgO of 49.56 and 13.84 wt.% for the hourglass and 47.13 and 12.53 wt.% for the

prism sectors) compared to the prism sector. The concentric zonation within prism sectors displays an alternation of CL1, CL2, and CL5 (Figs. 2b and Figs. 3a–d). In detail, comparing these three clusters with each other, it can be seen that CL1, among the three, exhibits a higher median content in FeO (8.95 wt.%), TiO_2 (2.21 wt.%), and Al_2O_3 (6.78 wt.%), and a lower content in SiO_2 (46.47 wt.%), CaO (21.97 wt.%), and MgO (12.35 wt.%). CL2 and CL5 present a higher and similar median content of SiO_2 (47.6 wt.% for both) and MgO (12.74 wt.% for CL2 and 12.70 wt.% for CL5). While they are similar in SiO_2 and MgO, CL5 show a higher content in CaO (CL5: 22.71 wt.%, CL2: 22.05 wt.%) and Al_2O_3 (CL5: 6.57 wt.%, CL2: 5.77 wt.%) and lower FeO (CL5: 8.02 wt.%, CL2: 8.71 wt.%) and TiO_2 (CL5: 1.68 wt.%, CL2: 1.93 wt.%; Fig. 3e). The pie charts in Fig. 4a shows the distribution of CL1, CL2, and CL5 through different lava fountains. Analyses collected on clinopyroxene erupted on February 16 and 19 show a similar proportion of the three clusters consisting of an almost equal amount of CL1 and CL5, and minor content of CL2. The CL2 content increases abruptly in clinopyroxene from the February 28 event, before decreasing gradually through the March 2 and 10 paroxysms (Fig. 4a).

Cluster distribution within crystals

The variations of the relative proportions of clusters over time are associated with changes in the distribution of clusters from the core to the rim of the crystals (Fig. 4b). For the February 16 episode, the majority of the rims are CL2 or CL1 in an almost equal amount, the mantle is mainly constituted of CL5 and the core of CL2 or CL1. In the February 19 crystals, the rims are mainly CL1, the middle part is mostly an alternation of CL1 and CL2, and the cores are predominately CL5.

The February 28 paroxysm shows more uniform crystals, constituted mostly of CL2; CL5 can be found mainly in the cores and in the middle portion of the crystals, and CL1 is in all crystals as a small concentric zone. Most of the March 2 clinopyroxene presents CL1 in the rims and CL2 in the central portion (mantles), with CL5 as a dominant phase in cores. Lastly, the March 10 clinopyroxene presents an oscillating distribution of CL1, which alternates with CL5 and CL2 in complex concentric zoning from rim to core. CL2 can be found everywhere from the rim to the core and CL5 mainly in the rims. Overall, we can observe a shift from rim to cores of CL5 from February 16 to the event of March 2, and CL5 mostly in the rim in the March 10 event. CL1 usually represents the rims, with the exception of the February 16 paroxysm where its distribution is generally constant through the profiles, and the March 10 event where an oscillating distribution is observed. CL2 can be found both in cores and in alternation with CL1 and CL5 in rims

Fig. 4 **a** Prism sector cluster distribution in different eruptive events. **b** cluster distribution between cores and rims of clinopyroxene, within all the studied episodes, normalized between 0 and 1
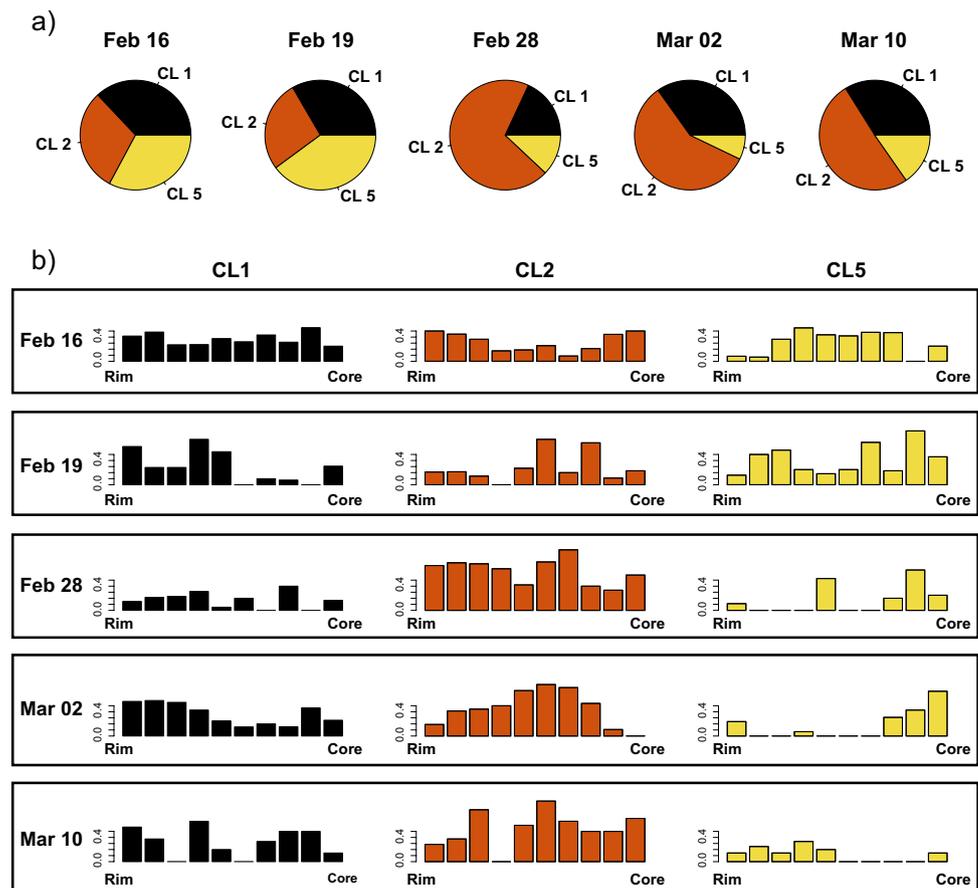
and in the central part of the crystals, and it is the dominant phase everywhere on February 28.

Random forest thermobarometry

The random forest thermobarometry results for our dataset are shown in Fig. 5a, where the estimates of P and T median values are plotted and color-contoured according to their respective clusters. Most analyses record a pressure between 1 and 3 kbar (SEE = 2.6 kbar) and a temperature between 1070 and 1130 °C (SEE = 58.9 °C). Such values agree with literature estimates of Etnean magma storage (Murru et al. 1999; Spilliaert et al. 2006; Giacomoni et al. 2014; Mollo et al. 2022). While the pressure is similar for the three clusters, they seem separated in terms of temperature, with CL1 recording the lowest temperatures and CL5 the highest. Even if the single estimates are within the uncertainty of our approach, differences are still visible in the distributions (Fig. 5a). While CL1 presents lower temperatures and more evolved compositions, CL2 and CL5 register higher temperatures and show a similar Mg# (Fig. 5a, b). CL2 shows higher CaO/Al_2O_3 values with respect to CL5. The Mg# and the

crystallization temperature reach higher values in CL5 clinopyroxenes. (Fig. 5).

Random forest cluster classifier

We tested the potential of a random forest classifier to identify clusters in previous eruptions, using the data we collected in 2021 as a training dataset. The results of the random forest cluster classification on previous SEC eruptions are shown in Fig. 6. The clusters identified for the 2021 eruptive sequence describe sufficiently well the chemical variability of past eruptions, with the exception of a group of data in the 2006 eruption. The chemistry of this group of clinopyroxenes is in agreement with xenocrystals, potentially in equilibrium with the mantle (Fig. 6b). Even if the xenocrystic nature of these crystals would require petrographic confirmation, this represents a good example of something that might happen during an eruption when clinopyroxenes of a previously unrecognized group are analyzed. To construct a reliable geochemical classifier for all Etna crystals, it is hence necessary to build a larger geochemical dataset that collects all existing clinopyroxene analyses and a description of their textural location. Once such a dataset will be constructed, the classifier we

Fig. 5 a P and T estimates for each analyzed point, color contoured according to the cluster number, the error bars are based on the interquartile range. On top of the plot, the kernel density estimate normalized between 0 and 1 show the density distribution of T estimates for the three clusters. **b** Magnesium number (Mg#) plotted vs CaO/Al$_2$O$_3$
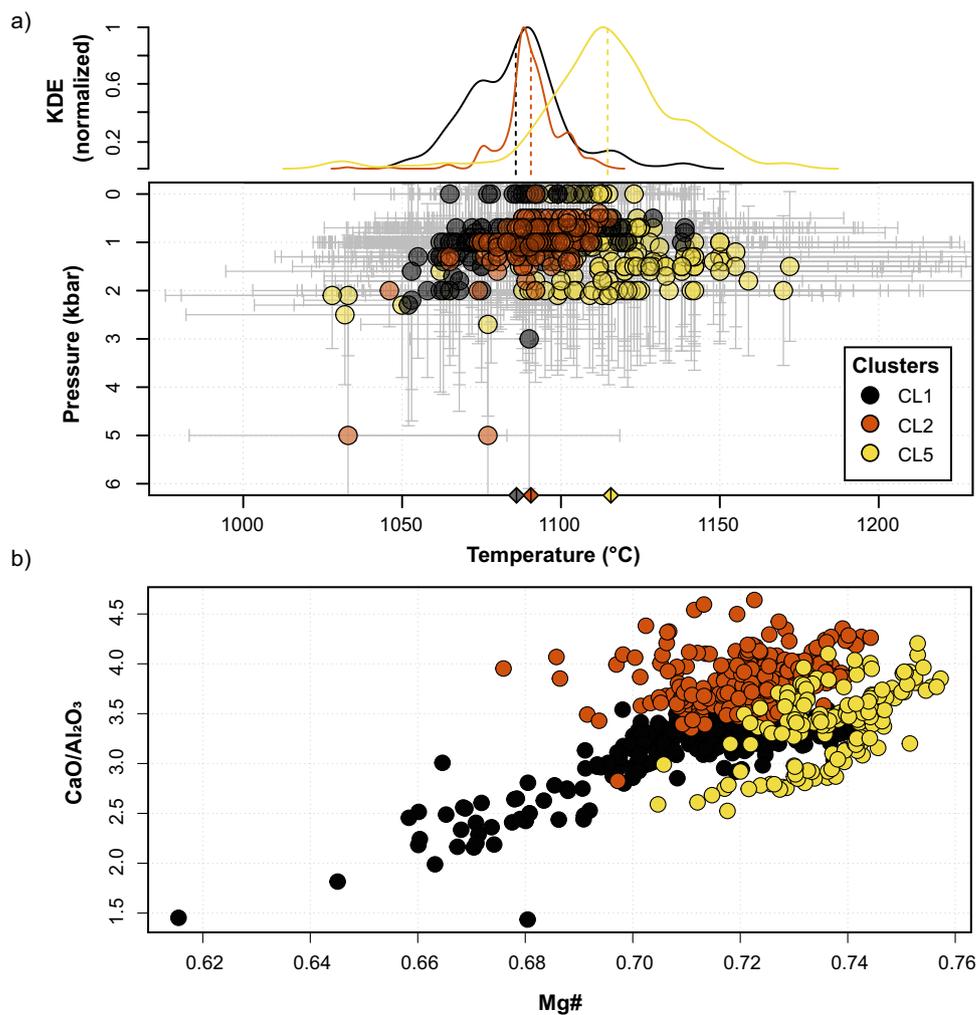
present here will allow the identification of the cluster for each collected analysis in clinopyroxene and therefore the rapid quantification of the proportions of different clusters in the erupted material. This approach would have three major advantages: (1) the rapid quantification of the proportions of analyses belonging to different clusters; (2) the quantitative comparison with previous eruptive events; (3) anticipate the evolution of an ongoing eruptive sequence on the base of previously identified quantitative temporal trends in the fraction of analyzed clusters.

Discussion

Clusters as recorders of the evolution of crystallization conditions

The February 28 events have the lowest counts for the lower temperature cluster (CL1) and thus the highest amount of analyses assigned to CL2 and CL5 (Fig. 4a). The increased contribution of mafic magma to this event is confirmed by the

glass analyses (Fig. S2 and Supplementary Table 3), showing a peak of CaO/Al$_2$O$_3$ and MgO content on February 28. This trend was also observed by Corsaro and Miraglia (2022), who analyzed the glass composition of the SEC 2021 eruptive products over a wider time span (Dec 2020–Apr 2021) showing the same peak of higher mafic compositions on February 28. The observed trends in glass and crystal-chemical variation suggest the interaction and mixing between an increasing proportion of a hotter and more mafic melt with a slightly colder and more evolved magma until February 28, followed by a decrease of the hot magma input. This hypothesis is also confirmed by the observed clinopyroxene concentric zonation, which could be an expression of magma mixing (Ubide and Kamber 2018; Ubide et al. 2019; Mollo et al. 2022).

We suggest that declining rates of mafic magma input after February 28 are reflected by increasing fraction of lower temperature CL1 clinopyroxene analyses. This is also confirmed by the core-to-rim distribution of clusters and the increasingly evolved nature of clinopyroxene rims after the episode of February 28 (Fig. 7b).
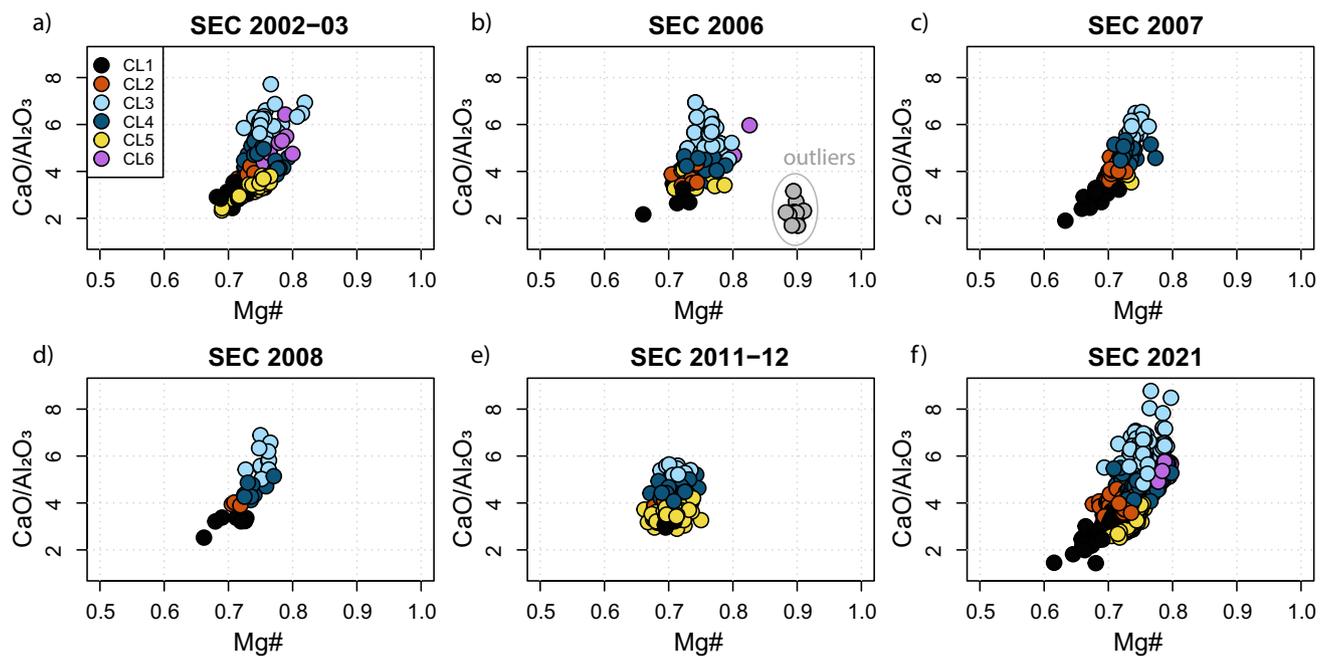
Fig. 6 Mg# vs CaO/Al$_2$O$_3$ composition of clinopyroxene from different SEC eruptions. The colors of the points represent the cluster estimated by an RF classifier which was trained on the dataset from this paper. Circled in light gray are the chemical outliers revealed in the chemical analysis of cpx from the 2006 eruption

Mollo et al. (2022) studied the growing conditions of clinopyroxene during the same eruptive period, giving us a broader view of the fountaining sequence. These authors note that clinopyroxene that erupted on February 28 records higher crystallization temperature compared with the previous events, in agreement with our observations. They also analyzed crystals from the final two episodes of the eruptive sequence (March 24 and 31), where clinopyroxene crystallization temperatures were similar to those that erupted on February 28. Altogether these results suggest that the period from February to the end of March was characterized by two major events of input of mafic magma from depth at the end of February and at the end of March 2021.

While the chemical characteristics and estimated temperature for CL1 can be reconciled with the crystallization of these clinopyroxenes at relatively low temperatures in a cooling superficial reservoir, it remains unclear which process can produce the difference in CaO/Al$_2$O$_3$ and Mg# observed between CL2 and CL5. While more data would be required, we provide in the following a speculative interpretation. Looking at the thermobarometry result obtained from clinopyroxene chemistry, the lack of high-pressure estimates is evident (Fig. 5a). This can be either the product of lack of crystallization or resorption. Considering the water content of magmas at Mt. Etna (Métrich et al. 2004), their liquidus temperature increases with increasing pressure from a few kilometer depth. As the liquidus phase for these magmas is clinopyroxene (Fig. S3; calculations performed with

rhyolite-MELTS; Gualda et al. 2012), magmas extracted deeper will have lower CaO/Al$_2$O$_3$ content because of more extensive clinopyroxene fractionation with respect to magmas extracted shallower at the same temperature. Thus, we could speculate that CL2 clinopyroxene crystallized from a melt extracted from a shallower depth with respect to the melt from which CL5 clinopyroxenes crystallized (Fig. 5b).

A second possibility is that partial resorption of clinopyroxene upon ascent resulted in an increased CaO/Al$_2$O$_3$ content of the melt from which CL2 clinopyroxenes crystallized at shallower depths. In this scenario, clinopyroxenes from CL5 crystallized from a melt that was extracted near liquidus, and therefore, no resorption could occur. This would agree with the higher temperatures recorded by clinopyroxenes from CL5 (Fig. 5a). The resorption of about 1.2 wt.% of clinopyroxene is sufficient to explain the difference in the CaO/Al$_2$O$_3$ in the glass between the 28 of February and the other fountaining events (Fig. S4).

Cluster variabilities and eruptive dynamics for the February 28 episode

The proportion of clusters in samples from the 28 of February lava fountain is clearly different from the other episodes we investigated (Fig. 4a). Additionally, this event shows the lowest and least variable textural complexity of all eruptions, together with the lowest oxide (e.g. MgO and CaO) variance (Fig. 7a, b, and c). Hence, our results suggest that

Fig. 7 Box plots of: **a** textural complexity; **b** and **c** MgO and CaO variance for each episode; **d** and **e** MgO and Al_2O_3 clinopyroxene content in rim and cores for each paroxysm
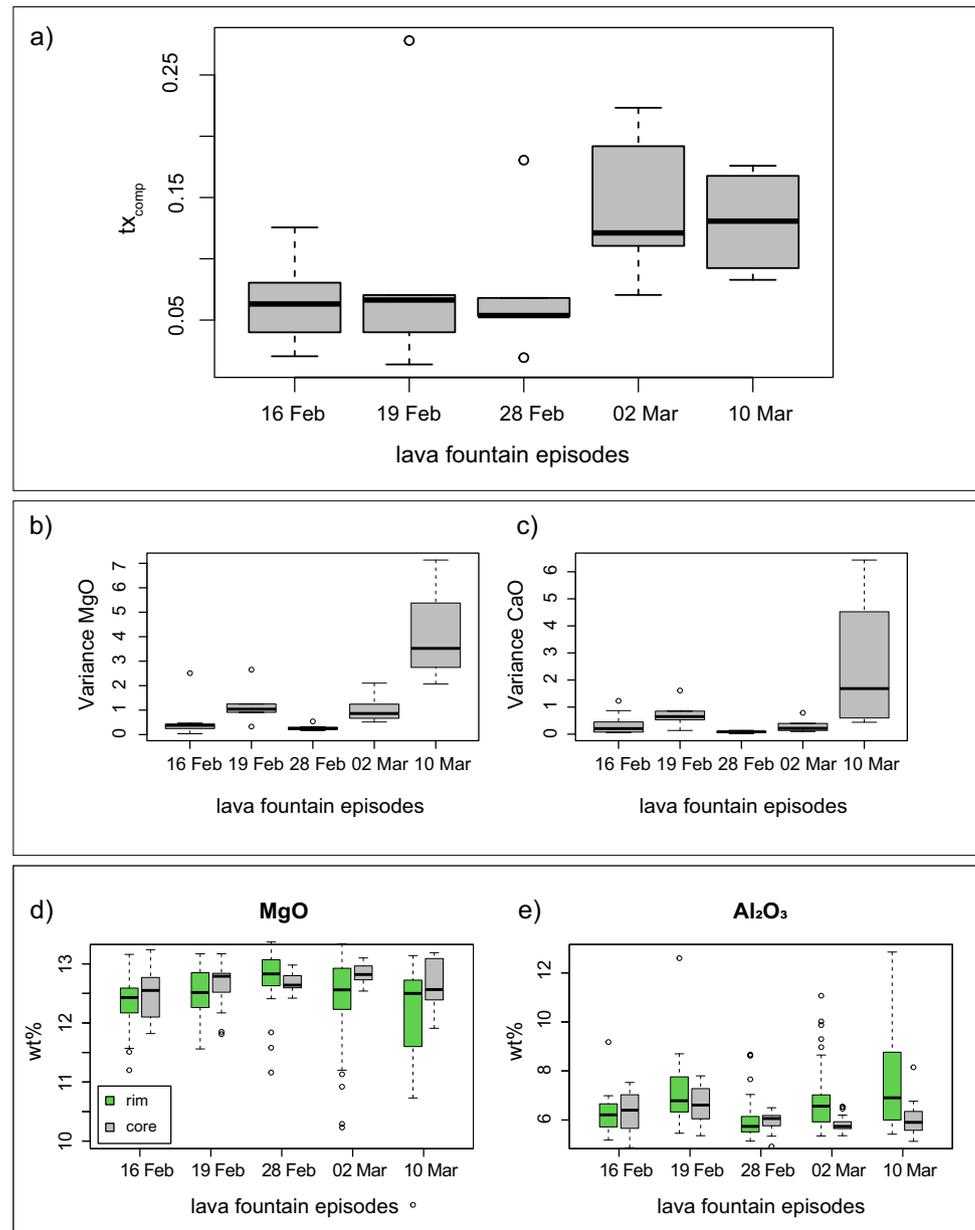
the February 28 event was fed by a more mafic magma, which potentially ascended faster (quasi adiabatically leading to an increase of the fraction of high CaO/Al_2O_3 CL2 clinopyroxenes). The lower textural and chemical complexity (low oxide variance along the crystal profile) of clinopyroxene from this event could be related to the growth of crystals in a more homogeneous environment with a larger volume of mafic magma. The presence of greater amounts of mafic magma is also confirmed by the results of our glass analyses (Fig. S2). The eruption of February 28 also shows a greater proportion of microlites, an observation confirmed by the CSD analyses conducted by Mollo et al. (2022). This could be due to a higher nucleation rate caused by the higher

ascent velocity of a more mafic and hotter magma (Vona and Romano 2013; Vetere et al. 2022).

A comparison between clinopyroxene core and rim chemistry through the entire eruptive sequence shows that cores are more chemically homogeneous with respect to rims and the chemical variance of rims reaches a minimum in the products of February 28 before increasing again until the event of March 10 (Fig. 7d, e). This suggests that progressive input of mafic magma led to the most mafic lava fountain of February 28, after which the input of mafic magma declined, allowing for the growth of rims reaching the most chemically evolved composition with the last event we investigated. Existing data (Mollo et al. 2022) suggest

another increase of the mafic input from depth occurred at the end of March 2021.

Linking quantitative petrology to eruptive dynamics

Differences in the rate of magma input, chemistry, and temperature of the erupted magma can affect the eruptive dynamics because of the control of these intensive parameters on magma viscosity (Giordano et al. 2008). To investigate this hypothesis, volcanological and geophysical parameters were collected from the INGV multidisciplinary weekly bulletins (INGV-OE, Bollettini Settimanali sul monitoraggio vulcanico, geochimico e sismico del vulcano Etna at: https://www.ct.ingv.it/index.php/monitoraggio-esorveglianza/prodotti-del-monitoraggio/bollettini-settimanalimultidisciplinari) and from published literature (Andronico et al. 2021; Calvari and Nunnari 2022). Specifically, we analyzed the variation of repose time before the eruptive event, the mean fountain height, the erupted volume during the fountaining activity, the time-averaged discharge rate (TADR), and the cumulative reduced displacement (RD_{cum}) between the 5 eruptive episodes studied in this work. RD_{cum} is the root mean square (RMS) of the volcanic tremor amplitude normalized for the distance of the measuring station from the source (Andronico et al. 2021). The RD value is cumulative and so accounts for values of volcanic tremor amplitude collected over a well-defined time window in which the tremor was recorded (Andronico et al. 2021). TADR represents the flux of volume of erupted tephra averaged over a specific time interval (Harris et al. 2007; Calvari and Nunnari 2022).

This analysis shows that the fraction of mafic magmas (i.e., clinopyroxenes from CL2 and CL5), which peaks on the 28 of February, follows a similar pattern to repose time, mean height of the lava fountain, and TADR (Fig. 8a). A good correlation is observed between the variance of MgO and Al_2O_3 in clinopyroxene, RD_{cum} and the volume erupted by the lava fountain (Fig. 8b). Finally, we observe an anticorrelation between the fraction of CL2 and RD_{cum} (Fig. 8b). These quantitative relationships should be further explored and potentially implemented for effective petrological monitoring in periods of protracted magmatic activity.

In agreement with the results of La Spina et al. (2021), our results show that the hotter and more mafic magma (i.e., lower viscosity) erupted on 28 February 2021, which led to the highest lava fountain event and TADR of the investigated period. The lower magma viscosity of this event probably caused faster rates of magma ascent. Moreover, higher rates of ascent (corresponding to a higher magma flow) increase crystal nucleation (Vona and Romano 2013), which would account for the more abundant microlite content of the magmas that erupted on February 28 (Fig. 3c and supplementary Fig. S1).

A model for the February–April SEC eruptive sequence

We suggest that the paroxysmal activity which occurred during the February–April 2021 eruptive sequence at Etna was driven by the injection of mafic and deep magma into a region of the plumbing system that extends over a pressure range of 1–3 kbar, as suggested by our barometric estimates, some of which erupted to the surface. Here, new magma is mixed with slightly more evolved magma left from previous eruptions in January (Corsaro and Miraglia 2022). The mixing explains the complex concentric zoning in the rim characterized by alternating CL1 and CL2. After the February 28 eruption, the magma supply to the storage area decreased, and the occurrence of fractional crystallization processes produced magmas that gradually evolved to compositions comparable to those produced in January 2021. The February 28 eruption represents the peak of the February-April eruptive series involving the highest portion of mafic and hotter magma, and it is characterized by a higher ascent rate and fountain height.

The RF cluster classifier

We have shown and validated on a known dataset that using a clustered dataset is a useful tool for robustly identifying discrete chemical clusters, which in turn can be related to magmatic processes as well as volcanological and geophysical monitoring parameters. We suggest that the RF cluster classifier method has the potential to recover the eruptive dynamics of past eruptions and to quickly establish a link between petrology and volcanological parameters.

To demonstrate the wider applicability of our method to historical eruptions, and to compare the chemical variability of clinopyroxene crystals measured in this work with those of past eruptions, we investigate a large dataset of clinopyroxene chemical composition from Mt. Etna, taken from GEOROC (DIGIS Team 2021). In Fig. 6, we see that the RF classifier distinguishes the proportions of clusters present in products from different eruptive periods. This result underlines the capability of this method to be used as a tool for reconstructing the chemical variations over time as recorded by crystals. To link mineral chemistry and eruptive dynamics, as we did in this study, the exact date of sample collection is necessary. However, this is not always available in the GEOROC database, where often only the year of the eruption is reported. As an example, in 2000, there were 64 lava fountain episodes (Andronico and Corsaro 2011). The connection between mineral chemistry and eruptive dynamics using our approach is thus only possible for samples for which the exact collection date is reported. Nevertheless, the results obtained on historical
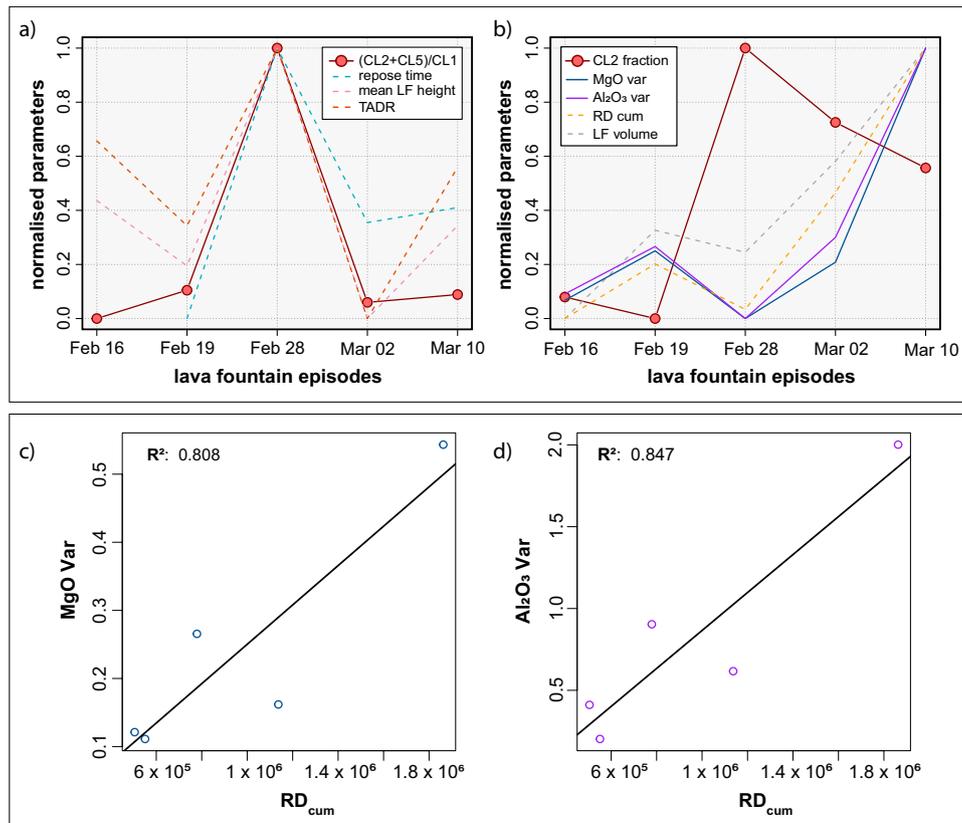
Fig. 8 **a** Ratio between the mafic (CL2 and CL5) and more evolved (CL1) cluster fractions plotted for each lava fountain episode and compared with volcanological parameters such as: the repose time before the eruptive event, the mean lava fountain (LF) height, and the time-averaged discharge rate (TADR). All the parameters have been normalized, the information about repose time has been collected thanks to the multidisciplinary weekly bulletin available on the INGV website, while mean LF height and TADR have been taken from Cal-

vari and Nunnari (2022). **b** The fraction of cluster 2 and the normalized variance of MgO and Al_2O_3 in clinopyroxene are plotted for each paroxysm and compared with the RD_{cum} (the RD_{cum} data has been taken from Andronico et al. 2021) and the LF erupted volume (from Calvari and Nunnari 2022). All the parameters have been normalized. **c** and **d** Regression lines and R^2 are reported to mark the relationship between RD_{cum} and the MgO and Al_2O_3 variance from core to rim of the crystals

eruptions allow us to make several conclusions: first, we note that while six clusters might be enough to describe the chemical variability of products erupted by Mt. Etna over the last 20 years, an additional cluster for xenocrystals is necessary (Fig. 6). Additionally, our method can easily discriminate geochemical clusters and identify whether the analysis was collected on the prism or on the hourglass section. Finally, we propose a set of guidelines to make our approach more robust for quantitative petrological monitoring. During prolonged eruptions, the comparison should be facilitated by the collection of samples with similar characteristics (e.g., tephra or lavas for which the emission date is known) and a consistent analytical approach to keep track of the zone of the crystals in which the chemical transects and/or spot analysis are acquired (i.e., core and rim). Clearly, the assembly of a larger dataset collected with the sampling and analytical strategy we describe above will serve to improve the accuracy of our

approach in linking the chemical evolution of minerals and eruptive dynamics.

Conclusions

This study shows that the application of robust data transformation and normalization combined with HC analysis on geochemical datasets is a useful tool for identifying clusters of similar chemical-textural zones in minerals. This, in tandem with RF thermobarometry, can be used to trace specific P-T-X growth conditions and constrain the distribution of clusters on inter-eruption and intra-eruption base. The HC analysis on clinopyroxene underlines the presence of six chemically and texturally distinct clusters within the February 16 to April 1 Etna eruptive sequence. CL3, CL4, and CL6 are related to hourglass zones and CL1, CL2, and CL5 are related to prism sectors and to

the concentric zoning. CL2 and CL5 represent a more mafic endmember, which proportionally increases up to the February 28 eruptive event before then decreasing. The application of RF thermobarometry to the clustered dataset allowed us to relate chemical clusters to P and T conditions. We estimate clinopyroxene crystallization to occur in a pressure range of 1–3 kbar and in a temperature range of 1070 and 1130 °C, in line with previous studies. CL1, CL2, and CL5 formed in the same pressure range but at different temperatures, where CL1 is a low-T cluster, and CL2 and CL5 are high-T clusters. We conclude that the 2021 Etna eruptive sequence was sustained by the intrusion of deep mafic magma into a storage area located at 1–3 kbar. Here, the mixing between a new fresh magma and a more evolved magma remaining from previous eruptions in January took place.

RF algorithms offer an effective and rapid way to classify and study the distribution of unique chemical clusters in different eruptive events. This represents a useful step in identifying the emergence of chemical compositions related to specific eruptive dynamics during an ongoing eruptive series, bringing a further contribution to the importance of syn-eruptive petrological monitoring. This has the potential to unveil magmatic processes at a depth that can control eruptive dynamics. More generally, the methodology shown in this work represents a fundamental tool for analyzing chemical variations in a robust and quantitative manner. This approach offers the possibility of quantifying chemical variations both within and between different eruptive events, allowing the direct correlation of the chemical variations with other monitoring parameters and enabling the construction of new parameters capable of tracing the chemical and textural complexity recorded by crystals. Moreover, we have shown the use of this technique on 1D profiles with its limitations. Extending the approach to 2D maps, in the future, could lead to a quantitative identification of zoning patterns and crystal families that underwent similar growth histories (e.g., Jerram and Martin 2008; Cheng et al. 2017; Sheldrake and Higgins 2021; Higgins et al. 2021). Finally, the use of clustering in conjunction with thermo-barometric methods allows rapid correlation between representative chemical compositions and their forming conditions.

Declarations

References

Abdi H, Williams LJ (2010) Principal component analysis. Wiley interdiscip rev: comput stat 2:433–459

Aitchison J (1986) The statistical analysis of compositional data. Springer, Netherlands, Dordrecht

Aitchison J, Egozcue J (2005) Compositional data analysis: where are we and where should we be heading? Math Geol 37:829–850

Alparone S, Andronico D, Lodato L, Sgroi T (2003) Relationship between tremor and volcanic activity during the Southeast Crater eruption on Mount Etna in early 2000. J Geophys Res Solid Earth 108(B3)

Andronico D, Corsaro RA (2011) Lava fountains during the episodic eruption of South–East Crater (Mt. Etna), 2000: insights into magma-gas dynamics within the shallow volcano plumbing system. Bull Volcanol 73:1165–1178

Andronico D, Cannata A, Di Grazia G, Ferrari F (2021) The 1986–2021 paroxysmal episodes at the summit craters of Mt. Etna: insights into volcano dynamics and hazard. Earth Sci Rev 220:103686

Bachmann O, Bergantz G (2008) The magma reservoirs that feed supereruptions. Elements 4:17–21

Barboni M, Schoene B (2014) Short eruption window revealed by absolute crystal growth rates in a granitic magma. Nat Geosci 7:524–528

Behncke B, Neri M (2003) Cycles and trends in the recent eruptive behaviour of Mount Etna (Italy). Can J Earth Sci 40:1405–1411

Bisciotti A, Comodi P, Fastelli M et al (2022) Transport amphoras from Gravisca to the eastern Aegean during the Greek archaic period (7th-6th c. BC): a multidisciplinary approach for characterization, manufacturing, provenance and trade. Mediter Archaeol Archaeom 22:179–202

Blundy J, Cashman K (2008) Petrologic reconstruction of magmatic system variables and processes. Rev Mineral Geochem 69:179–239

Bonaccorso A, Carleo L, Currenti G, Sicali A (2021) Magma migration at shallower levels and lava fountains sequence as revealed by borehole dilatometers on Etna volcano. Frontiers. Earth Sci 800

Boschetty FO, Ferguson DJ, Cortés JA et al (2022) Insights into magma storage beneath a frequently erupting arc volcano

(Villarrica, Chile) from unsupervised machine learning analysis of mineral compositions. Geochem Geophys Geosyst 23:e2022GC010333

Branca S, Del Carlo P (2004) Eruptions of Mt Etna during the past 3.200 years: a revised compilation integrating the Historical and stratigraphic records. In: Mt. Etna: Volcano Laboratory. American Geophysical Union (AGU), pp 1–27

Branca S, Del Carlo P (2005) Types of eruptions of Etna volcano AD 1670–2003: implications for short-term eruptive behaviour. Bull Volcanol 67:732–742

Branca S, Coltelli M, Groppelli G (2004) Geological evolution of Etna volcano. Washington DC Am Geophys Union Geophys Monog Ser 143:49–63

Breiman L (2001) Random forests. Machine Learning 45:5–32. https://doi.org/10.1023/A:1010933404324

Butler JC (1976) Principal components analysis using the hypothetical closed array. J Int Assoc Math Geol 8:25–36

Calvari S, Nunnari G (2022) Comparison between automated and manual detection of lava fountains from fixed monitoring thermal cameras at Etna volcano. Italy Remote Sensing 14:2392

Calvari S, Salerno GG, Spampinato L et al (2011) An unloading foam model to constrain Etna's 11–13 January 2011 lava fountaining episode. Journal of Geophysical Research: Solid. Earth 116

Calvari S, Cannavò F, Bonaccorso A et al (2018) Paroxysmal explosions, lava fountains and ash plumes at Etna Volcano: eruptive processes and hazard implications. Frontiers. Earth Sci 107

Calvari S, Bonaccorso A, Ganci G (2021) Anatomy of a paroxysmal lava fountain at Etna volcano: the case of the 12 March 2021. Ep Rem Sens 13:3052

Cappello A, Bilotta G, Neri M, Negro CD (2013) Probabilistic modeling of future volcanic eruptions at Mount Etna. J Geophys Res Solid Earth 118:1925–1935

Caricchi L, Petrelli M, Bali E et al (2020) A data driven approach to investigate the chemical variability of clinopyroxenes from the 2014–2015 Holuhraun–Bárdarbunga Eruption (Iceland). Front Earth Sci 8:18

Carranza EJM (2011) Analysis and mapping of geochemical anomalies using logratio-transformed stream sediment data with censored values. J Geochem Explor 110:167–185

Cashman K, Blundy J (2013) Petrological cannibalism: the chemical and textural consequences of incremental magma body growth. Contrib Mineral Petrol 166:703–729

Cassidy M, Manga M, Cashman K, Bachmann O (2018) Controls on explosive-effusive volcanic eruption styles. Nat Commun 9:2839. https://doi.org/10.1038/s41467-018-05293-3

Charrad M, Ghazzali N, Boiteau V, Niknafs A (2014) NbClust: an R package for determining the relevant number of clusters in a data set. J Stat Softw 61:1–36. https://doi.org/10.18637/jss.v061.i06

Cheng L, Costa F, Carniel R (2017) Unraveling the presence of multiple plagioclase populations and identification of representative two-dimensional sections using a statistical and numerical approach. Am Mineral: J Earth Planet Mater 102:1894–1905

Corsaro RA, Miraglia L (2022) Near real-time petrologic monitoring on volcanic glass to infer magmatic processes during the February–April 2021 Paroxysms of the South-East Crater, Etna. Front Earth Sci 10:828026. https://doi.org/10.3389/feart.2022.828026

Cortés JA, Palma JL, Wilson M (2007) Deciphering magma mixing: the application of cluster analysis to the mineral chemistry of crystal populations. J Volcanol Geotherm Res 165:163–188

Costa F (2008) Residence times of silicic magmas associated with calderas. Developments in volcanology 10:1–55

De Gori P, Giampiccolo E, Cocina O et al (2021) Re-pressurized magma at Mt. Etna, Italy, may feed eruptions for years. Communications Earth & Environment 2:1–9

Di Renzo V, Corsaro RA, Miraglia L et al (2019) Long and short-term magma differentiation at Mt Etna as revealed by Sr-Nd isotopes and geochemical data. Earth-Sci Rev 190:112–130

DIGIS Team (2021) GEOROC compilation: minerals. GRO.data. https://doi.org/10.25625/SGFTFN

DIGIS Team (2022) 2022-03-SGFTFN_CLINOPYROXENES. GRO. data. https://doi.org/10.25625/SGFTFN/XDJHPT

Eesa AS, Arabo WK (2017) A normalization methods for backpropagation: a comparative study. Sci J University of Zakho 5:319–323

Egozcue JJ, Pawlowsky-Glahn V, Mateu-Figueras G, Barcelo-Vidal C (2003) Isometric logratio transformations for compositional data analysis. Math Geol 35:279–300

Ferlito C, Coltorti M, Lanzafame G, Giacomoni PP (2014) The volatile flushing triggers eruptions at open conduit volcanoes: evidence from Mount Etna volcano (Italy). Lithos 184:447–455

Filzmoser P, Hron K, Reimann C (2009) Univariate statistical analysis of environmental (compositional) data: problems and possibilities. Sci Total Environ 407:6100–6108

Fry JM, Fry TR, McLaren KR (2000) Compositional data analysis and zeros in micro data. Appl Econ 32:953–959

Giacomoni PP, Ferlito C, Coltorti M et al (2014) Plagioclase as archive of magma ascent dynamics on "open conduit" volcanoes: The 2001–2006 eruptive period at Mt. Etna Earth-Science Reviews 138:371–393. https://doi.org/10.1016/j.earscirev.2014.06.009

Ginibre C, Wörner G, Kronz A (2007) Crystal zoning as an archive for magma evolution. Elements 3:261–266

Giordano D, Russell JK, Dingwell DB (2008) Viscosity of magmatic liquids: a model. Earth Planet Sci Lett 271:123–134

Gualda GAR, Ghiorso MS, Lemons RV, Carley TL (2012) Rhyolite-MELTS: a modified calibration of MELTS optimized for silica-rich, fluid-bearing magmatic systems. J Petrol 53:875–890. https://doi.org/10.1093/petrology/egr080

Halkidi M, Batistakis Y, Vazirgiannis M (2001) On clustering validation techniques. J Intell Inf Syst 17:107–145

Harris AJ, Dehn J, Calvari S (2007) Lava effusion rate definition and measurement: a review. Bull Volcanol 70:1–22

Higgins O, Sheldrake T, Caricchi L (2021) Quantitative chemical mapping of plagioclase as a tool for the interpretation of volcanic stratigraphy: an example from Saint Kitts. Lesser Antilles Bull Volcanol 83:51. https://doi.org/10.1007/s00445-021-01476-x

Higgins O, Sheldrake T, Caricchi L (2022) Machine learning thermobarometry and chemometry using amphibole and clinopyroxene: a window into the roots of an arc volcano (Mount Liamuiga, Saint Kitts). Contrib Mineral Petrol 177:1–22

Hirschmann MM, Ghiorso MS, Davis FA et al (2008) Library of experimental phase relations (LEPR): a database and web portal for experimental magmatic phase equilibria data. Geochem Geophys Geosyst 9. https://doi.org/10.1029/2007GC001894

Hotelling H (1933) Analysis of a complex of statistical variables into principal components. J Edu Psyc 24:417

Hsieh WW (2009) Machine learning methods in the environmental sciences: neural networks and kernels. Cambridge university press

Huber PJ (2004) Robust statistics, vol 523. John Wiley & Sons

Jerram DA, Martin VM (2008) Understanding crystal populations and their significance through the magma plumbing system. Geological Society, London, Special Publications 304:133–148. https://doi.org/10.1144/SP304.7

Jolliffe IT (ed) (2002) Choosing a Subset of Principal Components or Variables. In: Principal Component Analysis. Springer, New York, NY, pp 111–149

Jolliffe IT, Cadima J (2016) Principal component analysis: a review and recent developments. Philos Trans R Soc A Math Phys Eng Sci 374:20150202

Jorgenson C, Higgins O, Petrelli M et al (2022) A machine learning-based approach to clinopyroxene thermobarometry: model

optimization and distribution for use in earth sciences. J Geophys Res Solid Earth 127:e2021JB022904

La Spina G, Arzilli F, Llewellin EW et al (2021) Explosivity of basaltic lava fountains is controlled by magma rheology, ascent rate and outgassing. Earth Planet Sci Lett 553:116658. https://doi.org/10.1016/j.epsl.2020.116658

Leung IS (1974) Sector-zoned titanaugites: morphology, crystal chemistry, and growth. Am Mineral 59:127–138

Lierenfeld MB, Zhong X, Reusser E et al (2019) Species diffusion in clinopyroxene solid solution in the diopside–anorthite system. Contrib Mineral Petrol 174:46. https://doi.org/10.1007/s00410-019-1571-9

Mahalanobis PC (1936) On the generalized distance in statistics. National Institute of Science of India

Marchese F, Filizzola C, Lacava T et al (2021) Mt. Etna paroxysms of February–April 2021) monitored and quantified through a multi-platform satellite observing system. Remote Sens 13:3074

Martín-Fernández JA, Barceló-Vidal C, Pawlowsky-Glahn V (2000) Zero replacement in compositional data sets. In: Data analysis, classification, and related methods. Springer, pp 155–160

Métrich N, Allard P, Spilliaert N et al (2004) 2001 flank eruption of the alkali-and volatile-rich primitive basalt responsible for Mount Etna's evolution in the last three decades. Earth Planet Sci Lett 228:1–17

Mollo S, Putirka K, Misiti V et al (2013) A new test for equilibrium based on clinopyroxene–melt pairs: clues on the solidification temperatures of Etnean alkaline melts at post-eruptive conditions. Chem Geol 352:92–100

Mollo S, Giacomoni PP, Andronico D, Scarlato P (2015) Clinopyroxene and titanomagnetite cation redistributions at Mt. Etna volcano (Sicily, Italy): footprints of the final solidification history of lava fountains and lava flows. Chem Geol 406:45–54

Mollo S, Blundy J, Scarlato P et al (2018) An integrated PT-H2O-lattice strain model to quantify the role of clinopyroxene fractionation on REE+ Y and HFSE patterns of mafic alkaline magmas: application to eruptions at Mt. Etna Earth-Science Reviews 185:32–56

Mollo S, Pontesilli A, Moschini P et al (2022) Modeling the crystallization conditions of clinopyroxene crystals erupted during February–April 2021 lava fountains at Mt. Etna: Implications for the dynamic transfer of magmas. Lithos 420–421:106710. https://doi.org/10.1016/j.lithos.2022.106710

Morgavi D, Arienzo I, Montagna C et al (2017) Magma mixing: history and dynamics of an eruption trigger. Volcanic Unrest:123–137

Müller T, Dohmen R, Becker HW et al (2013) Fe–Mg interdiffusion rates in clinopyroxene: experimental data and implications for Fe–Mg exchange geothermometers. Contrib Mineral Petrol 166:1563–1576

Murru M, Montuori C, Wyss M, Privitera E (1999) The locations of magma chambers at Mt. Etna, Italy, mapped by b-values. Geophys Res Lett 26:2553–2556

Nazzareni S, Molin G, Peccerillo A, Zanazzi P (2001) Volcanological implications of crystal-chemical variations in clinopyroxenes from the Aeolian Arc, Southern Tyrrhenian Sea (Italy). Bull Volcanol 63:73–82

Nazzareni S, Morgavi D, Petrelli M et al (2022) Magmatic processes at Euganean Hills (Veneto Volcanic Province, Italy): clinopyroxene investigation to unravel magmatic interactions. Geosciences 12:108

Neri M, Acocella V, Behncke B et al (2011) Structural analysis of the eruptive fissures at Mount Etna (Italy). Ann Geophys 54:464–479. https://doi.org/10.4401/ag-5332

Parfitt EA, Wilson L (1994) The 1983–86 Pu'u 'O'o eruption of Kilauea Volcano, Hawaii: a study of dike geometry and eruption mechanisms for a long-lived eruption. J Volcanol Geotherm Res 59:179–205

Pawlowsky-Glahn V, Olea RA (2004) Geostatistical analysis of compositional data. Oxford University Press

Petrelli M, Caricchi L, Perugini D (2020) Machine learning thermobarometry: application to clinopyroxene-bearing magmas. Journal of Geophysical Research: Solid. Earth 125:e2020JB020130

Petrone CM, Bugatti G, Braschi E, Tommasini S (2016) Pre-eruptive magmatic processes re-timed using a non-isothermal approach to magma chamber dynamics. Nat Commun 7:12946. https://doi.org/10.1038/ncomms12946

Pichavant M, Di Carlo I, Le Gac Y et al (2009) Experimental constraints on the deep magma feeding system at Stromboli Volcano, Italy. J Petrol 50:601–624. https://doi.org/10.1093/petrology/egp014

Probst LC, Sheldrake TE, Gander MJ et al (2018) A cross correlation method for chemical profiles in minerals, with an application to zircons of the Kilgore Tuff (USA). Contrib Mineral Petrol 173:1–11

Putirka KD (2008) Thermometers and barometers for volcanic systems. Rev Mineral Geochem 69:61–120

Putirka KD, Johnson M, Kinzler R et al (1996) Thermobarometry of mafic igneous rocks based on clinopyroxene-liquid equilibria, 0–30 kbar. Contrib Mineral Petrol 123:92–108. https://doi.org/10.1007/s004100050145

R Core Team (2021) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria https://www.R-project.org/

Reimann C, Filzmoser P, Garrett R, Dutter R (2011) Statistical data analysis explained: applied environmental statistics with R. John Wiley & Sons

Rousseeuw PJ, Hubert M (2011) Robust statistics for outlier detection. Wiley Interdiscip Rev: Data Min And Knowl Discov 1:73–79

Ruprecht P, Bergantz GW, Cooper KM, Hildreth W (2012) The crustal magma storage system of Volcán Quizapu, Chile, and the effects of magma mixing on magma diversity. J Petrol 53:801–840

Shea T, Costa F, Krimer D, Hammer JE (2015) Accuracy of timescales retrieved from diffusion modeling in olivine: a 3D perspective†. Am Mineral 100:2026–2042. https://doi.org/10.2138/am-2015-5163

Sheldrake T, Higgins O (2021) Classification, segmentation and correlation of zoned minerals. Comput Geosci 156:104876. https://doi.org/10.1016/j.cageo.2021.104876

Simm J, De Abril IM, Sugiyama M (2014) Tree-based ensemble multi-task learning method for classification and regression. IEICE Trans Inf Syst 97:1677–1681

Spilliaert N, Allard P, Métrich N, Sobolev AV (2006) Melt inclusion record of the conditions of ascent, degassing, and extrusion of volatile-rich alkali basalt during the powerful 2002 flank eruption of Mount Etna (Italy). J Geophys Res Solid Earth 111:B04203. https://doi.org/10.1029/2005JB003934

Spina L, Cannata A, Morgavi D, Perugini D (2019) Degassing behaviour at basaltic volcanoes: New insights from experimental investigations of different conduit geometry and magma viscosity. Earth Sci Rev 192:317–336

Templ M, Filzmoser P, Reimann C (2008) Cluster analysis applied to regional geochemical data: Problems and possibilities. Appl Geochem 23:2198–2213. https://doi.org/10.1016/j.apgeochem.2008.03.004

Ubide T, Kamber BS (2018) Volcanic crystals as time capsules of eruption history. Nat Commun 9:1–12

Ubide T, Mollo S, Zhao J et al (2019) Sector-zoned clinopyroxene as a recorder of magma history, eruption triggers, and ascent rates. Geochim Cosmochim Acta 251:265–283. https://doi.org/10.1016/j.gca.2019.02.021

Unglert K, Radić V, Jellinek AM (2016) Principal component analysis vs. self-organizing maps combined with hierarchical clustering for pattern recognition in volcano seismic spectra. J Volcanol Geotherm Res 320:58–74

Vetere F, Iezzi G, Perugini D, Holtz F (2022) Rheological changes in melts and magmas induced by crystallization and strain rate. Comptes Rendus Géoscience 354:1–22. https://doi.org/10.5802/crgeos.125

Vona A, Romano C (2013) The effects of undercooling and deformation rates on the crystallization kinetics of Stromboli and Etna basalts. Contrib Mineral Petrol 166:491–509. https://doi.org/10.1007/s00410-013-0887-0

Ward JH Jr (1963) Hierarchical grouping to optimize an objective function. J Am Stat Assoc 58:236–244

Weber G, Caricchi L, Arce JL, Schmitt AK (2020) Determining the current size and state of subvolcanic magma reservoirs. Nat Commun 11:1–14

Winpenny B, Maclennan J (2011) A partial record of mixing of mantle melts preserved in Icelandic phenocrysts. J Petrol 52:1791–1812 10.1093/petrology/egr031

Zellmer GF (2021) Gaining acuity on crystal terminology in volcanic rocks. Bull Volcanol 83:1–8

Zellmer GF, Sakamoto N, Hwang S-L et al (2016) Inferring the effects of compositional boundary layers on crystal nucleation, growth textures, and mineral chemistry in natural volcanic tephras through submicron-resolution imaging. Front Earth Sci 4:88

Zhou X, Lu P, Zheng Z et al (2020) Accident prediction accuracy assessment for highway-rail grade crossings using random forest algorithm compared with decision tree. Reliab Eng Syst Saf 200:106931. https://doi.org/10.1016/j.ress.2020.106931