communications earth & environment

Article

A Nature Portfolio journal

https://doi.org/10.1038/s43247-025-02425-8

Understanding the evolution of scoria cone morphology using multivariate models

Gabor Kereszturi[1] ✉, Pablo Grosse[2,3], Melody Whitehead[1], Marie-Noëlle Guilbaud[4], Drew T. Downs[5], Rina Noguchi[6] & Matthieu Kervyn[7]

Scoria cones are the most abundant type of volcano in the Solar System. They occur in every tectonic setting and often overlap with human populations, yet our ability to provide complete geochronology within volcanic fields remains limited. Appropriate geochronology underpins the reconstruction of size-frequency distribution and is a key input for robust volcanic hazard assessment. Morphometric data have long been used to estimate relative ages of scoria cones; however, they have only shown promise at single volcanic fields and simple cones with homogenous pyroclastics. Here, we present a new global inventory of dated scoria cones ($n = 572$) from 71 volcanic fields formed under diverse magmatic, tectonic and climatic regimes, and build data-driven age models for dating scoria cones using easily accessible morphometric, reflectance and climatic variables. Our models suggest chemical composition of ascending magma may influence the initial scoria cone morphology which is then gradually modified by erosion over time.

Unlike "classical" geomorphology[1], volcano geomorphology results from the interplay between eruptive and erosional processes[2]. Volcano morphology is conceptually linked to time-dependent erosion[3], and hence provides first-order proxies for eruption age[4–7]. Scoria cones are small-volume volcanoes (<1 km^3) that are excellent natural laboratories to fingerprint erosional processes and for tracking the evolution of volcanic fields. The order in which scoria cones erupt (their geochronology) is important for reconstructing eruption size-frequency distributions[8], and for spatio-temporal volcanic hazard analysis[9], with fields typically hosting dozens to thousands of scoria cones. However, flank slope angle or cone height-to-width ratio, often used as measures of relative age, do not provide sufficient information necessary for developing unequivocal and globally valid age models (Fig. 1).

Scoria cone erosion has been numerically modelled as a slope-dependent diffusion process[10,11], which can yield accurate estimates of the degree of erosion for pyroclastic-dominated scoria cones. However, stratigraphic and sedimentological configurations of individual cones can widely differ due to welding/agglutination, vent migration, and eruption style changes[12–14]. These processes can influence initial cone morphology (and diversity of pyroclastic successions) and therefore, induce controls on the rates and styles of subsequent erosion. The unknown initial morphology of older cones remains a largely understudied issue in comparative volcano geomorphology. The only feasible approach to address such uncertainty is by sampling and analysing many scoria cones formed by diverse eruption styles and subsequent erosional histories across multiple volcanic fields.

We present new inherently multidimensional models for scoria cone erosion and ages via multivariate statistical approaches that integrate temporal-geomorphic-climatic relationships. Our models are based on a new global catalogue of dated scoria cones ($n = 572$) distributed across 71 volcanic fields (Fig. 1). The proposed models allow us to draw conclusions on both erosional and eruptive controls on scoria cone morphology, and to develop a globally valid morphometry-based dating tool.

Results

Scoria cones as "simple" volcanic edifices

Small-volume (<1 km^3), or monogenetic, volcanoes, forming through a series of eruptions over a short time (days to decades), are abundant in many geological and tectonic settings, including intraplate settings and convergent and divergent plate margins[15]. They typically erupt mafic magmas; however, examples of intermediate[16] and more evolved silicic magmas are also

[1]Volcanic Risk Solutions, School of Agriculture and Environment, Massey University, Palmerston North, New Zealand. [2]Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Buenos Aires, Argentina. [3]Fundación Miguel Lillo, Miguel Lillo 251, 4000 Tucumán, Argentina. [4]Departamento de Vulcanología, Instituto de Geofísica, Universidad Nacional Autónoma de México, México, D.F., México. [5]U.S. Geological Survey, Hawaiian Volcano Observatory, Hilo, HI, 96720, USA. [6]Faculty of Science, Niigata University, Nishi-ku, Niigata, 950-2181, Japan. [7]Department of Geography, Vrije Universiteit Brussel, Brussels, Belgium. ✉e-mail: G.Kereszturi@massey.ac.nz

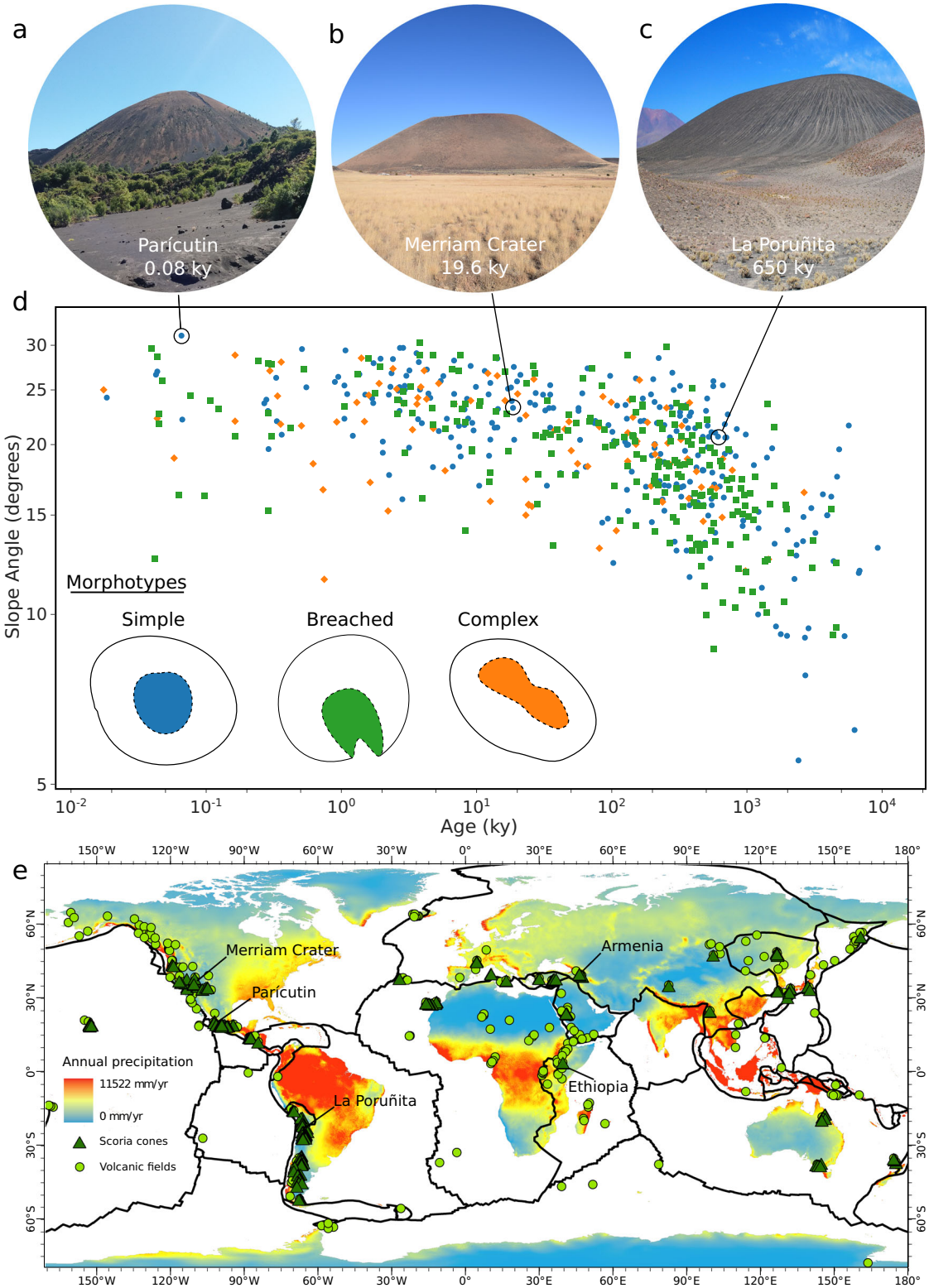

Fig. 1 | Scoria cone slope angle depicting a wide scatter over 10 million years and the distribution of scoria cones used in this study. a–c Examples of scoria cones of different ages. **d** Flank slope angles tend to decrease over time but show large scatter regardless of cone morphotypes, including a simple cone with one crater (blue), a simple cone with breached crater (green) and a complex cone with multiple craters (orange). **e** The distribution of volcanic fields globally based on the Smithsonian Institution's Global Volcanism Program[61] (green circles) and scoria cones analysed in this study (green triangles) shown on a basemap of annual precipitation from CHELSA[32] and plate boundaries[62]. Location of scoria cones shown in (**a–c**) are labelled along with volcanic fields from Armenia and Ethiopia, parametrised from SRTM Digital Terrain Models are also shown.

known[17]. Small-volume volcanoes can form a wide continuum of volcanic landforms, constructed dominantly by pyroclastic materials, including spatter ramparts, spatter cones, scoria cones, tuff rings, maar-diatremes and tuff cones[14,15], with an increasing influence of ground- and/or external water.

Scoria cones are conical-shaped volcanic constructs that are often conceptualized as morphologically and geologically "simple" landforms. However, scoria cones can have a variety of morphologies (e.g., simple, horseshoe-shaped with breached crater, elongated, and complex)[18], and form through a combination of magmatic eruption styles, including Hawaiian, Strombolian[14], violent Strombolian[19] and sub-Plinian[20]. In extreme cases, scoria cone forming eruptions can also include magma-water interaction driven phreatomagmatic phases[21–24]. The former magmatic eruption styles are heavily influenced by the coupling of volcanic gases within the ascending magma that fragments due to rapid near-surface expansion of the exsolved gases. Magmatic fragmentation styles result in near-vent accumulation of pyroclasts, typically consisting of sequences of poorly to moderately sorted scoriaceous deposits with median lapilli grain size[14,25]. The shape and grain size of the resultant volcanic deposits are closely linked to magma explosivity[26,27], with more intense fragmentation creating much finer deposits. The pyroclastic fragments are often transported via ballistic ejection[28] and fallout from eruption clouds[29,30], where they can generate dry grain avalanches after landing on the outer flanks of the growing scoria cone. Pyroclasts must be solidified and small to maintain a granular behaviour after landing, otherwise, deposits can show local (m-scale) to extensive (e.g., entire flanks of the cones) welding and agglutination due to high ejection rates and inadequate cooling before landing[14,31]. The competition of such a diverse range of eruption styles can create unique volcanic stratigraphy with sedimentologically variable deposits, and consequently diverse cone morphologies.

To reflect the diverse evolutions and eruption styles of scoria cones, this work defines scoria cones in a broad sense, as any volcanic landform that formed through dominantly magmatic eruption styles (>90%), has a broadly conical edifice with or without a summit crater, and is made of predominantly scoriaceous deposits. We prefer this definition to reduce the subjectiveness around defining scoria cones purely on a sedimentological basis (i.e., cones made of scoriaceous deposits). This fresh perspective can highlight the eruptive and geomorphic diversity of scoria cones, but it is still clear enough to exclude analysis of tuff cones and any phreatomagmatic volcanoes (e.g., formed dominantly through magma-water interactions) that often co-occur with scoria cones in volcanic fields.

Construction of a global database

There is a lack of empirical models that capture scoria cone morphological changes on timescales from years to millions of years, regardless of compositional, volcanological and tectonic settings. Here, we address this by analysing 12 m spatial resolution Digital Terrain Model (DTM, including WorldDEM, locally available Light Intensity and Ranging and photogrammetry-based DTMs) derived morphometric parameters ($n = 34$), surface reflectance from Sentinel-2 satellites ($n = 11$), and climatic variables from CHELSA[32] (Climatologies at High resolution for the Earth's Land Surface Areas) and Interpolated Global Rainfall Erosivity[33] databases ($n = 20$) (Supplementary Table 1), as a function of available absolute ages from 572 scoria cones. The morphometric, reflectance and climatic datasets are variables chosen to capture the geomorphology (via DTMs), surface cover type (e.g., vegetation cover and lithology) and climatic influences (via global climate models), respectively. The climatic variables represent present-day values; however, we anticipate that they are still important to discriminate erosion rates and styles of surface processes even for cones older than Holocene[34]. These variables are assumed in this study to change directly with age (i.e., degradation/erosion) of the edifice, and hence they can inform our age prediction models.

The target variables include the eruption age of scoria cones and their age uncertainties (i.e., upper and lower bounds of age determinations). The range of input target variables spans 0.018 to 9950 ka, dated using a range of methods (e.g., $^{40}Ar/^{39}Ar$, $^{40}K/^{40}Ar$, U-Th/He, thermoluminescence, optically stimulated luminescence (OLS), ^{10}Be, ^{21}Ne, ^{36}Cl, ^{3}He, and ^{14}C techniques, and eyewitness accounts (Supplementary Data 1)). The compiled absolute ages show a bimodal distribution with peaks centred at 20–30 ka and >200–300 ka (Fig. 2). Most ages included in the database have supporting ancillary information in the source publications (e.g., matching stratigraphic observations), confirming the geological ages. To eliminate uncertainty of absolute ages, we have further filtered the database to only include scoria cones that have ≤25% age errors.

The cones span 71 volcanic fields covering a large range of climatic and tectonic environments, although cones from tropical regions are under-represented due to an overall scarcity of geochronological studies in the volcanic fields of these regions (Fig. 1). We used multiple linear regression (MLR), partial least squares regression (PLSR), and support vector regression (SVR) to develop cone age prediction models (target variable) using morphometric, reflectance and climatic variables as predictor variables. See "Methods" for a detailed analysis and workflow.

Modelling geomorphic age

Integrating morphometric, reflectance and climatic data provides a first-order model for expressing geomorphic ages (Fig. 2). Age models perform evenly after independent validation across multiple methods (Supplementary Table 2), with an increasing accuracy when morphology, reflectance and climatic predictors are combined up to R^2_{val} of 0.73 and mean absolute percentage error (MAPE) of 47.5%. Models using only cone morphology as inputs perform at moderate levels (R^2_{val} of 0.54–0.58; MAPE of 58.9–63.5%), and overpredict ages on Holocene cones (Supplementary Fig. 1). We attribute this mismatch to initial eruptive morphometric variability associated with different eruption styles and resultant pyroclastics[12]. Such bias mostly disappears on age models using a combination of morphometric and reflectance data, suggesting rapid surface modifications on scoria cones during the first 10 ky is well captured by multispectral satellite images. The models built here assume that the scoria cone population age is unimodal, and we are not aware of any geological rationale/evidence for any temporal global dip(s) in volcanism. However, the scoria cone sample used to train and test the models shows bimodality with a dip at ~50–80 ka (Fig. 2). This likely reflects the uneven sampling of the previously published ages due to limitations of absolute dating methods. For example, the upper limit of ^{14}C dating is around 50 ka, while $^{40}Ar/^{39}Ar$ can return reliable absolute ages on scoria cones that are >50 ka. The temporal bias in the geochronology is also combined with spatial bias due to the insufficient number of scoria cones dated in humid and tropical volcanic fields (e.g., Fig. 1). The inherent bias requires future attention to improve the database fitness for dating using morphology.

We further attribute the moderate performance of the models measured by absolute error metrics, such as R^2_{val} and $MAPE_{val}$ (Supplementary Table 2), to the non-linearity of erosional processes[35]. Such non-linearity agrees with either expedited[36] or delayed[37] erosional response of pyroclastics immediately after eruption due to either their unconsolidated nature or surface permeability, respectively. In log–log space, however, we find statistically significant rank-correlation (Spearman $R > 0.85$, $p < 0.05$) between measured absolute ages and predicted ages. Hence, our age prediction models reflect the order of eruptions, providing a first-order geochronology across volcanic fields.

To test the robustness of the model, we applied our training model to a continental USA subset ($n = 72$) and two volcanic fields from Armenia ($n = 14$) and Ethiopia ($n = 3$) in which WorldDEM data are not available; instead, we parametrised the cones on 30 m Shuttle Radar Topography Mission (SRTM) DTMs. Predictions perform well with Spearman R of 0.67–0.73 ($p \approx 0$), depending on the model (Supplementary Table 2). The results show that observed relationships scale to coarser resolution input DTMs and verify our new tool for providing the order of eruptions using globally available SRTM data to improve the geochronology of poorly dated volcanic fields.

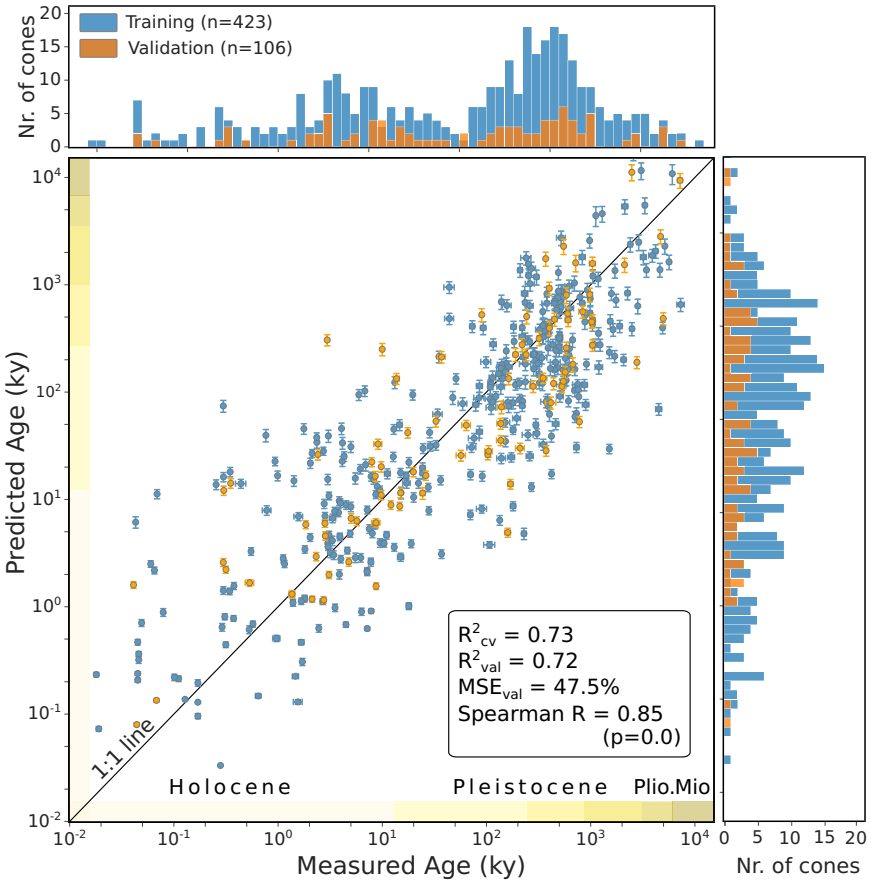

Fig. 2 | Prediction models using morphometric, reflectance, and climate input variables show good agreement with measured absolute ages of scoria cones from the Holocene to the Upper Miocene. The model is derived using Partial Least Squares Regression (PLSR) using morphometric, reflectance, and climatic variables ($n = 65$). The horizontal histogram shows the age distribution of training and validation, while the vertical histogram shows the predicted age for those two populations. Colours correspond to training (blue) and validation populations (orange). Horizontal error bars are one standard deviation of the absolute age, whereas the vertical error bars from PLSR are a prediction of the upper and lower bounds of age errors, which are often smaller than the symbols due to the log–log scale.

Discussion

Factors influencing geomorphic age

Spearman R analysis and Variable of Importance (VIP) calculated from PLSR confirm the age sensitivity of some previously applied morphometric variables, including cone height-to-width ratio (HcoWcoMean) and flank slope angle (FlSlopeMean), and identify additional age-sensitive variables, including crater infill volume (CrVol), crater slope angle (CrSlopeMean), flank irregularity (CrIrrMean), and reflectance at 704, 1613, and 2202 nm (Fig. 3). Results indicate that outer flank and crater morphology, including slope angle and its irregularity, are more sensitive morphometric variables than cone height-to-width ratio (e.g., HcoWcoMean). Our study generalises such relationships, showing that they occur under contrasting tectonic, magmatic, climatic, and associated erosional regimes. Due to the multivariate nature of the analysis, our models can better capture erosional trends through flank and crater variables prevalent under a large range of geologic and climatic regimes.

Both Spearman R and PLSR-VIP scores indicate that Sentinel-2 Band 5 (704 nm), Band 11 (1613 nm), and Band 12 (2202 nm) correlate well and positively with absolute age (Fig. 3), which has never been described before. We interpret this as a signal of vegetation establishment together with soil formation, captured by the increased reflectance at Band 5 and light absorption by hydrated minerals at Bands 11–12[38]. Removal of unconsolidated pyroclastics due to rapid initial erosion can expose the often oxidised and variously welded/agglutinated interior cone deposits[39], which can increase the surface reflectance. Change in the otherwise spectrally "flat" basaltic to andesitic scoria of pristine cones (e.g., Parícutin and Alumbrera on Fig. 3) can improve all age prediction models (Supplementary Table 2). This suggests that reflectance data can capture the early steps of erosion, whereas morphology-related variables are sensitive to erosional modifications on timescales of thousands to millions of years (e.g., reduce model bias for cones older than the Holocene). Furthermore, tephra blanketing, tectonic activity (e.g., faulting), and anthropogenic modification can generate morphologic/morphometric modifications that may translate to 'noise' or uncertainty in the prediction models (i.e., it can decrease the obtained correlations), producing over- or underpredictions.

Climate is an important factor for scoria cone degradation[34], and accordingly, our models show improvements mostly by including diurnal temperature ranges and annual precipitation variables. Climatic variables are inherently "regional" in scale (e.g., many cones located within the same volcanic field share similar values), while morphometric and reflectance-derived variables change cone-to-cone, better explaining underlying variability, as indicated by the lack of importance of most climatic variables in Spearman correlation analysis (Fig. 3). Indirectly, climatic variables can contribute to models by adding proxies for weathering and erosion style. For example, under colder and more arid climates, physical weathering and seasonally heterogeneous sediment transport dominate, including wind-erosion[36], freeze-thaw cycles[40], and debris flows[39], while tropical and humid climates promote chemical weathering and subsequent surface wash, rainfall-driven fluvial erosion[41], and cation leaching[42]. Given these are "regional" variables, they cannot capture intra-field variations but are useful for discriminating short- and long-term erosional trends among volcanic fields.

When erosion kicks in

Volcano morphology at a given time is a function of eruptive heritage[43] and erosional processes[13]. First, the cone shape reflects eruptive style (eruption-controlled regime), but then, as time passes, it increasingly reflects the post-eruptive erosional processes (erosion-controlled regime). Our understanding of timescales during this transition is fragmented due to insufficient geochronological data and limited physical volcanology data due to the lack of exposure and high weathering rates.

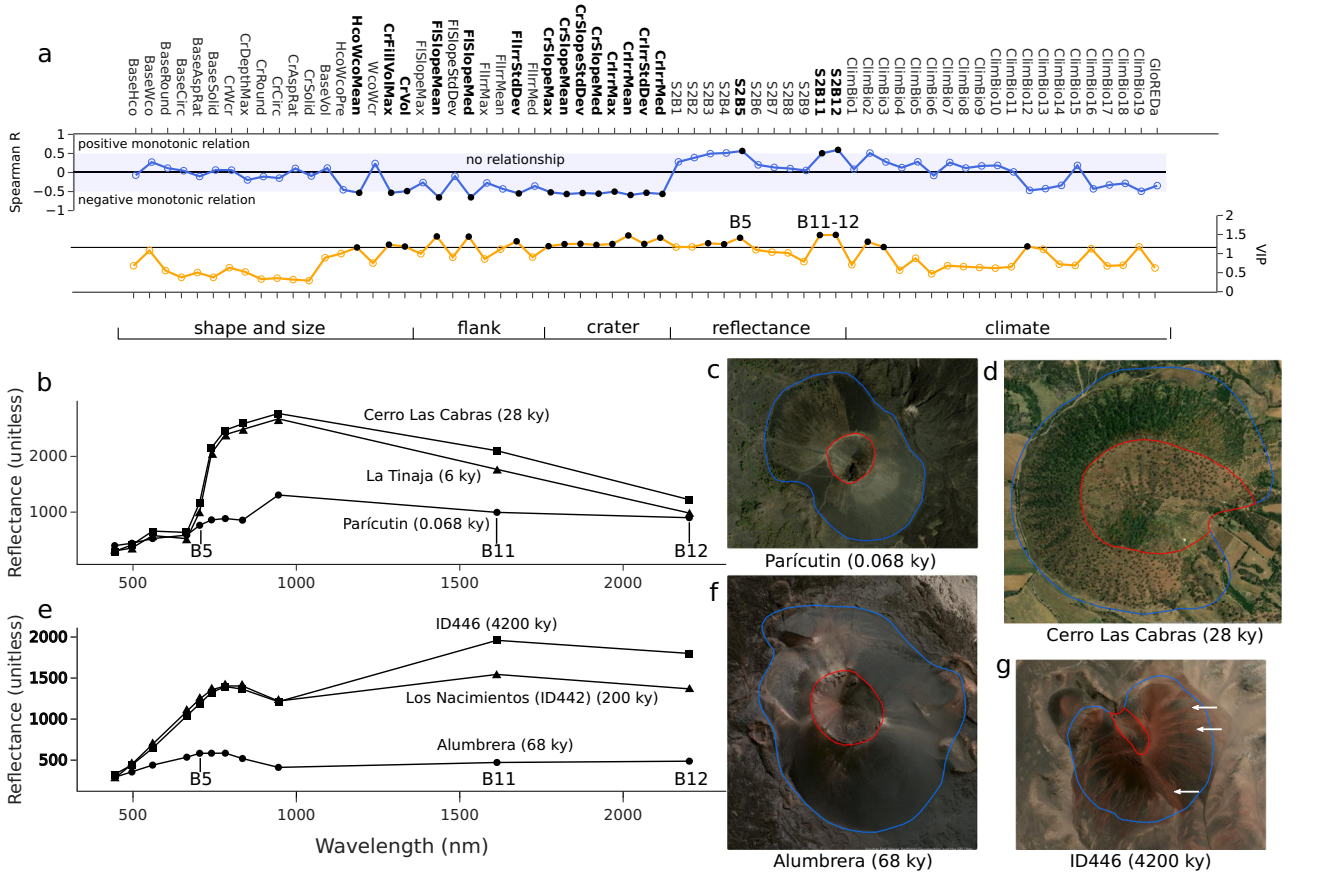

Fig. 3 | Factors influencing scoria cone erosion. a Plot showing Spearman R rank correlation coefficients (blue) and Variable of Importance (VIP; orange) for input variables (defined in Supplementary Table 1). Important variables are shown as solid black circles and bold text on graph. **a** Strong monotonic correlation is present when Spearman R is either larger or smaller than 0.5 or −0.5. The light blue area indicates no relationship on the Spearman R plot. The black line on the VIP plot shows the threshold of importance (VIP = 1.2)[63]. **b**–**g** Observed importance of Sentinel-2 bands 5, 11, and 12 coincides with reflectance changes on scoria cones due to revegetation over time (**b**–**d**), and deposit oxidation state (**e**–**g**). True colour Sentinel-2 satellite images are provided as examples. White arrows on g indicate oxidised outcrops and welded/agglutinated beds.

We hypothesize that eruptive processes should be reflected in at least some morphological parameters for young cones; this information will then be variably modified by erosional processes under a given climate. To test this hypothesis, we employed the non-parametric Mann–Whitney *U* test on PLSR-based age predictions, based on morphometric variables only. Using predictions from threefold cross-validation, we measured whether under- and overpredicted cones have morphometric attributes that are statistically different (Fig. 4 and Supplementary Data 2). Iteratively, we calculated Mann–Whitney *U* test statistics between 0 and 500 ky with a cumulative increment of 1000 years. We assume the first 500 ky period has the highest chance of witnessing any transition from initial eruptive- to more erosion-dominated morphology[13]. For a given morphometric variable, if the Mann–Whitney *U* test *p*-values are substantially >0.05, a null hypothesis holds, the under- and overpredicted cones share similar morphologies, whereas *p*-values of ≤0.05 mean the populations are statistically significantly different.

Our calculations show that morphometric variables are selectively impacted by both eruptive and erosional processes. However, the most important morphometric variables, identified via PLSR-VIP, converge at zero *p*-values (Fig. 4), meaning cones come from populations that are morphometrically different. We find that cone height-to-width ratio (HcoWcoMean) and median flank slopes (FlSlopeMed) are statistically different between under- and overpredicted cones immediately after cone formation, and become gradually similar over time, reaching statistical similarity at 115 ky and 209 ky, respectively (95% confidence interval; Fig. 4). In contrast, crater depth (CrFillVolMax), crater slope angle (CrSlopeMed), and crater volume (CrVol) show initial morphological similarities due to eruptive processes (e.g., similarity in crater excavation mechanism) that transitions into a period of morphological dissimilarity after ~10 ky and then similarity again after 200 ky. This is consistent with field observations, such as removal of unconsolidated tephra, exposure of more erosion-resistant pyroclastics[39], and accumulation of aeolian deposits[44]. Erosional morphologies after 200 ky become more alike due to a combination of overland flow and dry ravel sediment transport processes.

Compositional forcing on volcano morphology?

When models based only on morphometric variables are validated using leave-one-group-out cross-validation, age predictions show considerable variability between volcanic fields and volcanic regions (Supplementary Table 3). Age predictions for scoria cones formed along subduction-related systems (e.g., volcanic fields in the continental USA, Mexico, Turkey, and New Zealand) always outperform intracontinental settings (e.g., Harrat Rahat in Saudi Arabia; Wudalianchi, Kaluo, and Tangchong in China; Newer and McBride Volcanic Provinces in Australia) and volcanic islands (e.g., Hawaii, Canary Islands and Azores). Given statistically significant morphological differences among cones younger than 100 ky (Fig. 4), erosion cannot explain the observed departures from the "expected" geomorphic age (that is the 1:1 line in Fig. 5), suggesting rather deeper controls on the cone morphologies (e.g., magmatic, crustal, or tectonic processes).

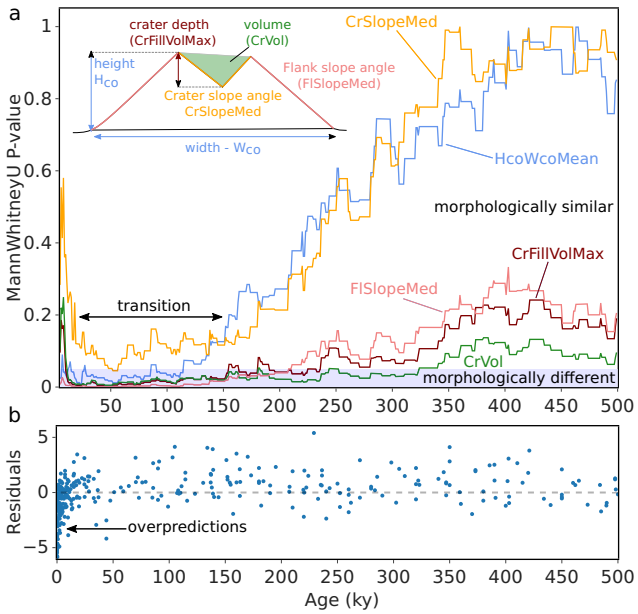

Fig. 4 | The initial eruptive morphology gradually transitions into erosional morphology in the first 100–200 ky. a Observed timescales of erosional modification over the first 500 ky using a partial least squares regression (PLSR) derived age model. The PLSR model was developed on morphometric data only. The *y*-axis shows the distribution of *p*-values from the Mann–Whitney *U* test, the blue zone shows time-steps in which the over- and underpredicted cones are morphologically different at the 95% confidence interval. Transition from eruption-dominated morphologies is observed in the first 100 ky, then erosional processes gradually make scoria cones more morphologically similar. The inset sketch shows the location of morphometric data within a 2D cross-section of a scoria cone. **b** Graph showing the residuals in prediction ages for the first 500 ky. Note the overprediction (older than measured) of ages for Holocene cones.

We hypothesize that magma composition and degassing history can be important pre-conditions (Fig. 5), driving bubble and crystal nucleation and growth, and strongly guide sub-Plinian[20], violent Strombolian[19], Strombolian, and Hawaiian eruption styles[45]. This is further supported by the moderate correlation of age prediction residuals with whole-rock K_2O content across 13 volcanic fields from the western and southwestern USA (Fig. 5 and Supplementary Table 4). While this apparent correlation requires further studies, we putatively interpret it as a combination of varying partial mantle melting in the asthenosphere[46], and/or decompression of magmas en-route to surface[47,48]. In the latter case, exsolution of magmatic volatiles (e.g., H_2O and CO_2) can induce deep-seated degassing, usually accompanied by fractional crystallization[47]. Such volatile-poor magmas might either stall in mid-crustal magma reservoirs[49] or ascend to the surface at slower rates compared to volatile-rich melts[47]. Contrary to this, volatile-rich mafic magmas are generally less viscous and tend to ascend to the surface faster due to increased buoyancy and overpressure[50], following closed-system degassing[51].

Deep-seated controls on magma ascent and degassing are suggested to impact eruption dynamics. For violent Strombolian eruptions, gas and melt are well coupled, whereas lower intensity Strombolian eruptions are related to magmas with decoupled bubbles that move as a slug flow in a cylindrical conduit[26]. This inherently influences fragmentation style and the thermal history of pyroclastics[14]. Violent Strombolian eruptions can produce smaller grain sizes and particles that are sufficiently cool to behave as a 'dry' granular media, whereas lower intensity Strombolian eruptions can eject coarser and still-molten particles that can be subjected to intensive agglutination and welding, predominantly within the crater and its rim[14]. Contrasting eruption styles and resulting deposits in terms of grain size and susceptibility for welding/agglutination can explain volcano morphologies well. Overpredicted cones tend to have lower median flank slope angles by 1.7° at 10 ky time ($p = 0.01$), whereas underpredicted cones with steeper slopes can be explained by interlocking coarser particles that can be further impacted by both agglutination and welding. Hence, eruption styles driven by compositional heterogeneity and degassing history appear to correlate with the observed variations in erosionally intact volcano morphologies.

Our results suggest deeper controls on primary scoria cone morphology than have previously been proposed. Such heterogeneity can induce diverse degradation paths in balance with physical and chemical weathering and erosion of pyroclastic successions. Additional work that links volcano morphology with volatile budget, geochemistry and eruption rates, especially to constrain magma ascent and crustal degassing processes, is needed.

Broader implications

The continental USA alone hosts at least 2229 Quaternary small-volume volcanoes (most of which are scoria cones); one-quarter of them have absolute dates, of which 15% are by ^{14}C, Ar-Ar, or surface-exposure techniques[52]. A more pronounced lack of geochronology of volcanic fields is found worldwide; for example, less than 1% of the total scoria cone population of Africa has absolute ages[53]. The lack of absolute ages often reflects the presence of cones that are too young for Ar–Ar dating, scarce organic material for ^{14}C dating, limited resources for dating, or security issues that severely restrict field sampling. The limited number of absolute ages from humid and tropical regions implies that the presented age models should be employed with care in such cases and should also be a call for new geochronological research, targeting volcanic fields in these regions (e.g., Cameroon, Colombia, and Indonesia).

Our globally trained models are well-positioned to improve the geochronology of volcanic fields in which scoria cones dominate. This can allow for first-order spatio-temporal patterns of volcanism, and potentially provide vital age inputs for probabilistic volcanic hazard assessment[54]. Our models can be used in combination with existing ages from stratigraphy, exposure dating, lake sedimentation and crystallization (e.g., $^{40}Ar/^{39}Ar$ and $^{40}K/^{40}Ar$), to augment temporal models for volcanic fields using probabilistic approaches, as has been employed for the Auckland volcanic field, New Zealand[8]. By integrating morphometric ages with existing datasets in a probabilistic framework, we can overcome the current limitations of dating within volcanic fields.

The ability to constrain scoria cone ages on a global scale opens opportunities to improve our understanding of the volume-frequency and geochemical evolution of distributed volcanic fields to better constrain magma generation and transfer through the crust. Studies that integrate this information with physical volcanology data for a greater understanding of the most common volcano type on Earth are further encouraged, which will aid to fingerprint sediment transport and landscape evolution[10,11], and probe the origins of extraterrestrial volcanism through morphology[55].

Methods

Input data

In total, 572 scoria cones from 71 volcanic fields were selected based on available absolute ages and then parameterized using Digital Terrain Models (DTM), satellite-based reflectance, and globally available climatic models (Supplementary Table 1). Regression models were built with cone ages and their upper and lower bounds (i.e., age uncertainties) as target variables, and with morphometric, reflectance, and climatic attributes used as predictor variables.

Morphometric variables ($n = 34$) were parameterized using DTMs. Basal and crater outlines of each studied cone were manually drawn using DTM-derived slope angle maps, hillshade images, and high-resolution satellite imagery. In total, 34 morphometric attributes (Supplementary Table 1) were estimated following standard methods[43,56,57], and computed from the 12-m resolution WorldDEM, and higher resolution Light Intensity

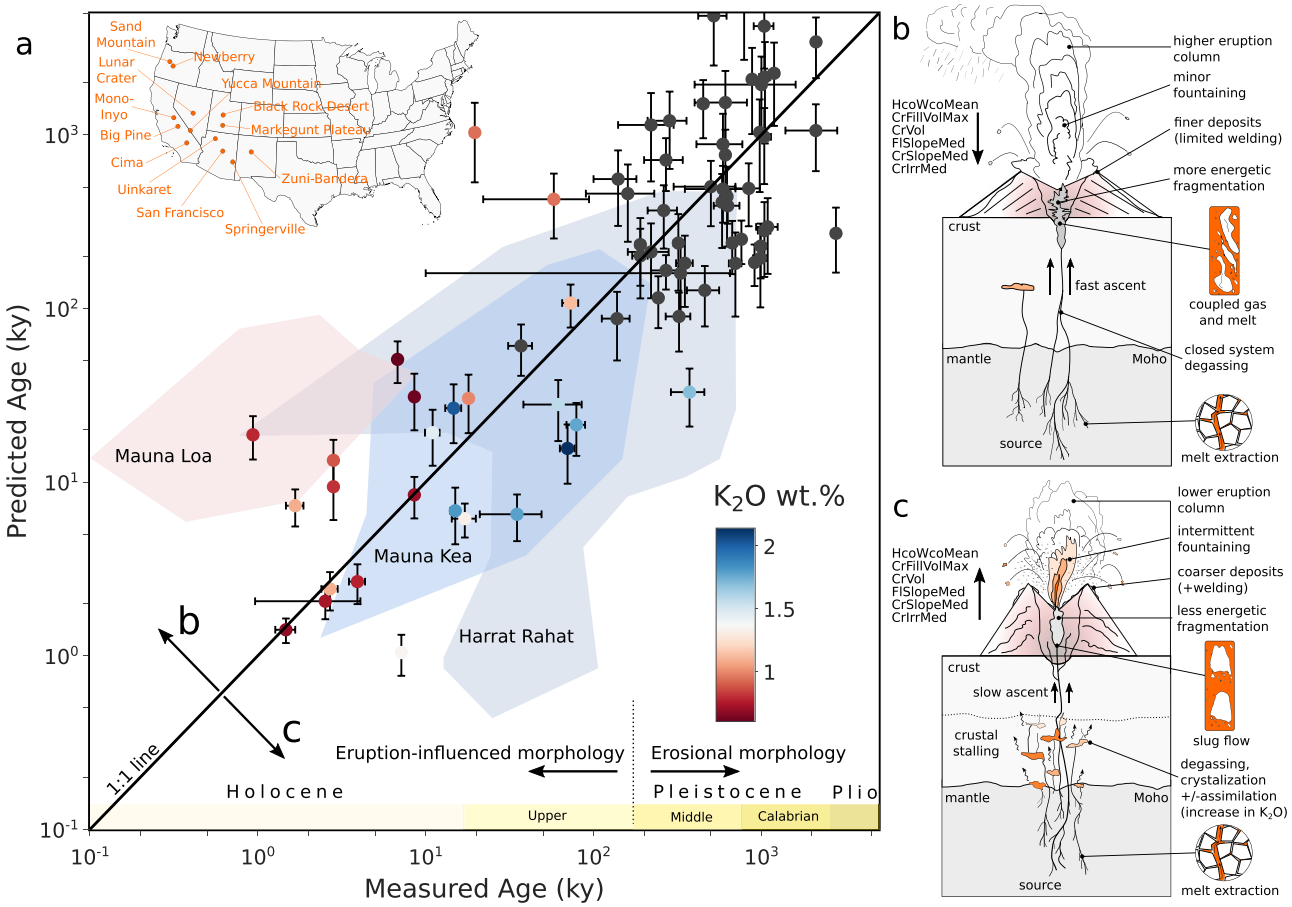

Fig. 5 | Eruption style controls on scoria cone morphology. a Absolute (*x*-axis) and predicted ages (*y*-axis) for scoria cones of 13 basaltic volcanic fields across the western and southwestern USA. Cones are colour-coded by K_2O content; black dots indicate the lack of geochemical data. Prominent over- and underpredictions moderately correlate with primary melt K_2O content. The coloured fields in (**a**) are for Mauna Loa (erupts lower K_2O, volatile-poor tholeiitic lavas), Mauna Kea (erupts mafic alkalic lavas with higher K_2O)[64], and Harrat Rahat (impacted by mid-crustal storage and degassing)[65], which show similar geochemical and morphological trends, supporting observations within the continental USA. The error bars show one standard deviation for both the measured and predicted ages. **b** Cartoon showing eruption scenario of volatile-rich magma (low K_2O) that might erupt explosively from gas-coupled melts (i.e., churn flow). The resultant cone morphology is dominated by a lower cone height-to-width ratio, slope angle, crater volume and depth, variables that are less susceptible to welding and agglutination. **c** Alternative eruption scenario where a volatile-poor magma (high K_2O) ascending as a slug-flow feeds eruptions dominated by lower fragmentation efficiency and welding/agglutination, forming cones with higher height-to-width ratios, steeper flank slope angles, and larger and deeper crater morphologies.

and Ranging (LiDAR) and photogrammetry-based DTMs. All DTMs were reprojected to the universal transverse mercator (UTM) coordinate system and resampled using the bilinear method to 12 m, an appropriate spatial resolution for analysing small-volume volcanoes[58].

Reflectance variables ($n = 11$) were extracted from the European Space Agency's Sentinel-2 satellite images with spatial resolutions of 10, 20 and 60 m, depending on spectral bands, and bandwidths between 15 nm and 175 nm, assessed via the Google Earth Engine (Supplementary Table 1). The database contains median atmospherically corrected reflectance from visible shortwave to infrared within each scoria cone outline. The median value per cone was extracted from a median reduced time series of Sentinel-2 images between 2019 and 2022 with cloud cover of ≤5%. We only considered images from the driest and warmest months of the year at each location to avoid potential snow cover.

Climatic variables ($n = 20$) were extracted from CHELSA Version 1.2, which are a 30 arcsecond (~1 km) resolution monthly and annual mean temperature and precipitation raster layers collected between 1979 and 2013[32] (Supplementary Table 1). The CHELSA-derived variables were complemented by interpolated global rainfall erosivity[33].

Target variables include the eruption age of scoria cones and their age uncertainties (i.e., upper and lower bounds of age determinations). The range of input target variables spans 0.018 to 9950 ka, dated using $^{40}Ar/^{39}Ar$, $^{40}K/^{40}Ar$, and U-Th/He, thermoluminescence, optically stimulated luminescence (OLS), ^{10}Be, ^{21}Ne, ^{36}Cl, ^{3}He, and ^{14}C techniques, and eyewitness accounts (Supplementary Data 1).

Age models

Predictor variables (morphometry, reflectance, climate) and target variables (age, including age uncertainty) were analysed through multivariable statistical analysis. We built age models using multiple linear regression (MLR), partial least squares regression (PLSR) and support vector regression (SVR). Before scaling and centring the data, we age-filtered the scoria cone catalogue ($n = 572$), removing cones with >50% uncertainty in the absolute ages (excluded data; $n = 42$). The filtered population was split into a training ($n = 423$; 80%) and an independent validation dataset ($n = 106$; 20%) using random sampling. Predictor variables were log10 transformed to normalize input data distributions and centred and scaled to unit variance before analysis, whereas target variables were only log10 transformed. Each regression model was validated using both k-fold cross-validation ($k = 3$; indexed as "cv") and independent validation datasets (indexed as "val"), and reported as R^2, mean absolute percentage error (MAPE) and Spearman's R coefficient. By changing the predictor variable groups (e.g., morphometry only, morphometric + reflectance data, and morphometric + reflectance data + climatic data), we built three models with each regression algorithm,

giving nine models in total. Furthermore, we tested age prediction models using coarser resolution Shuttle Radar Topography Mission (SRTM) DTM data (version GL1 from OpenTopography), which has a near-global coverage at 30 m spatial resolution. We resampled using bilinear resampling and reprojected SRTM data across the continental USA ($n = 72$) and two volcanic regions (not included in training models) from Armenia ($n = 14$) and Ethiopia ($n = 3$) with reliable Ar–Ar absolute ages.

Multiple linear regression (MLR) models the relationship between an array of predictors **X** and target variables **Y** using the least squares method:

$$\boldsymbol{Y} = \boldsymbol{X}\beta + \varepsilon \tag{1}$$

where β is a vector of regression coefficient and ε is the model error. Regression coefficients are calculated to minimize the residual sum of squares SS_{Res}, between observed $\boldsymbol{Y}$ and predicted $\hat{\boldsymbol{Y}}$ target variables:

$$SS_{Res} = \frac{1}{N}||\boldsymbol{Y} - \hat{\boldsymbol{Y}}||^2 \tag{2}$$

MLR is straightforward to implement in the scikit-learn library in Python, and no (subjective) input parameters are needed. Thus, it is used here to produce a simple baseline model. However, MLR requires the mean response (target variable) to be linear in the parameters, such that each regression coefficient can be estimated independently. Additionally, any multicollinearity (where predictors are highly correlated with each other) can cause unstable estimates of regression coefficients as the same variance space in the target variable is taken up by multiple predictors.

Partial least squares regression (PLSR)[59] handles multicollinearity using an iterative procedure that effectively removes the influence of predictors that are causing more noise than information. Predictor and target variables are decomposed into a set of orthogonal scores ($\boldsymbol{T}$, $\boldsymbol{U}$), and loadings ($\boldsymbol{P}$, $\boldsymbol{Q}$) as $\boldsymbol{X} = \boldsymbol{T}\boldsymbol{P}^{\boldsymbol{T}}$, and $\boldsymbol{Y} = \boldsymbol{U}\boldsymbol{Q}^{\boldsymbol{T}}$, respectively. Scores and loadings are then estimated through simple regressions iteratively (starting with each x on $\boldsymbol{Y}$), until values are reached that maximize the covariance structure between $\boldsymbol{X}$ and $\boldsymbol{Y}$. The final regression coefficients β are then found with the regression model $\boldsymbol{U} = \beta\,\boldsymbol{T}$, and target variables $\boldsymbol{Y}$ are then estimated with:

$$\hat{\boldsymbol{Y}} = \boldsymbol{X}\boldsymbol{P}\beta\boldsymbol{Q}^{T} \tag{3}$$

PLSR was implemented in Python using the scikit-learn library (PLS2 algorithm), and the optimal number of PLS components was determined using R^2 values and threefold cross-validation.

Support Vector Regression (SVR)[60] approximates an underlying relationship by finding a function $f(\boldsymbol{X})$ that describes **Y** within a tolerance, ϵ, (that controls model complexity and fit), in a n-dimensional space, where p is the number of predictor variables. SVR derives a function that maximizes the margin while minimizing prediction error as:

$$\min \frac{1}{2}||w||^2 + C\sum_{i=1}^{p}\left(\xi_i^- + \xi_i^+\right) \tag{4}$$

where, $\frac{1}{2}||w||^2$ represents the squared norm of the weight vectors, C is the regularization parameter (user-defined), and ξ represents the deviation above (ξ^+), and below (ξ^-) the predicted values in relation to ϵ. Weight vector w defines the magnitude and direction of influence of each predictor variable on the predictions. C is chosen to regulate model fit, with high values in favour of minimizing the error (i.e., overfit), and low values allowing for a larger margin (i.e., underfit). To predict target variables, SVR uses a generalized formula:

$$y_i = w^T \varnothing\left(x_i\right) + b \tag{5}$$

which is the dot product of the weight vector w and $\varnothing(x_i)$ transformed predictor variable for observation (cone) i, and b is the model bias. To handle the non-linear relationship between $\boldsymbol{X}$ and $\boldsymbol{Y}$, we used a radial basis function (or Gaussian) kernel. This projects the input predictor variables into n-dimension space, in which a linear function can be found. The width of the kernel, γ, defines the influence of individual data points (i.e., small γ means narrower Gaussian functions and shorter influence distance). We used the grid search method to select optimal parameters ϵ, C and γ using the training data with 3-fold cross validation. We used the SVR implemented in the scikit-learn library in Python.

Variable Importance

To analyse the contribution of each predictor variable to the model, we conducted two tests: Spearman's rank correlation and PLSR's variable of importance in projection (VIP) scores. Spearman's rank correlation coefficient, ρ, is a non-parametric method (i.e., no assumption for normality of input data) to assess monotonic relationships between two variables:

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)} \tag{6}$$

where d_i is the rank difference between observation pairs, and n is the number of observations (cones). ρ ranges between -1 and 1, with 0 representing no correlation, and -1 and 1 representing perfect negative and positive monotonic correlations, respectively. This was implemented through the SciPy library in Python.

PLSR automatically ranks input predictors by their contributions to modelled target variables $\boldsymbol{Y}$. This can be expressed using VIP. For predictor variable j, VIP_j can be calculated as:

$$VIP_j = \sqrt{\frac{p\sum_{k=1}^{h}\left(SSY_k\, w_{kj}^2\right)}{h\, SSY_{Total}}} \tag{7}$$

where p is the number of predictor variables, h is the number of PLS components, SSY_k is the sum of squares for the kth PLS component, and w_{kj} is the weight of X_j in the kth PLS component.

Statistical inference

We used the Mann–Whitney U test to determine any statistical differences between positive (underpredicted age or younger than actual) and negative residuals (overpredicted age or older than actual) using the model developed purely on morphometric data. The Mann–Whitney U test is a non-parametric test assessing whether two populations are from the same underlying distributions (the null hypothesis) and is formulated as:

$$U_1 = n_1 n_2 + \frac{n_1(n_1+1)}{2} - R_1 \tag{8a}$$

and

$$U_2 = n_1 n_2 + \frac{n_2(n_2+1)}{2} - R_2 \tag{8b}$$

where, R is the sum of ranks for each group, and the test statistic is $U = \min(U_1, U_2)$. This value is then used to calculate p-values using the 'exact' method in SciPy. This test was implemented through the SciPy library in Python with a statistical significance level set at $\alpha = 0.05$.

Besides predicting ages, we assessed the model's performance by leave-one-group-out cross-validation, where we train a model without the hold-out group and then apply the model to the hold-out dataset. This approach can shed light on inter-regional differences among volcanic fields, including geochemical differences. Geochemical data were compiled from the literature and included as an arithmetic mean if multiple samples were available. The source reference is included in Supplementary Table 4.

Data availability

The scoria cone database is shared in Supplementary Data 1. The Digital Terrain Models are either commercially available products (e.g., World-DEM) or have been accessed through data portals (e.g., Open Topography, Land Information New Zealand).

Code availability

Morphometric data were extracted from Digital Elevation Models using a custom-built workflow in ArcGIS Pro Model Builder. Sentinel-2 reflectance data were extracted using Google Earth Engine, and the database was analyzed through Python. All tools and codes are shared through Figshare: https://doi.org/10.6084/m9.figshare.28973333.v1.

Received: 17 November 2024; Accepted: 29 May 2025;
Published online: 06 June 2025

References

1. Marder, E. & Gallen, S. F. Climate control on the relationship between erosion rate and fluvial topography. *Geology* **51**, 424–427 (2023).
2. Grosse, P., van Wyk de Vries, B., Petrinovic, I. A., Euillades, P. A. & Alvarado, G. E. Morphometry and evolution of arc volcanoes. *Geology* **37**, 651–654 (2009).
3. Hooper, D. M. & Sheridan, M. F. Computer-simulation models of scoria cone degradation. *J. Volcanol. Geotherm. Res.* **83**, 241–267 (1998).
4. Nieto-Torres, A. & Martin Del Pozzo, A. L. Spatio-temporal hazard assessment of a monogenetic volcanic field, near México City. *J. Volcanol. Geotherm. Res.* **371**, 46–58 (2019).
5. Conway, F. M. et al. Recurrence rates of basaltic volcanism in SP cluster, San Francisco volcanic field, Arizona. *Geology* **26**, 655–658 (1998).
6. Pedrazzi, D. et al. Morphometric analysis of monogenetic volcanoes in the Garrotxa Volcanic Field, Iberian Peninsula. *Geomorphology* **465**, 109400 (2024).
7. Zhang, R., Brenna, M., White, J. D. L. & Kereszturi, G. Variable controlling factors lead to contrasting patterns of volcanism in the Changbaishan volcanic area (Tianchi-Longgang), China-North Korea: Insights from morphometry and spatial-temporal analyses. *J. Volcanol. Geotherm. Res.* **451**, 108116 (2024).
8. Bebbington, M. S. & Cronin, S. J. Spatio-temporal hazard estimation in the Auckland Volcanic Field, New Zealand, with a new event-order model. *Bull. Volcanol.* **73**, 55–72 (2011).
9. Zimmerer, M. J. A temporal dissection of late Quaternary volcanism and related hazards within the Rio Grande rift and along the Jemez lineament of New Mexico, USA. *Geosphere* **20**, 505–546 (2024).
10. Pelletier, J. D. & Cline, M. L. Nonlinear slope-dependent sediment transport in cinder cone evolution. *Geology* **35**, 1067–1070 (2007).
11. de' Michieli Vitturi, M. & Arrowsmith, J. R. Two-dimensional nonlinear diffusive numerical simulation of geomorphic modifications to cinder cones. *Earth Surf. Processes Landf***38**, 1432–1443 (2013).
12. Kervyn, M., Ernst, G. G. J., Carracedo, J.-C. & Jacobs, P. Geomorphometric variability of "monogenetic" volcanic cones: evidence from Mauna Kea, Lanzarote and experimental cones. *Geomorphology* **136**, 59–75 (2012).
13. Zarazúa-Carbajal, M. C., Valentine, G. A. & De la Cruz-Reyna, S. Scoria cone erosional degradation by incision: different behaviors in three volcanic fields reflect environmental conditions. *Geology* **52**, 565–569 (2024).
14. Valentine, G. A. & Gregg, T. K. P. Continental basaltic volcanoes—-processes and problems. *J. Volcanol. Geotherm. Res.* **177**, 857–873 (2008).
15. Németh, K. & Kereszturi, G. Monogenetic volcanism: personal views and discussion. *Int. J. Earth Sci.* **104**, 2131–2146 (2015).
16. Guilbaud, M.-N. et al. Geology, geochronology, and tectonic setting of the Jorullo Volcano region, Michoacán, México. *J. Volcanol. Geotherm. Res.* **201**, 97–112 (2011).
17. Clarke, B. et al. Fluidal pyroclasts reveal the intensity of peralkaline rhyolite pumice cone eruptions. *Nat. Commun.* **10**, 2010 (2019).
18. Bemis, K. G. & Ferencz, M. Morphometric analysis of scoria cones: the potential for inferring process from shape. *Geolog. Soc. Lond. Spec. Publ.* **446**, 61–100 (2017).
19. Pioli, L. et al. Explosive dynamics of violent Strombolian eruptions: The eruption of Parícutin Volcano 1943–1952 (Mexico). *Earth Planet. Sci. Lett.* **271**, 359–368 (2008).
20. Alfano, F. et al. Subplinian monogenetic basaltic eruption of Sunset Crater, Arizona, USA. *GSA Bull.* **131**, 661–674 (2018).
21. Martin, U. & Németh, K. How Strombolian is a "Strombolian" scoria cone? Some irregularities in scoria cone architecture from the Transmexican Volcanic Belt, near Volcán Ceboruco (Mexico), and Al Haruj (Libya). *J. Volcanol. Geotherm. Res.* **155**, 104–118 (2006).
22. Planagumà, L., Bolós, X. & Martí, J. Hydrogeologic and magmatic controls on phreatomagmatism at the La Garrotxa monogenetic volcanic field (NE of Iberian Peninsula). *J. Volcanol. Geotherm. Res.* **441**, 107894 (2023).
23. Doubik, P. & Hill, B. E. Magmatic and hydromagmatic conduit development during the 1975 Tolbachik Eruption, Kamchatka, with implications for hazards assessment at Yucca Mountain, NV. *J. Volcanol. Geotherm. Res.* **91**, 43–64 (1999).
24. Romero, J. E. et al. The eruptive history and magma composition of Pleistocene Cerro Negro volcano (Northern Chile): Implications for the complex evolution of large monogenetic volcanoes. *J. Volcanol. Geotherm. Res.* **429**, 107618 (2022).
25. Kereszturi, G. & Németh, K. Sedimentology, eruptive mechanism and facies architecture of basaltic scoria cones from the Auckland Volcanic Field (New Zealand). *J. Volcanol. Geotherm. Res.* **324**, 41–56 (2016).
26. Pioli, L., Azzopardi, B. J. & Cashman, K. V. Controls on the explosivity of scoria cone eruptions: Magma segregation at conduit junctions. *J. Volcanol. Geotherm. Res.* **186**, 407–415 (2009).
27. Di Traglia, F., Cimarelli, C., de Rita, D. & Gimeno Torrente, D. Changing eruptive styles in basaltic explosive volcanism: examples from Croscat complex scoria cone, Garrotxa Volcanic Field (NE Iberain Peninsula). *J. Volcanol. Geotherm. Res.* **180**, 89–109 (2009).
28. McGetchin, T. R., Settle, M. & Chouet, B. A. Cinder cone growth modeled after Northeast Crater, Mount Etna, Sicily. *J. Geophys. Res.* **79**, 3257–3272 (1974).
29. Riedel, C., Ernst, G. G. J. & Riley, M. Controls on the growth and geometry of pyroclastic constructs. *J. Volcanol. Geotherm. Res.* **127**, 121–152 (2003).
30. Mannen, K. & Ito, T. Formation of scoria cone during explosive eruption at Izu-Oshima volcano, Japan. *Geophys. Res. Lett*. **34**, L18302 (2007).
31. Head, J. W. & Wilson, L. Basaltic pyroclastic eruptions: influence of gas-release patterns and volume fluxes on fountain structure, and the formation of cinder cones, spatter cones, rootless flows, lava ponds and lava flows. *J. Volcanol. Geotherm. Res.* **37**, 261–271 (1989).
32. Karger, D. N. et al. Climatologies at high resolution for the earth's land surface areas. *Sci. Data* **4**, 170122 (2017).
33. Panagos, P. et al. Global rainfall erosivity assessment based on high-temporal resolution rainfall records. *Sci. Rep.* **7**, 4175 (2017).
34. Wood, C. A. Morphometric analysis of cinder cone degradation. *J. Volcanol. Geotherm. Res.* **8**, 137–160 (1980).
35. O'Hara, D. et al. Time-varying drainage basin development and erosion on volcanic edifices. *Earth Surf. Dynam.* **12**, 709–726 (2024).
36. Fornaciai, A. et al. Detecting short-term evolution of Etnean scoria cones: a LIDAR-based approach. *Bull. Volcanol.* **72**, 1209–1222 (2010).

37. White, J. D. L., Houghton, B. F., Hodgson, K. A. & Wilson, C. J. N. Delayed sedimentary response to the A.D. 1886 eruption of Tarawera, New Zealand. *Geology* **25**, 459–462 (1997).
38. van der Meer, F. D., van der Werff, H. M. A. & van Ruitenbeek, F. J. A. Potential of ESA's Sentinel-2 for geological applications. *Remote Sens. Environ.* **148**, 124–133 (2014).
39. Kereszturi, G. & Németh, K. Post-eruptive sediment transport and surface processes on unvegetated volcanic hillslopes—a case study of Black Tank scoria cone, Cima Volcanic Field, California. *Geomorphology* **267**, 59–75 (2016).
40. Holness, S. D. Sediment movement rates and processes on cinder cones in the maritime Subantarctic (Marion Island). *Earth Surf. Processes Landf.* **29**, 91–103 (2004).
41. Franco-Ramos, O. et al. Reconstruction of gully erosion based on exposed tree roots in a recent landform of Paricutin Volcano, Mexico. *Earth Surf. Processes Landf.* **47**, 742–755 (2022).
42. Rad, S. D., Allègre, C. J. & Louvat, P. Hidden erosion on volcanic islands. *Earth Planet. Sci. Lett.* **262**, 109–124 (2007).
43. Kereszturi, G., Jordan, G., Németh, K. & Dóniz-Páez, J. Syn-eruptive morphometric variability of monogenetic scoria cones. *Bull. Volcanol.* **74**, 2171–2185 (2012).
44. Dohrenwend, J. C., Wells, S. G. & Turrin, B. D. Degradation of Quaternary cinder cones in the Cima volcanic field, Mojave Desert, California. *Geol. Soc. Am. Bull.* **97**, 421–427 (1986).
45. Parfitt, E. A. A discussion of the mechanisms of explosive basaltic eruptions. *J. Volcanol. Geotherm. Res.* **134**, 77–107 (2004).
46. Putirka, K. & Busby, C. J. The tectonic significance of high-K2O volcanism in the Sierra Nevada, California. *Geology* **35**, 923–926 (2007).
47. Barth, A. et al. Magma decompression rate correlates with explosivity at basaltic volcanoes — Constraints from water diffusion in olivine. *J. Volcanol. Geotherm. Res.* **387**, 106664 (2019).
48. Johnson, E., Wallace, P., Chashman, K., Granados, H. D. & Kent, A. Magmatic volatile contents and degassing-induced crystallization at Volcán Jorullo, Mexico: Implications for melt evolution and the plumbing systems of monogenetic volcanoes. *Earth Planet. Sci. Lett.* **269**, 478–487 (2008).
49. Stelten, M. E. et al. Timescales of magmatic differentiation from alkali basalt to trachyte within the Harrat Rahat volcanic field, Kingdom of Saudi Arabia. *Contrib. Mineral. Petrol.* **173**, 68 (2018).
50. Allison, C. M., Roggensack, K. & Clarke, A. B. Highly explosive basaltic eruptions driven by CO2 exsolution. *Nat. Commun.* **12**, 217 (2021).
51. Cassidy, M., Manga, M., Cashman, K. & Bachmann, O. Controls on explosive-effusive volcanic eruption styles. *Nat. Commun.* **9**, 2839 (2018).
52. Valentine, G. A., Ort, M. H. & Cortés, J. A. Quaternary basaltic volcanic fields of the American Southwest. *Geosphere* **17**, 2144–2171 (2021).
53. Schmidt, C. et al. The complexities of assessing volcanic hazards along the Cameroon Volcanic Line using spatial distribution of monogenetic volcanoes. *J. Volcanol. Geotherm. Res.* **427**, 107558 (2022).
54. Bebbington, M. S. Assessing spatio-temporal eruption forecasts in a monogenetic volcanic field. *J. Volcanol. Geotherm. Res.* **252**, 14–28 (2013).
55. Brož, P., Čadek, O., Hauber, E. & Rossi, A. P. Shape of scoria cones on Mars: Insights from numerical modeling of ballistic pathways. *Earth Planet. Sci. Lett.* **406**, 14–23 (2014).
56. Fornaciai, A., Favalli, M., Karátson, D., Tarquini, S. & Boschi, E. Morphometry of scoria cones, and their relation to geodynamic setting: a DEM-based analysis. *J. Volcanol. Geotherm. Res.* **217–218**, 56–72 (2012).
57. Favalli, M., Karátson, D., Mazzarini, F., Pareschi, M. T. & Boschi, E. Morphometry of scoria cones located on a volcano flank: A case study from Mt. Etna (Italy), based on high-resolution LiDAR data. *J. Volcanol. Geotherm. Res.* **186**, 320–330 (2009).
58. Zhang, R., Kereszturi, G., Brenna, M. & Ahn, U. S. Sensitivity assessment of morphometric parameters of monogenetic volcanic landforms with global free DEMs. *Geomorphology* **415**, 108408 (2022).
59. Wold, S., Sjöström, M. & Eriksson, L. PLS-regression: a basic tool of chemometrics. *Chemometr. Intell. Lab. Syst.* **58**, 109–130 (2001).
60. Vapnik, V. N. *The Nature of Statistical Learning Theory*. 1–314 (Springer-Verlag, 1995).
61. Global Volcanism Program (2024).
62. Bird, P. An updated digital model of plate boundaries. *Geochem. Geophys. Geosyst*. **4**, 1027 (2003).
63. Chong, I.-G. & Jun, C.-H. Performance of some variable selection methods when multicollinearity is present. *Chemometr. Intell. Lab. Syst.* **78**, 103–112 (2005).
64. Frey, F. A., Rhodes, J. M., Cox, K. G., McKenzie, D. P. & White, R. S. Intershield geochemical differences among Hawaiian volcanoes: implications for source compositions, melting process and magma ascent paths. *Philos. Trans. R. Soc. Lond. Ser. A***342**, 121–136 (1993).
65. Stelten, M. E. et al. The timing and compositional evolution of volcanism within northern Harrat Rahat, Kingdom of Saudi Arabia. *GSA Bull.* **132**, 1381–1403 (2019).

Acknowledgements

The authors acknowledge support from the 'EVoLvE: Modelling temporal evolution of volcano landscape through erosion' project, funded by Research Foundation—Flanders (FWO)—FWO Junior Project Grant G029820N. GK has further been supported by Massey University Research Fund and MBIE's National Science Challenges—Resilience to Nature's Challenges programme (GNS-RNC047). Any use of trade, product, or firm names is for descriptive purposes only and does not imply endorsement by the U.S. Government. The authors gratefully acknowledge the constructive comments provided by Kellie Wall and Eric Breard on an earlier draft of this manuscript.

Author contributions

G.K. conceived the idea and compiled the database with inputs from P.G., M.N.G., D.T.D., R.N., and M.K. G.K. led the data analysis, with contributions from M.W., and wrote the first paper. All authors have contributed to the interpretations, discussions and conclusions.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s43247-025-02425-8.

Correspondence and requests for materials should be addressed to Gabor Kereszturi.

Peer review information *Communications Earth & Environment* thanks Dario Pedrazzi, Hugo Murcia and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Domenico Doronzo and Carolina Ortiz Guerrero. A peer review file is available.