




Eruption plumes extended more than 30 km in altitude in both phases of the Millennium eruption of Paektu (Changbaishan) volcano

Antonio Costa¹[✉], Leonardo Mingari², Victoria C. Smith³, Giovanni Macedonio⁴, Danielle McLean³, Arnau Folch², Jeonghyun Lee⁵ & Sung-Hyo Yun⁵

The Millennium Eruption of Paektu volcano, on the border of China and North Korea, generated tephra deposits that extend >1000 km from the vent, making it one of the largest eruptions in historical times. Based on observed thicknesses and compositions of the deposits, the widespread tephra dispersal is attributed to two eruption phases fuelled by chemically distinct magmas that produced both pyroclastic flows and fallout deposits. We used an ensemble-based method with a dual step inversion, in combination with the FALL3D atmospheric tephra transport model, to constrain these two different phases. The volume of the two distinct phases has been calculated. The results indicate that about 3–16 km³ (with a best estimate of 7.2 km³) and 4–20 km³ (with a best estimate of 9.3 km³) of magma were erupted during the comendite and trachyte phases of the eruption, respectively. Eruption rates of up to 4×10^8 kg/s generated plumes that extended 30–40 km up into the stratosphere during each phase.

¹Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna, Bologna, Italy. ²Geociencias Barcelona (GEO3-BCN), CSIC, Barcelona, Spain. ³Research Laboratory for Archaeology and the History of Art, School of Archaeology, University of Oxford, Oxford OX1 3TG, UK. ⁴Istituto Nazionale di Geofisica e Vulcanologia, Osservatorio Vesuviano, Naples, Italy. ⁵Volcano Specialized Research Center, Pusan National University, Busan, South Korea. ✉email: antonio.costa@ingv.it

Large explosive volcanic eruptions cover wide areas in thick pyroclastic deposits and release volatiles into the atmosphere that can perturb climate for years^{1–3}. Although the impacts of large volcanic eruptions can be far reaching, they vary considerably. The impact on climate is linked to the amount of sulfur injected into the stratosphere, but models show that the latitude and timing of the injection have implications for the longevity of the sulfate in the stratosphere and thus also control their effects^{2,4}. Despite a wealth of detailed models, there are only a few large eruptions for which there is enough data on the aspects thought to control climate impacts, and even still there are considerable gaps in understanding and large uncertainties⁵. Quantifying the volume of magma erupted, volatiles released and plume height of a variety of past eruptions is critical to provide empirical evidence to further interrogate the relationship between explosive eruptions and their climate effects, as well as testing the accuracy of dispersal models.

Sulfate layers preserved in the polar ice cores provide a detailed chronology of volcanic eruptions in the last few thousand years, many of which have been attributed to particular eruptions using glass compositions that act as a chemical fingerprint⁶. This time anchored history allows for the impact of various eruptions to be assessed using paleoenvironmental records (e.g., tree rings and pollen⁷) and written records for more recent events. It is important that these impacts are linked to eruption parameters to further inform our understanding on what controls the impact on climate. A notable eruption for which we have good constraints on both the date and climate response is the 946 CE Millennium

eruption (ME) of Paektu, an active caldera situated on the border of the Democratic People's Republic of Korea (DPRK) and People's Republic of China (Fig. 1). The tephra erupted during this event is found over large areas, including over Greenland that is more than 7000 km from vent⁸, suggesting that it was a large magnitude eruption of ~M6.7, but the climatic impact appears limited^{9,10}. This is at odds with what is observed from other large eruptions (e.g., the 1815 CE eruption of Tambora, Indonesia¹¹) so it is not clear whether the eruption parameters for this event have been accurately constrained. Here we combine tephra fallout thickness measurements with a new ensemble-based method as part of a dual step inversion, which involves the FALL3D atmospheric tephra transport and deposition model^{12,13}, to constrain the magnitude of the eruption and the eruption parameters such as mass eruption rate, duration, and plume height. For the first time we model the different phases of the eruption which are known from the proximal and distal deposits. By coupling these new volume constraints with published petrological data on the volatile content¹⁴ we can update the estimates of the amount of the volatiles released into the stratosphere and ultimately help understand the limited climatic impact.

A large eruption from Paektu. The discovery of a fine and widespread volcanic ash (tephra) layer in northern Japan was the first evidence to suggest there was a recent eruption from Paektu volcano¹⁵, also referred to as Baekdu (South Korean), Baegdusan (Japanese), or Changbaishan (Chinese). This visible tephra was widely identified in sequences in northern Japan, including

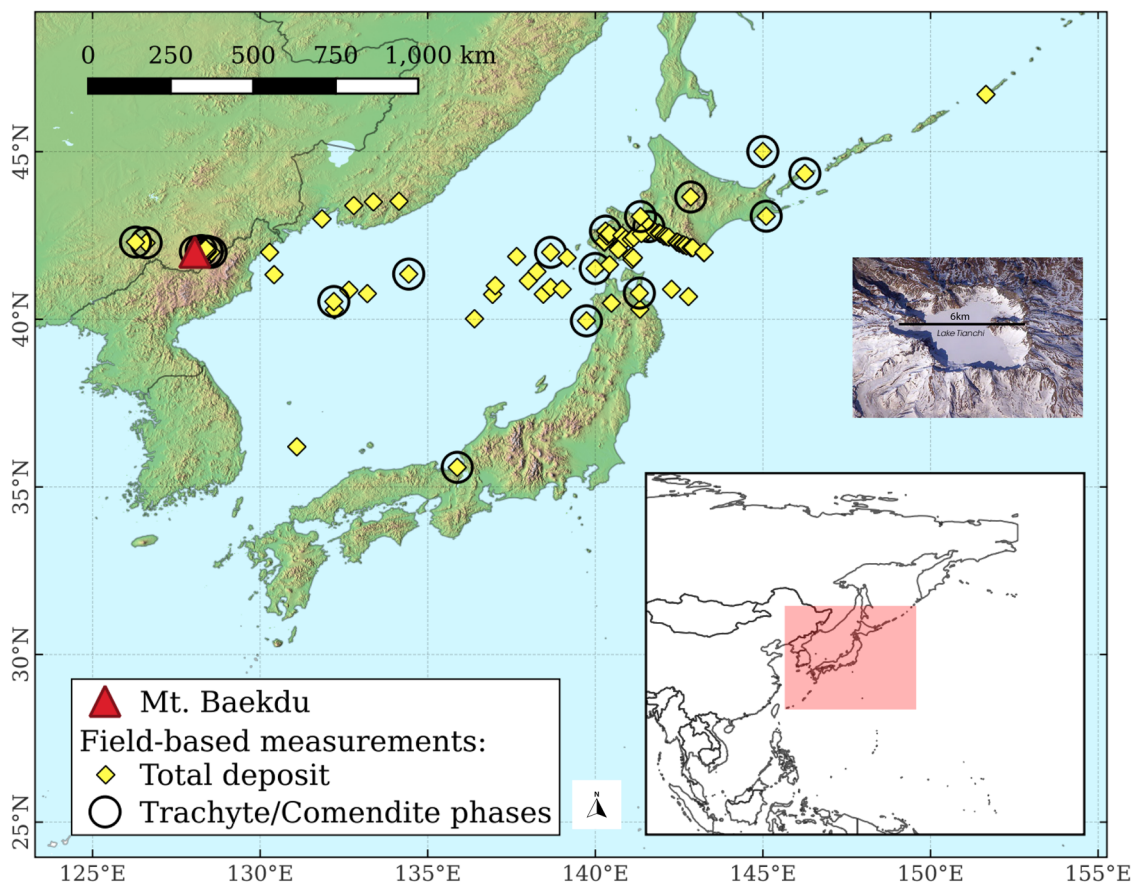


Fig. 1 Studied area. Studied area and locations of tephra observations used for the dispersal modelling, compiled from terrestrial exposures and marine cores, for the entire deposit and for the points where the two magmatic phases can be distinguished (see Supplementary Table 1 for further details and references). The map was created using the QGIS open software (<https://qgis.org/>). Paektu caldera, around Lake Tianchi, in the inset is from astronaut photograph ISS006-E-43366 acquired April 4, 2003 (ISS Crew Earth Observations experiment and the Image Science & Analysis Group, Johnson Space Center).

southern Hokkaido, northern Honshu, and the Sea of Japan¹⁵. Based on the decrease in tephra thickness to the east and the presence of K-feldspar, consistent with observations from pyroclastic density current (PDC) deposits around the volcano¹⁶, Paektu was identified as the source by Machida and Arai¹⁵. This distal tephra layer was labelled B-Tm following Japanese tephrostratigraphy conventions, with B referring to Baegdusan (Paektu) and Tm relating to the type locality for the deposit at Tomakomai, Hokkaido¹⁵. Since these initial discoveries, the B-Tm tephra layer has been identified in numerous marine, lacustrine and archaeological sequences across northern Japan, northeast China, and coastal regions of Russia (see McLean et al.¹⁷ and references cited in Supplementary Table 1). In addition, glass shards with B-Tm geochemical compositions have been identified as a non-visible layer (cryptotephra) in ice cores from Greenland⁸, over 7000 km from its source. The glass compositions of the tephra deposits across Japan verify that the occurrences are indeed from the same eruption of Paektu volcano and have correlated the B-Tm tephra to the Millennium eruption (ME) deposits¹⁸.

Geological history of Paektu volcano. Volcanism at the intra-plate Mt. Paektu (42.00°N, 128.05°E, 2750 m a.s.l.) is fuelled by the upwelling of mantle material. Initial basaltic fissure eruptions commenced at the location ~28 My ago and the volcanic edifice started to develop at ~4 Ma with activity divided into three stages. A basaltic shield volcano developed in Stage 1 and more explosive activity began in Stage 2 at around 1 Ma and built the trachytic-rhyolitic stratovolcano¹⁹. The Cheonji (Tianchi in Chinese) caldera-forming Stage 3^{19,20} generated several large eruptions in the last 80 ka²¹. Deposits are found in distal records testifying to multiple explosive eruptions during this most recent stage^{22–24}, which have spanned a range of magnitudes from small vulcanian eruptions, such as the 1903 event²⁵, to the Plinian Millennium eruption²⁶.

Precise date of the Millennium eruption. Radiocarbon dates of tree remains found with the B-Tm ash initially indicated the eruption occurred between 794–1192 CE¹⁵. The date of the tephra was shown to be consistent with that of the deposits close to the source, which was subsequently constrained further over the years using radiocarbon measurements of organic material recovered in the ME deposits around the vent. A precise wiggle-matched radio-carbon age of 946 ± 3 CE was obtained from a tree felled by the eruption that was recovered from the PDC on the western flank of the volcano⁹. Further confirmation that the eruption occurred in 946 CE was established by Oppenheimer et al.¹⁰ by obtaining radiocarbon measurements and identifying the 775 CE cosmogenic event²⁷ across a similar tree stump recovered from the same location within the deposits. The B-Tm tephra is also found in Greenland ice cores and a date of 939–940 CE was assigned to the eruption using the Greenland Ice Core Chronology⁸. Subsequent research established the GICC05 ice core timescale was offset by –6 years²⁸ so the ice core date is consistent with that obtained using the tree stump.

Written records from across the region document phenomena consistent with a large volcanic eruption in both 946 and 947 CE, with references such as the “roll of drums within the sky” in 946 CE (record from the Koryŏ dynasty, Korean Peninsula^{29,30}), “white ash falls as snow in the night” in Nara, Japan, on 3rd November 946 CE^{10,30,31}, and two records from 7th February 947 CE that noted “sound in the sky, like thunder”^{30,32}. These historic accounts and the tree and ice records indicate that the eruption certainly started in 946 CE, with the further historical accounts

suggesting there may have been another eruption phase or other activity in early 947 CE.

Proximal deposits of the Millennium eruption. The Millennium eruption stratigraphy around Paektu has been the focus of numerous studies over the last three decades²¹. The studies have been focussed on the deposits on the separate sides of the border, with seemingly no research group working on those from both China and DPRK. Various pyroclastic units are identified around the volcano but their attribution to a particular eruptive event had remained somewhat ambiguous. Both Horn and Schmincke³³ and Pan et al.²¹ indicated the ME deposits had both a comendite and trachytic phase. However, Horn and Schmincke³³ noted that the comenditic component was the most dominant and only ~5% of the melt erupted was trachytic in composition, found in mingled clasts within the deposit. Their ratio of comendite to trachyte is inconsistent with the similar thickness measurements of the comenditic and trachytic fallout units recorded by Pan et al.²¹, and at odds with the distal records that reveal that up to 50% of the glass shards are trachytic in composition¹⁷. Pan et al.²¹ interpreted that all the proximal units labelled Baguamiao, Millennium and some others, attributed by other researchers^{34,35} to later eruptions, were all phases of the ME. As such, Pan et al.²¹ labelled the deposits associated with the 946 CE eruption the Generalized Millennium Eruption, here referred to as ME, and note that the sequence includes a light grey and dark grey coloured comenditic fallout and PDC and a much darker coloured trachytic fallout and PDC. Both comenditic and trachytic fallout phases are decimetres in thickness in sections ~30 km east of the vent. The comenditic PDC is more than 10 m thick at sections <15 km east of the vent and it is partially welded in places^{21,33}. At least 0.4 m of a trachytic PDC was deposited after a fallout deposit from the same magma, but it has been eroded in most sections suggesting that it is a minimum thickness for the deposit²¹. The observed deposit thicknesses both proximally and distally suggest that the comendite and trachyte fallout phases were similar in size, but the comenditic PDC was greater in volume relative to the trachytic PDC.

A lahar deposit was observed by Pan et al.²¹ between the two phases of the ME at a locality ~15 km east of the volcano. They suggested that this could reflect a hiatus of a few wet months, which is consistent with historic accounts that record phenomena commonly related to volcanic eruptions from 3rd November 946 through to 7th February 947 CE²¹. Although the distal sedimentary records seemingly indicate a single layer¹⁵, the sedimentation rates at the sites are not high enough to record detectable sedimentation between deposits of the different phases if they were indeed separated by months. In addition, current ice core sampling routines also mean that datasets are unlikely to show a clear separation in the different phases.

Tephra fallout volumes were recently reassessed by Yang et al.³⁶ based on known proximal and distal thicknesses of the deposit. They dealt with the entire deposit thickness (both comenditic and trachytic) and estimated that the fallout equated to 13.4 to 37.4 km³ Dense Rock Equivalent (DRE) of magma³⁶. The PDC volumes associated with the eruption were estimated by Horn and Schmincke³³ to be between 5.3–7.6 km³ DRE (12.3 to 17.5 km³ bulk volume) based on assumptions of their extent in China. Recently, Zhao et al.³⁷ mapped the PDC deposits in China and estimated they accounted for ~ 3 km³ DRE (7 km³ bulk volume). These previous estimates were reconciled by Yang et al.³⁶ to constrain the total extra-caldera PDC volume to ~4.1–5.7 km³ DRE (9.3–13.0 km³ bulk) and estimated up to ~2.1 km³ DRE (maximum of 4.9 km³ bulk tephra) of intra-caldera PDC. These estimates indicate that the bulk volume of the

entire ME deposit is between 40.2 and 97.7 km³, which equates to 17.5 to 42.5 km³ DRE magma assuming a tephra deposit density of 1000 kg/m³ and a magma density of 2300 kg/m³ that is appropriate for rhyolite. Nonetheless, these represent two different magma compositions and the ratio of comendite to trachytic magma was not quantified.

Impact. Proxy records of past climate such as the maximum Latewood Density of tree rings, which is sensitive to summer temperatures⁷, from a range of Northern Hemisphere locations, including China, do not show noticeable changes following the Millennium eruption¹⁰. Furthermore, the non-sea salt sulfur (nssS; i.e. that deposited from volcanic emissions) recorded in the Greenland ice core is ~30 ppb at 946 CE, roughly double the background levels, with no perturbation in nssS in Antarctic records²⁸. These nssS are less than half that recorded in Greenland (80 ppb; 80–130 ppb in Antarctic ice cores) following the 1815 CE eruption of Tambora, Indonesia, but the latitudes are very different which would have affected sulfate deposition²⁸. This muted climatic response indicates either low S volatile concentrations or tropospheric plume heights.

New insights into the Millennium eruption. Here we use all the published thickness and compositional data of the deposits that are found over an area >700,000 km² east of the volcano (see Fig. 1 and Supplementary Table 1) with a dual-step inversion method to model the tephra dispersal from the two phases of the Millennium eruption. A new ensemble-based inversion modelling strategy³⁸ was used to gain insights into the eruption dynamics (e.g., mass eruption rate and column height) and estimate the volume of the comenditic and trachytic magma erupted during the ME. We then use these total volume estimations with existing data on the volatile compositions of the melts^{14,33} to estimate the likely total volatile release. The probabilistic nature of the ensemble-based method³⁸ used for solving this inverse problem allowed us to properly assess the uncertainties.

Results and discussion

Tracing the dispersal of the comenditic and trachytic phases. Although the comenditic and trachytic phases of the ME deposit are visually distinct in proximal locations within 40 km of the volcano, at distal locations (more than a few hundred kilometres from the vent) the two phases cannot be distinguished visually¹⁵. However, since the melts are compositionally distinct the glass chemistry can be used to establish which phase is present at a site. Many distal sites have deposits from both phases, and the percentage of which phase is present can be estimated based on the proportion of glass shards of each composition¹⁷. We used this published glass chemistry to attribute a thickness to each phase at distal sites where the tephra thickness is also reported. This assumes that the glass compositions analysed are representative, which may not be the case, but given that they are similarly easy to analyse there is no likely bias. The B-Tm ash in some distal records does not form a visible layer, with the layers (termed cryptotephra) being identified through density extraction techniques³⁹. McLean et al.^{17,40} carried out detailed cryptotephra analysis of the Lake Suigetsu sediment cores (central Japan, ~950 km from Mt Paektu) and noted that the B-Tm is only visible in some core sections and has 1 mm in thickness. The tephra corresponds to shard concentrations of 25,000 shards/gram dry sediment in the 1 cm (core depth) sample⁴⁰. Based on these measurements, we have assumed that 25,000 shards/gram of dry sediment equates to 1 mm in thickness and used this to assess likely tephra thickness where only shard concentrations are known (see Supplementary Table 1).

We have 23 locations for which we know thicknesses for both the comenditic (first) and trachytic phases of the ME deposit^{17,21,40}. These data points were used for the inversion (details below), and the solutions related to each eruptive phase then used to compare and validate the results with the total thicknesses of the deposit for which more points are available ($n = 83$; see Fig. 1 and Supplementary Table 1).

Solving an inverse problem using an ensemble-based method.

To estimate Eruption Source Parameters (ESPs) from observed deposit data we used a dual step inversion method. The first-step inversion used a fast semi-analytical code, such as PARFIT-2.3.1 (<http://datasim.ov.ingv.it/download/parfit/download-parfit-2.3.1.html>), to identify suitable wind conditions and ranges of ESPs for the subsequent ensemble-based inversion. For this purpose, we considered more than 18,000 daily-averaged wind profiles near the eruptive vent location (Supplementary Table 1) over a period of 50 years (1961 to 2010), which were obtained from the ERA5 Reanalysis⁴¹.

We solved three separate inverse problems, one for the comenditic phase, one for the trachytic phase, and another for the total deposit (two phases together; Supplementary Table 1). For each case we used the weighted least square method, considering both proportional and statistical weights (i.e. $\theta_{1,i} \propto 1/T_i^2$ and $\theta_{2,i} \propto 1/T_i$, being θ the weights and T deposit loads⁴²), and allowed wide ranges of ESP and meteorological conditions. In this step, we considered only the tephra load at distances >40 km from the vent where the comendite and trachyte fraction have been estimated (see Supplementary Table 1; proximal outcrops were excluded as the model is oversimplified at these distances⁴³). The results of this step generated the ranges of input needed for the second step (see “Methods”).

The results of the first step inversion are reported in Table 1. For the comenditic phase, the total erupted mass (TEM), expressed as DRE, ranges from 16 to 20 km³ and column height is about 22–26 km above vent level (a.v.l.), whereas, for the trachytic phase, the volume ranges from 17 to 36 km³ DRE with a column height from 26 to 35 km (a.v.l.). Using the dataset consisting of all measurement points available at a distance from the volcano greater than ~40 km, we found an erupted magma volume of about 15–19 km³ DRE and a column height ranging from 38 to 45 km. The wide ranges obtained for the TEM and column height, and the fact that the total volume is smaller than each individual phases, reflects the intrinsic uncertainties in the solution of the inverse problem with simplified analytical models such as PARFIT, which are not well constrained due to the scarcity and sparsity of the data^{44,45}. Furthermore, solving the column height and wind intensity with only thickness data is not ideal^{46,47}. For these reasons, these first-step ranges were used as a starting point to find an inverse problem solution using the second ensemble-based method.

In the second step, for each optimal meteorological configuration (Table 1) we performed six ensemble runs using the FALL3D-8.2 dispersal model^{38,48} considering an ensemble size of 128 members. The best meteorological configurations (start date reported in Table 1) were used to drive the dispersal model for each ensemble run. This approach provides an ensemble of deposit mass loading field estimations (first guess). Subsequently, observations are assimilated to reconstruct the deposit mass loading (analysis) using the GNC method described by Mingari et al.⁴⁸. The deposit was reconstructed for the individual phases considering the six meteorological configurations and assimilating the 23 observations (assimilation dataset) corresponding to each phase. The total tephra deposit is obtained as the sum of the comenditic and trachytic deposits. The total deposit resulting

Table 1 Results of PARFIT tephra dispersal inversion.

Eruptive parameters	Explored range	Combined (θ_1)	Combined (θ_2)	Comendite phase (θ_1)	Comendite phase (θ_2)	Trachyte phase (θ_1)	Trachyte phase (θ_2)
Optimal meteorological configuration	ERA5 reanalysis 1961–2010	30 September 1968	30 October 1992	29 November 1978	8 October 1975	23 September 1986	21 April 1994
TEM (10^{13} kg)	Calculated analytically	4.5	3.6	3.9	4.9	4.1	8.7
Volume (DRE; km^3)	Calculated analytically	18.9	15	16.4	20.4	17.4	36.5
Column height a.v.l. (H , km)	15–45	38	45	26	22	26	35
Suzuki coefficient ⁷⁰	1–9	6	2	4	2	6	3
Diffusion coefficient (K , $\text{m}^2 \text{s}^{-1}$)	1000–30,000	8000	16,000	20,000	28,000	19,000	20,000
Density of aggregates (kg m^{-3})	100–600	300	300	300	300	275	300
Deposit density (kg m^{-3})	Assumed	1000	1000	1000	1000	1000	1000
χ^2	Calculated	0.72	0.37	0.31	0.10	0.13	0.05

Results of PARFIT tephra dispersal (first step) inversion for the ME eruption. θ_1 and θ_2 represent proportional and statistical weights in the PARFIT inversion, respectively.

from every combination (36 possible configurations) was evaluated using different metrics computed using the full dataset of 83 measurements of total deposit thickness (independent validation dataset).

The root-mean-square error (RMSE) was used to measure the differences between the values observed and the analyses interpolated at the sampling sites. For this purpose, we used the weighted version of RMSE and other statistical indicators, as explained in the Methods section. An overview of all used statistical parameters is reported in Supplementary Table 2.

The optimal configuration, that minimises the *Confidence Ratio* and *Composite index* (that is on average all the other metrics), corresponds to the deposit reconstructed using the meteorological datasets starting on 23 September 1986 for the comendite phase and 30 October 1992 for the trachyte phase (see Supplementary Table 2). Similarly optimal solutions for the are comendite phase were obtained with the meteorological conditions starting on 23 September 1986, and additional optimal meteorological conditions for the trachytic phase were those observed on 21 April 1994 (RMSE and MAE) and 29 November 1978 (Bias and SMAPE) (see Supplementary Figures). In our approach these represent the most compatible meteorological conditions scanned across the ERA5 dataset. The reconstructed deposit fields for the comendite and trachyte phases and for the total deposit are shown in Fig. 2.

Figure 3 shows measurement and analysis data from the assimilation datasets (23 measurements for each eruptive phase) and for the validation dataset (83 measurements for the total deposit). Note that a RMSE close to 1, like in the presented case (Supplementary Table 2), means that deviations between measurements and analysis are similar to the observation errors, i.e. $|y_i^{\text{obs}} - y_i^{\text{an}}| \sim \epsilon_i$. On the other hand, a positive bias (0.39–0.40) for the optimal configuration means that the analysis tends to underestimate the observations, but the bias is within the observation error range ($|\text{Bias}| < 1$). Finally, more than 68% of the observations are within a factor 2.4 (that is one standard deviation for a lognormal distribution). The other two optimal solutions are reported in the Supplementary Figures.

The source term was inverted using the procedure based on the GNC method described by Mingari et al.⁴⁸. This method provides a set of weight factors (w_i) for every ensemble member (see “Methods” section). According to the inversion procedure, the

first-guess emission source term for the i th ensemble member (s_i) is then rescaled according to $s_i \rightarrow w_i s_i$. The resulting total source term (in $\text{kg s}^{-1} \text{m}^{-3}$), defined as the weighted sum of the individual source terms:

$$s = \sum_i w_i s_i \quad (1)$$

is represented in Fig. 4 in terms of the linear source emission strength δs (in $\text{kg s}^{-1} \text{m}^{-1}$), i.e. $\delta s dA$ being dA the area of the cell grid. The emission rate (i.e. δs vertically integrated) is also represented in Fig. 4 (dash-dotted blue line) along with the cumulative erupted volume (solid black line). The total erupted volume for the fallout phases according to the inverted source term was ~ 7.2 and 9.3 km^3 DRE for the comendite and trachyte phase, respectively (see Fig. 4 and Table 2), with a mean error of about a factor 2.4, that is a volume between 3 to 17 km^3 for the comendite phase and between 4 and 22 km^3 for the trachyte phase.

Volume and eruption parameters. The results from this new approach indicate that fallout (Plinian and co-ignimbrite) from each of the compositionally distinct fallout phases erupted roughly the same volume of magma, that is about $7\text{--}9 \text{ km}^3$ DRE (Table 2). This indicates that around 16 km^3 of magma was dispersed by the two fallout phases, but, considering the uncertainty, such a value could have been as low as 7 km^3 and up to 39 km^3 . For comparison, Horn and Schmincke³³ calculated a volume of $23 \pm 5 \text{ km}^3$ DRE for the comendite phase, and recent re-estimates by Yang et al.³⁶ for the fallout for both phases were estimated to be between 13.4 and 37.4 km^3 DRE. The tephra fallout volumes need to be added to the PDC estimates, which are $4.1\text{--}5.7 \text{ km}^3$ DRE for those deposited outside the caldera and an additional $\sim 2.1 \text{ km}^3$ DRE for PDC emplaced in the caldera³⁶ to estimate the total volume erupted. Our new fallout volumes combined with those estimated for the PDCs indicate the total volume of magma erupted during the ME was about 23 km^3 DRE, with at least 13 km^3 and up to 47 km^3 DRE erupted, corresponding to a M6.5 to M7.0 eruption. The average volume (23 km^3) is similar to the amount of material that missing from edifice (22 km^3) that would have collapsed into the void generated by evacuating the magma. The amount of material displaced was estimated using 2.8 km as the mean radius of the caldera, obtained averaging the maximum distances along the different

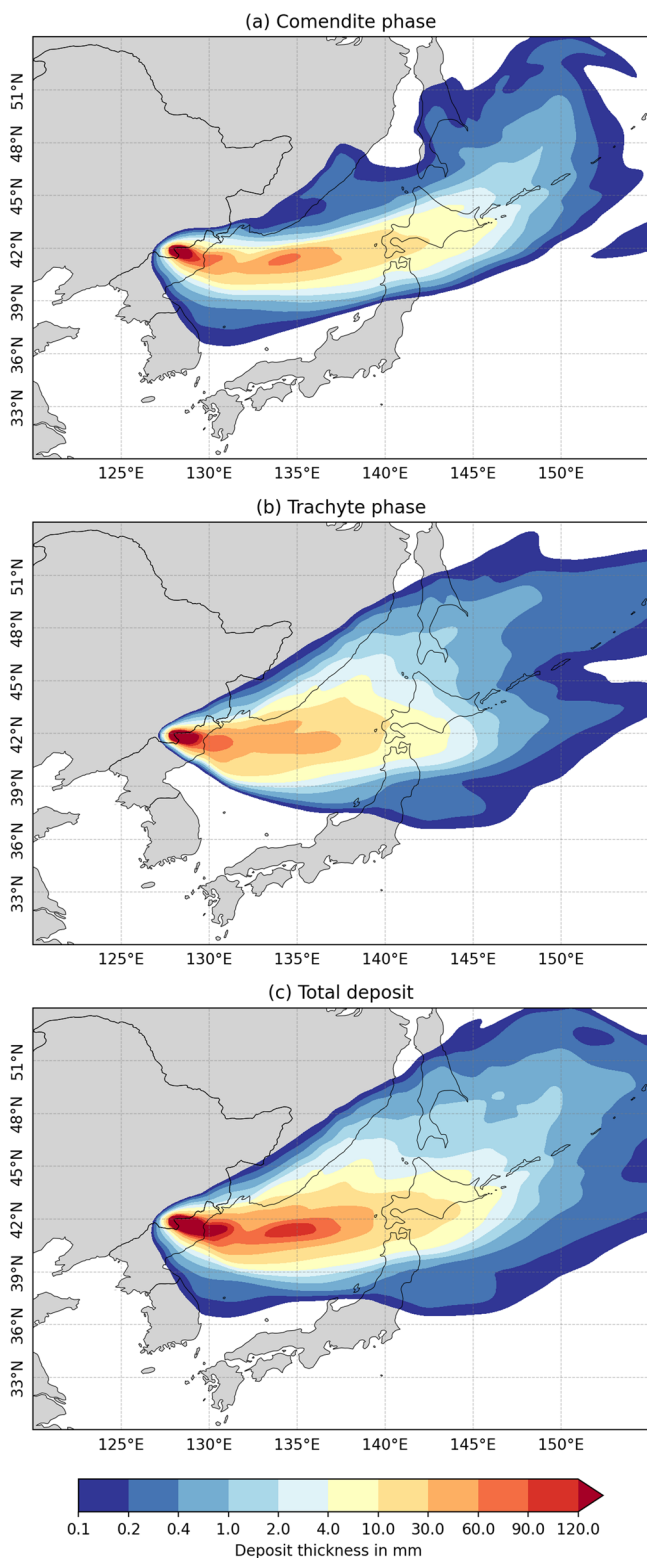


Fig. 2 Reconstructed tephra deposits. Reconstructed deposit thickness contours for the comendite (a), trachyte (b) eruptive phases, and for the total deposit (c) of the Millennium eruption (ME). The maps were created using the open software “Cartopy: a cartographic python library with a matplotlib interface” (<https://scitools.org.uk/cartopy/docs/latest/index.html>).

directions, and ~900 m height based on the difference between the maximum elevation and the bottom of the lake. The agreement between these values suggests that the optimal estimations are more realistic than the upper bounds.

Our results indicate that each of the two phases lasted less than a day (~15 and 20 h; Fig. 4) and peak mass flow rates were estimated to be about 4×10^8 kg/s, which are typical for Plinian eruptions⁴⁹. Similar values were estimated from observations of the climatic phase of the 1991 Pinatubo eruption^{50,51}. These high mass flow rates would have fed umbrella clouds⁵¹ and sustained columns that, for both phases, reached ~30–40 km (Fig. 4 and Table 2), which are well above the tropopause height of ~12 km at this latitude. This indicates that almost all of the eruptive mixture, including volatiles released during the climatic phase of the eruption, were injected into the stratosphere.

Volatile release and impact. The petrological method^{52,53} was previously used with the comendite volumes³³ (23 ± 5 km³ DRE) with volatile contents of the melt inclusions (MI) and matrix glasses to estimate the amount of S, Cl and F released. Measurements of the volatiles in MI and matrix glasses from the comenditic magma have been made by Horn and Schmincke³³, Iacovino et al.¹⁴ and very recently re-analysed by Scaillet and Oppenheimer⁵². The F and Cl contents of MI and matrix glasses cover a similar range^{14,33}, suggesting the melts were probably not saturated in F or Cl, and loss of these volatile phases could be negligible. Iacovino et al.¹⁴ modified the petrological approach used to also estimate the syn-eruptive degassing to quantify the total pre-eruptive volatile loss associated with the comenditic magma. This approach assumed the comendite erupted was linked to the trachyte erupted via fractional crystallisation, which is consistent with regression modelling of the major elements. They compared the observed volatile contents relative to incompatible elements (U in this case) to values estimated from fractional crystallisation models to estimate the volatiles lost from the melt during crystallisation¹⁴. Their data indicates that the pre-eruptive volatile release from the comenditic magma could have been substantial.

If we consider comendite magma volumes calculated with the new method presented here, which range from 3 to 17 km³ DRE, and the mean volatile contents measured by Iacovino et al.¹⁴ in the MI and matrix glass, we can rescale the volatiles that were removed from the melt prior to eruption to be between 5 and 30 Tg S, 6–32 Tg F, and 2–15 Tg Cl. While our scaled estimates for the syn-eruptive S released into the stratosphere is 0.5–2.6 Tg (1.0–5.2 Tg SO₂). Nonetheless, these estimates do not consider the volatile release associated with the trachytic melt (4–22 km³ DRE) as there are no published volatile data for trachytic matrix glass or MI in crystals exclusively associated with the trachytic magma. The new volumes suggest the pre-eruptive volatile release would have been smaller than previously reported by Iacovino et al.¹⁴. The volatile elements released prior to the eruption are likely to have been sequestered by the aquifers and hydrothermal fluids or passively degassed, and it is expected that most of these volatiles did not make it into the stratosphere. Nonetheless, it appears that some of the pre-eruptive S loss from the 1991 Pinatubo magma made its way into the stratosphere (which is higher at the more tropical latitude; ~18 km) as the satellite measured S content in the stratosphere was greater than the syn-eruptive load estimated by the petrological method⁵³.

The lack of recorded climatic impacts for the eruption^{7,10} suggests a low syn-eruptive volatile release, which is consistent with the above scaled syn-eruptive yields despite not including any release from the trachytic melt. The S yield is similar to the ~2 Tg estimated from the ice core non-sea salt sulfate record⁵⁴,

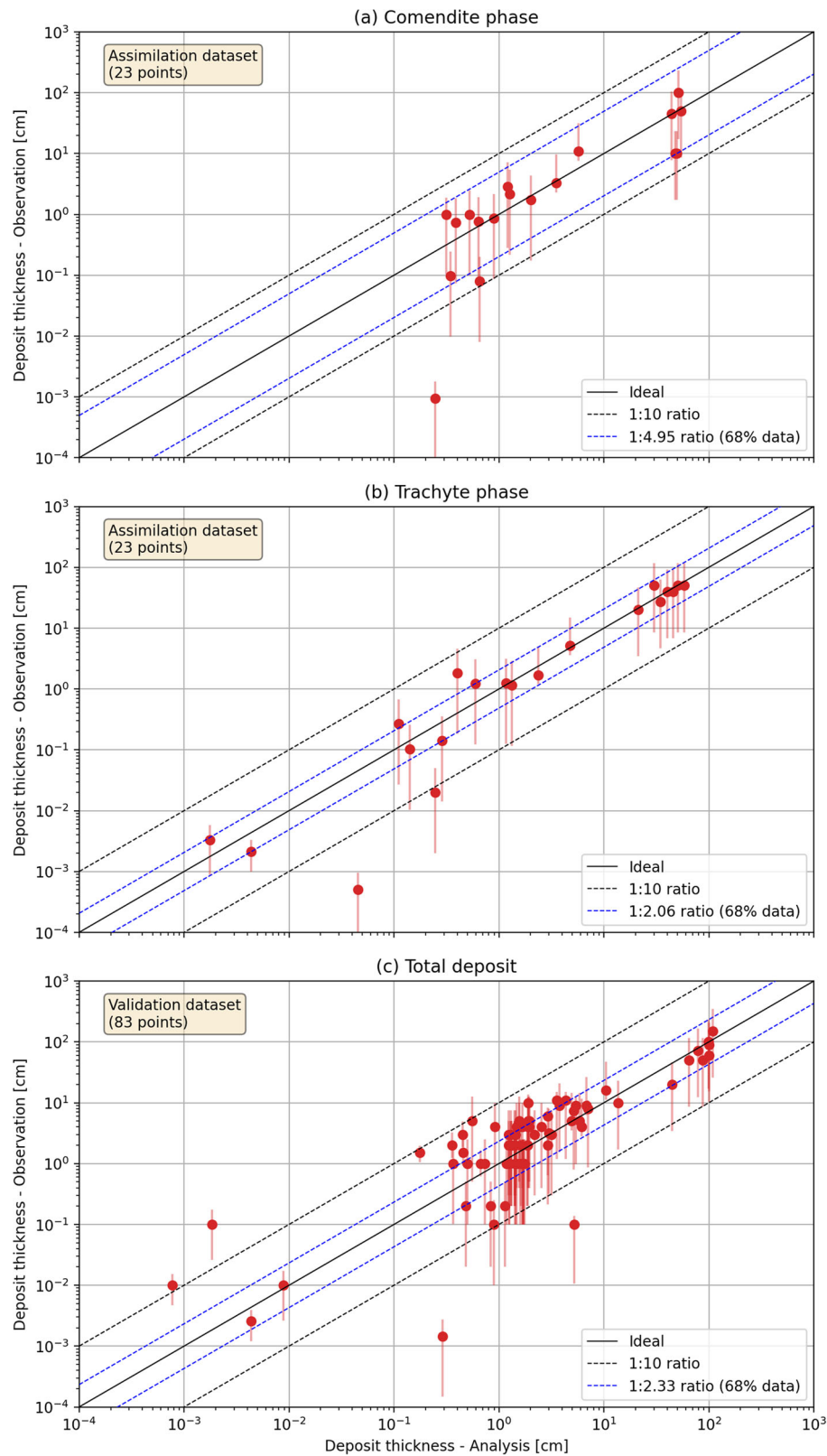


Fig. 3 Comparison models and observations. Quantitative comparison between observed and modelled (analysis) deposit thickness of the comendite (a) and trachyte (b) eruptive phases of the Millennium eruption (ME) using the assimilation dataset (23 measurements for each eruptive phase). In addition, the comparison between the total deposit thickness and the validation dataset (83 measurements) is also shown (c). Note that zero-valued observations for the comendite (P82 and P83, Supplementary Table 1) and trachyte (P34, P68, P72, P74, Supplementary Table 1) phases are not shown in these log-log plots (a) and (b), respectively. The plot corresponding to the total deposit (c) includes the full dataset of 83 data points. Errors are estimated in accord to Mingari et al.³⁸ assuming a minimum value of 30%. The plots were created using the open software Matplotlib (<https://matplotlib.org/stable/>).

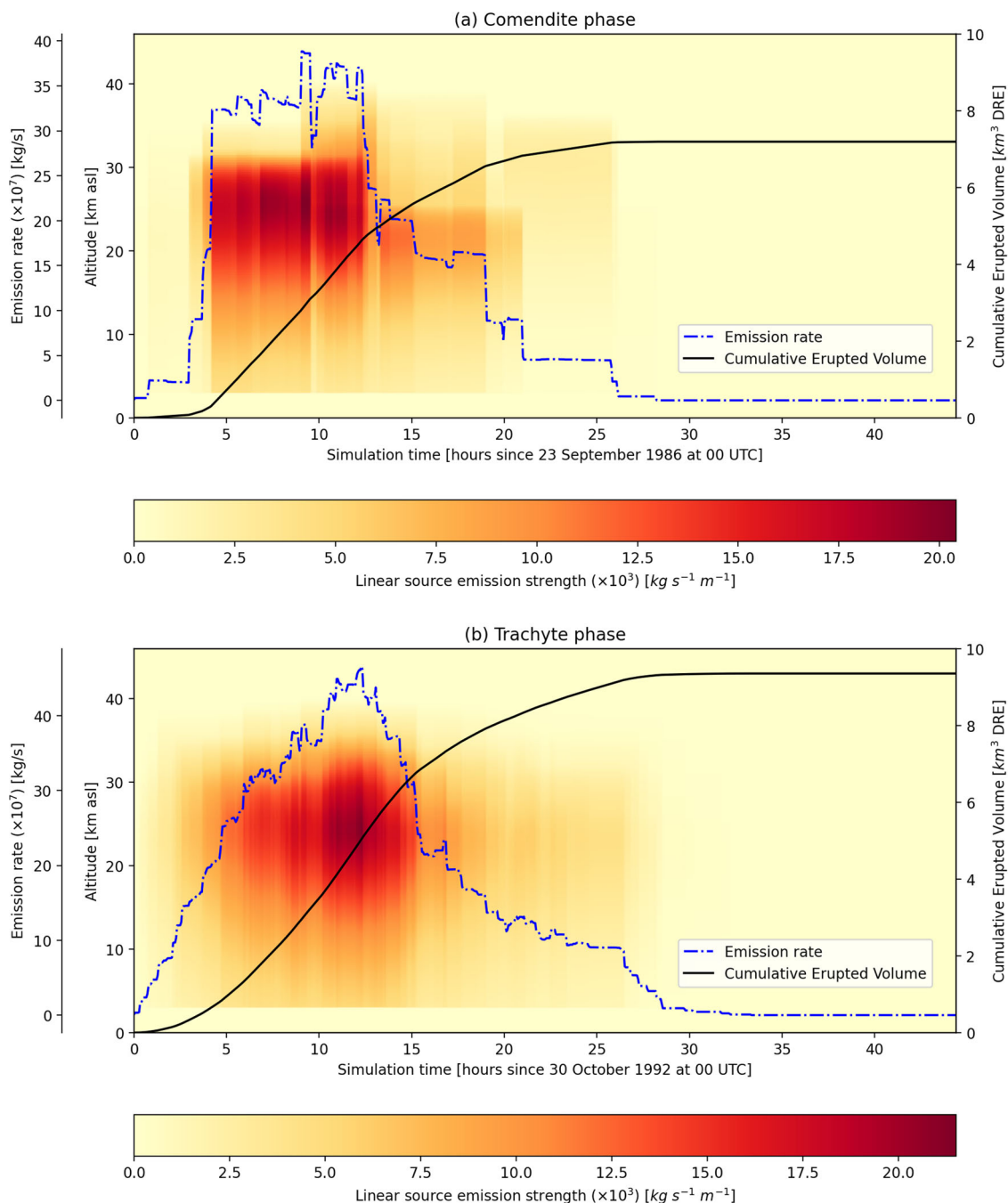


Fig. 4 Time evolution of the eruption columns. Time evolution of the emission source profiles derived for the comenditic first phase (a) and trachytic phase (b) of the Millennium eruption (ME). The plots were created using the open software Matplotlib (<https://matplotlib.org/stable/>).

which represents both the comenditic and trachytic phases. The modelled climate impacts of high latitude eruptions with S loads of <5 Tg are shown to be negligible, especially those in winter as insolation affects the removal of swimschool@abingdon.org.uk-sulfate from the stratosphere⁵⁵. Nonetheless, recent simulations have shown that stratospheric sulfur injections in the extratropics spend longer in the stratosphere than those from low latitudes, and suggest that cooling is more pronounced, but confined to the hemisphere in which it is injected⁴. Stratospheric halogens also exacerbate the radiation forcing of the sulfate⁵⁶, but measured values suggest the syn-eruption halogen release during the ME was low. In addition, the halogens are probably scavenged in the lower plume with only a few percent making it into the stratosphere⁵⁷.

Despite the lack of any significant climate impact, there would have been quite a substantial regional impact linked to the deposition of the thick deposits, lahars, and gas release. It is likely that the eruption led to acid rain, which would have affected crops, livestock, and high F potentially poisoned local freshwater supplies in the downwind regions.

Conclusions

We applied a dual step strategy for solving an inverse problem aimed at reproducing the observed tephra deposit thicknesses generated by the ME of Paektu volcano. The first step uses a fast efficient code for tephra sedimentation based on an analytical solution (PARFIT). The second step considers a recently

Table 2 Eruption parameter model output.

Eruption parameters	Comendite phase	Trachyte phase
Total erupted volume (km ³ DRE)	7.2 (3–17)	9.3 (4–22)
Plume height (km)	30–40	30–40
Peak mass flow rate (10 ⁸ kg/s)	~4	~4
Duration (h)	~15	~20

proposed ensemble-based inversion method with an High Performance Computing (HPC) numerical tephra dispersal model (FALL3D) to reconstruct the eruption dynamics and the volume of magma erupted during the comendite and trachyte phases. Such estimates of past eruption volumes and parameters, and their uncertainties, are key to understanding the impacts on climate. The new inversion method allows us, for the first time, to estimate the evolution of the mass eruption rate with time during the two main phases, and not only the time-averaged values, as commonly done. The total erupted mass of the two Plinian phases are similar, with best estimates of ~7 km³ DRE for the comendite phase and ~9 km³ DRE for the trachyte phase. Considering previous estimations for PDC deposits and intra-caldera filling, we can estimate a total magma volume of about ~23 km³ DRE for the eruption, which corresponds to a M6.7, consistent with other recent estimations. The determination of the contribution of the two different magmas could be used to better estimate the amount of released sulfur once volatile data from the trachyte phase (~9 km³ DRE) is available. The estimations for the comendite phase made with the proposed method are about one third smaller than the reference values used in the literature. A simple scaling suggests between 0.5 to 2.8 Tg syn-eruptive sulfur release, which is in agreement with S loads estimated from ice-core records, and consistent with negligible effects recorded in climate proxy records. This research highlights that in eruptions that tap compositionally distinct melts, with different volatile histories, it is important to understand the temporal relationship between changes in melt chemistry and eruption dynamics so that the likely impact of past eruptions can be accurately assessed.

Methods

Computational methodology

Tephra dispersal model. Tephra dispersal simulations were performed using the latest version (8.2) of the FALL3D-8.2 model^{12,38,48}. FALL3D-8.2 is an Eulerian model for atmospheric transport and deposition based on the solution of the so-called the Advection-Diffusion-Sedimentation (ADS) equations¹². This code has been significantly refactored over the last years, resulting in major improvements to the scalability and performance¹². In addition, the model physics and numerics have been revised in this version^{12,58} to allow aerosol and radionuclide dispersal modelling. The updated model includes new coordinate mapping options, a more efficient and less diffusive solving algorithm, and better memory management and parallelisation strategy based on a full 3-D domain decomposition, suitable for HPC applications. A set of case studies that use the model, including the simulation of the dispersal of long-range fine volcanic ash and volcanic SO₂, and the deposition and dispersal of tephra fallout and radionuclides, were recently presented by Prata et al.⁵⁸

In addition, the capabilities of FALL3D have been extended by making it possible to perform an ensemble of model runs. Ensemble modelling allows model uncertainties, associated with poorly constrained input parameters or errors in the underlying meteorological driver, to be characterised and quantified. Furthermore, it is possible to explore different ensemble-based data assimilation techniques or inversion modelling methods

based on ensemble approaches in order to produce improved simulations of volcanic clouds and tephra deposits consistent with real observations.

Inversion method. We used a novel strategy to model the tephra dispersal of the Millennium Eruption that improves, from a computational point of view, the approach of Costa et al.⁵⁹, which was based on a set of several hundreds of simulations of the FALL3D ash dispersal model¹² and 3D time-dependent meteorological fields across the region of interest. This new approach involves two separate inversion steps. The first step uses a simpler analytical solution^{60,61} (<http://datasim.ov.ingv.it/download/parfit/download-parfit-2.3.1.html>) to identify compatible meteorological conditions that describe the general observed tephra dispersal and estimate the most likely ranges of the Eruption Source Parameters (ESP). These first inversion estimates are then used in the subsequent step built on the ensemble-based inversion algorithm recently proposed by Mingari et al.³⁸ to further constrain dispersal and the ESPs. This dual inversion step method is computationally more efficient than that used by Costa et al.⁵⁹ and thus, allows a very wide range of meteorological conditions and ESP to be explored and allows for more extensive analysis.

First step inversion. The first inversion is based on a semi-analytical solution of the mass conservation equation, which is not computationally intensive and allows tens of thousands of simulations to be explored on a simple desktop computer. We used the most recent version of the PARFIT software package (PARFIT-2.3.1 code: <http://datasim.ov.ingv.it/download/parfit/download-parfit-2.3.1.html>) to solve an inverse volcanological problem, which consists of finding mean wind profile, column height, Total Erupted Mass (TEM), and Total Grain Size Distribution (TGSD), using thickness and grain-size measurements from available eruption deposit outcrops. PARFIT minimises the difference between the simulated and the real deposit loadings. The fitting is performed using a least-squares method which compares measured and calculated deposits thickness and grain-sizes. The function to be minimised⁴² is:

$$\chi^2 = \frac{1}{N - p} \sum_{i=1}^N \theta_i (Y_{\text{obs},i} - Y_{\text{mod},i})^2 \quad (2)$$

where θ_i are weighting factors, N is the number of measurements, p is the number of free parameters, $Y_{\text{obs},i}$ denote the observed ground loads (kg/m²) and $Y_{\text{mod},i}$ the values predicted by the model. The choice of the weighting factors, θ_i , depends upon the distribution of the errors⁴². To accurately constrain the input parameters, it is necessary to know thickness and grain-size values for a large number, N , of locations, such as $N \gg p$. Without a large statistical set of observational data there is a risk of finding parameters that are incorrect, especially when considering limited portions of the deposit⁶¹ that represent the tail of the distribution relative to a particle class^{44,46}. In order to reduce the degrees of freedom, avoiding these problems^{46,47}, the TGSD was estimated following Costa et al.⁶².

Tephra dispersal simulations are performed using the HAZMAP model that is embedded in PARFIT. The HAZMAP program simulates the tephra deposit generated by the sedimentation of volcanic particles from a sustained column^{60,61}. The user defines the ranges and resolution steps of the eruption parameters. This defines a multidimensional grid with combinations of all parameters. PARFIT performs a search on the grid for the minimum χ^2 between the simulated deposit and the field data.

Real wind profiles are used in the PARFIT-2.3.1 model that can be assessed from an ensemble of daily winds over the investigated area. In this case, wind data were obtained from ERA5 Reanalysis

Table 3 Eruption parameter ranges for PARFIT.

Eruption parameter	Range
Total erupted mass (TEM)	Calculated
Plume height	15–45 km
Suzuki coefficient ⁷⁰	1–9
Diffusion coefficient	1000–30,000 m ² /s
TGSD	Following Costa et al. ⁶²
Aggregate density	100–1000 kg/m ³

dataset⁴¹ and converted to the required format using specific routines within the software package. Due to the enhanced computational efficiency, tephra dispersals driven by tens of thousands of daily-averaged wind profiles were investigated to identify the most compatible meteorological conditions within the range explored (50 years, from 1961 to 2010).

For this first-step inversion we considered thickness data associated with the comenditic and trachytic phases, reported in Supplementary Table 1, and searched for the best fit eruption parameters for each distinct phase (similarly to Marti et al.⁶³) in the ranges of values shown in Table 3. Since there is extremely limited grain-size data, we estimated the TGSD using the Costa et al.⁵⁸ parameterisation that is based on magma viscosity and column height. For this purpose, we considered a magma viscosity of about 10⁷ Pa s for a rhyolitic composition^{14,17,19,64} and a plume height of 30 km, a typical value for Plinian eruptions⁴⁹. We used the parameterisation proposed by Bonadonna and Phillips⁶⁵ to characterise densities of particles from rhyolitic magma and used particle shapes that have been measured for other explosive eruptions⁶⁶. Ash aggregation⁶⁷ was considered using the parameterisation of Cornell et al.⁶⁸. The resulting TGSD used at this step is reported in Table 4.

Second-step ensemble-based inversion. As mentioned above, the parameter ranges explored in this second step are derived from the results obtained in the first step. The second step inversion reconstructed the spatial distribution of tephra deposit thickness (i.e. mass loading) associated with the ME using the ensemble-based assimilation method proposed by Mingari et al.³⁸. This algorithm provides a map of a tephra-fall deposit given a dataset of observations of a volcanic deposit at scattered points. To this purpose, an ensemble of model realisations is additionally required. Given a background ensemble of model states characterised by the state vectors x_i^b ($i = 1, \dots, m$), representing m model realisations (ensemble members), along with a vector of observations \vec{y}_o , the state estimate (analysis) is expressed as a linear combinations of the background ensemble:

$$x^a = w_1 x_1^b + \dots + w_m x_m^b \quad (3)$$

where the coefficients $w_i \geq 0$, $i = 1, \dots, m$ are weight factors to be determined which depend on the observations. In our case, the model state vector x is assumed to be the two-dimensional deposit mass loading on a regular grid. In the limit case in which no observations are available, the ensemble members are uniformly weighted according to the constant factors: $w_i = \frac{1}{m}$ and, consequently, the state vector x^a is simply the ensemble mean.

In order to determine the weight factors w_i , a cost function must be minimised in the subspace spanned by the ensemble members with the non-negative constraints $w_i \geq 0$. Let the vector model state \vec{x} be a linear combination of the corresponding ensemble member vectors expressed as in Eq. (3), i.e. $\vec{x} = \sum_i w_i x_i^b$. We look for the vector state minimising the following cost

Table 4 TGSD estimated according to Costa et al.⁶², and accounting for ash aggregation following Cornell et al.⁶⁸, for the PARFIT first step inversion.

(Φ) Diameter (mm)	Density (kg/m ³)	Shape	Percentage
(−5) 32.0	450.0	0.83	1.8
(−4) 16.0	450.0	0.83	1.7
(−3) 8.0	450.0	0.83	2.2
(−2) 4.0	600.0	0.83	2.6
(−1) 2.0	800.0	0.83	2.6
(0) 1.0	900.0	0.83	2.5
(1) 0.5	1100.0	0.83	3.1
(2) 0.25	1300.0	0.83	6.4
(3) 0.125	1500.0	0.83	13.9
(4) 0.0625	1700.0	0.83	10.7
(5) 0.03125	1900.0	0.83	5.3
(2.5) 0.2	300.0	1.00	47.2

function:

$$J(y) \propto (y - \bar{y})^T P^{-1} (y - \bar{y}) + (y_o - y)^T R^{-1} (y_o - y) \quad (4)$$

where the vectors y and \bar{y} are defined as $y = Hx$ and $\bar{y} = H\bar{x}$, being H a linear observation operator which translates the model state into the observation space; P is the ensemble-based error covariance matrix computed in the observation space and we assume a Gaussian distribution with error covariance matrix R for the measurements. In this work, the observation operator interpolates the modelled mass loading into the sampling site and computes the deposit thickness assuming a typical bulk density for consolidated deposits⁴⁴ of $\rho_b = 1000$ kg/m³. This optimisation problem leads to a problem of non-negative quadratic programming that can be solved using an iterative approach. See Mingari et al.³⁸ for further details.

The root-mean-square error (RMSE) was used to measure the differences between the values observed and the analyses interpolated at the sampling site:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n p_i^2 (y_i^{\text{obs}} - y_i^{\text{an}})^2} \quad (5)$$

being y_i^{obs} the i th observation and y_i^{an} the corresponding analysis for each individual phase. Notice that if the non-weighted ($p_i = 1$) version of (1) is used, only very proximal data (i.e. largest deposit thickness measurements) makes a significant contribution to RMSE. Instead, in order to treat the deviations more evenly, the observation errors (ϵ_i) are used to define the weights according to $p_i = 1/\epsilon_i$.

Similarly, we computed the bias using the following definition:

$$\text{Bias} = \frac{1}{n} \sum_{i=1}^n p_i (y_i^{\text{obs}} - y_i^{\text{an}}) \quad (6)$$

and the mean of absolute value of errors (MAE):

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n p_i |y_i^{\text{obs}} - y_i^{\text{an}}| \quad (7)$$

The mean absolute percentage error (MAPE) is also useful:

$$\text{MAPE} = \frac{100}{n} \sum_{i=1}^n \frac{|y_i^{\text{obs}} - y_i^{\text{an}}|}{y_i^{\text{obs}}}$$

where y^{obs} is a vector of 83 observations and y^{an} the vector of the corresponding analysis interpolated to the sampling sites. However, such definition of the error is very sensitive to the observations having very small values (close to zero) and for this reason is preferable using the symmetric mean absolute percentage error (SMAPE), which is another accuracy measure based on relative errors used to evaluate the performance of the

Table 5 FALL3D model configuration.

Model parameter	Value
Simulation time	143 h
Ensemble size	128
Grid resolution	0.23° × 0.16°
Domain size	150 × 150 × 60
Emission source shape	Suzuki distribution ⁷⁰
Species	10 ash bins + 1 aggregate bin
TGSD	Estimated from magma composition and column height (Costa et al. ⁶²)

Table 6 Ensemble configuration.

Model input	Central value	Perturbation range	PDF
Column height	30 km avl	25%	Uniform
Source start	$t = 24 \text{ h}^a$	24 h	Uniform
Source duration	7 h	3 h	Uniform
Suzuki coefficient ⁷⁰ A	4	50%	Uniform
MER	Degruyter ^b	-	-
U Wind	ERA5 ^c	25%	Gaussian
V Wind	ERA5 ^c	25%	Gaussian
TGSD	Costa ^d	-	-
Aggregate density	400 kg m ⁻³	300 kg m ⁻³	Uniform

^aTime in hours after the simulation start.

^bParameterisation from Degruyter and Bonadonna⁷¹.

^cWind field obtained from the single level and model level ERA5 datasets.

^dParameterisation from Costa et al.⁶², considering 10 ash bins from -3 Phi to 5 Phi and ash aggregates (200 micron diameter).

deposit reconstruction:

$$\text{SMAPE} = 100 \times \frac{2}{n} \sum_{i=1}^n \frac{|y_i^{\text{obs}} - y_i^{\text{an}}|}{|y_i^{\text{obs}}| + |y_i^{\text{an}}|} \quad (8)$$

We have also calculated the confidence ratio (CR), defined as the value discriminating the symmetric band on the log-log space of observations vs. analyses, containing 68% of the points represented in Fig. 3 (this percentage will correspond to a sigma if we assume a lognormal distribution). Finally, we considered a composite index (CI), defined as:

$$\text{CI} = \text{RMSE} + \text{Bias} + \text{MAE} + \text{CR} + \text{SMAPE}/100 \quad (9)$$

The values of the evaluation metrics described above are reported in Supplementary Table 2.

Ensemble modelling. The FALL3D computational domain was configured using a horizontal resolution of 0.23° (longitude) and 0.16° (latitude) and a domain size of $n_x \times n_y \times n_z = 150 \times 150 \times 60$ grid cells. Numerical simulations used the meteorological fields estimated using the PARFIT-2.3.1 code. Among the different meteorological configurations selected in the first step procedure, the conditions corresponding to 25–29 January 1962, for the comendite phase, and 29 February–3 March 1976 for the trachyte phase, showed the best agreement with observations during the second step (i.e. when using the ensemble-based FALL3D; see Supplementary Table 2). The Total Grain Size Distribution (TGSD) was computed from the column height using the parameterisation proposed by Costa et al.⁶² considering 10 ash bins with particle diameters ranging between -2Φ and 8Φ . Moreover, an additional bin aggregate of diameter $d_{\text{agg}} = 200 \mu\text{m}$ was

considered. Table 5 provides further details about the model configuration.

A background ensemble with 128 members was generated by perturbing Eruption Source Parameters (ESP), the horizontal wind components, and the aggregate density using either uniform or truncated normal distributions^{13,48}. A Latin Hypercube Sampling (LHS)^{13,69} was used to efficiently sample the parameter space. The ensemble configuration requires defining a central value (i.e. results from the first inversion step), a perturbation range and a probability density function (PDF) for each parameter to be perturbed. The ensemble is built by randomly sampling the corresponding PDF. Table 6 provides a summary of the configuration used to define the background ensemble along with the list of the perturbed parameters and model inputs.

The emission start time was uniformly sampled between $t = 0$ and $t = 48 \text{ h}$, where t is the time after the simulation starts in hours, as detailed in Table 6. The emission source duration, Δt , was sampled in the range from 4 to 10 h, which means that emission is only possible for $0 \leq t \leq 58 \text{ h}$. Since we aim to simulate both phases of the ME (i.e. November 946 CE and February 947 CE events), we invert independently the two phases and consider the formation of the total tephra fallout deposit as the sum of the first and second eruptive event for validation purposes.

Source term inversion. FALL3D solves an almost linear problem with weak nonlinearity effects (e.g., due to gravity current, wet deposition, or aggregation). Consequently, a re-scaling of the emission source term according to $s_i \rightarrow w_i s_i$ leads to a deposit mass loading re-scaled correspondingly: $x_i^b \rightarrow w_i x_i^b$, where s_i is the emission source term in $\text{kg m}^{-3} \text{ s}^{-1}$ for the i th ensemble member. This is the mass injected into the computational domain per unit time and depends on time and height above the vent. The analysed tephra deposit (mass loading) on the ground is described by Eq. (3) that gives:

$$x^a = \sum_i x_i^b \quad (10)$$

i.e., the analysed mass loading is a result from the contribution of multiple members of an ensemble defined by the first-guess emission source term s_i rescaled by w_i .

Data availability

The data used in this study are all published data and all the sources of information are reported in Supplementary Table 1 and cited in the Supplementary Reference section. Data sharing not applicable to this article as no original datasets were generated or analysed during the current study.

Code availability

The Fall3D configuration files and scripts used for the simulations are available at <https://doi.org/10.5281/zenodo.10159965>.

Received: 5 June 2023; Accepted: 5 December 2023;

Published online: 02 January 2024

References

- McCormick, M. P., Thomason, L. W. & Trepte, C. R. Atmospheric effects of the Mt-Pinatubo eruption. *Nature* **373**, 399–404 (1995).
- Robock, A. Volcanic eruptions and climate. *Rev. Geophys.* **38**, 191–219 (2000).
- Timmreck, C. Modeling the climatic effects of large explosive volcanic eruptions. *WIREs Clim. Change* **3**, 545–564 (2012).
- Toohey, M. et al. Disproportionately strong climate forcing from extratropical explosive volcanic eruptions. *Nat. Geosci.* **12**, 100–107 (2019).
- Marshall, L. R. et al. Volcanic effects on climate: recent advances and future avenues. *Bull. Volcanol.* **84**, 54 (2022).

6. Smith, V. C. et al. The magnitude and impact of the 431 CE Tierra Blanca Joven eruption of Ilopango, El Salvador. *Proc. Natl Acad. Sci. USA* **117**, 26061–26068 (2020).
7. Stoffel, M. et al. Estimates of volcanic-induced cooling in the Northern Hemisphere over the past 1,500 years. *Nat. Geosci.* **8**, 784–788 (2015).
8. Sun, C. et al. Ash from Changbaishan millennium eruption recorded in Greenland ice: implications for determining the eruption's timing and impact. *Geophys. Res. Lett.* **41**, 694–701 (2014).
9. Xu, J. D. et al. Climatic impact of the Millennium eruption of Changbaishan volcano in China: new insights from high-precision radiocarbon wigglematch dating. *Geophys. Res. Lett.* **40**, 54–59 (2013).
10. Oppenheimer, C. et al. Multi-proxy dating the Millennium eruption of Changbaishan to late 946 CE. *Quat. Sci. Rev.* **158**, 164–171 (2017).
11. Oppenheimer, C. Climatic, environmental and human consequences of the largest known historic eruption: Tambora volcano (Indonesia) 1815. *Prog. Phys. Geogr. Earth Environ.* **27**, 230–259 (2003).
12. Folch, A. et al. FALL3D-8.0: a computational model for atmospheric transport and deposition of particles, aerosols and radionuclides—part 1: model physics and numerics. *Geosci. Model Dev.* **13**, 1431–1458 (2020).
13. Folch, A., Mingari, L. & Prata, A. T. Ensemble-based forecast of volcanic clouds using FALL3D-8.1. *Front. Earth Sci.* **9**, 74 (1841).
14. Iacovino, K. et al. Quantifying gas emissions from the “Millennium Eruption” of Paektu volcano, Democratic People's Republic of Korea/China. *Sci. Adv.* **2**, e1600913 (2016).
15. Machida, H. & Arai, F. Extensive ash falls in and around the Sea of Japan from large late Quaternary eruptions. *J. Volcanol. Geotherm. Res.* **18**, 151–164 (1983).
16. Asano, G. Several discoveries obtained from the Hakutosan (Baegdusan). Expedition in 1942–1943. *Mineral Geol.* **1**, 228–231 (1948).
17. McLean, D. et al. Identification of the Changbaishan ‘Millennium’ (B-Tm) eruption deposit in the Lake Suigetsu (SG06) sedimentary archive, Japan: synchronisation of hemispheric-wide palaeoclimate archives. *Quat. Sci. Rev.* **150**, 301–307 (2016).
18. Machida, H., Moriwaki, H. & Zhao, D. The recent major eruption of Changbaishan volcano and its environmental effects. *Geogr. Rep. Tokyo Metropol. Univ.* **25**, 1–20 (1990).
19. Zhang, M. et al. The intraplate Changbaishan volcanic field (China/North Korea): a review on eruptive history, magma genesis, geodynamic significance, recent dynamics and potential hazards. *Earth Sci. Rev.* **187**, 19–52 (2018).
20. Yun, S. H., Won, C. K. & Lee, M. W. Cenozoic volcanic activity and petrochemistry of volcanic rocks in the Mt. Paektu area. *J. Geol. Soc. Korea* **29**, 291–307 (1993).
21. Pan, B., de Silva, S. L., Xu, J., Liu, S. & Xu, D. Late Pleistocene to present day eruptive history of the Changbaishan-Tianchi Volcano China/ DPRK: new field, geochronological and chemical constraints. *J. Volcanol. Geotherm. Res.* **399**, 106870 (2020).
22. Ikehara, K. Late Quaternary seasonal sea-ice history of the northern Japan Sea. *J. Oceanog.* **59**, 585–593 (2003).
23. Derkachev, A. N. et al. Tephra layers of large explosive eruptions of Baitoushan/Changbaishan Volcano in the Japan Sea sediments. *Quat. Int.* **519**, 200–214 (2019).
24. McLean, D. et al. Refining the eruptive history of Ulleungdo and Changbaishan volcanoes (East Asia) over the last 86 kyrs using distal sedimentary records. *J. Volcanol. Geotherm. Res.* **389**, 301–307 (2020).
25. Yun, S. H. Volcanological interpretation of historical eruptions of Mt. Baekdusan volcano. *J. Korean Earth Sci. Soc.* **34**, 456–469 (2013).
26. Wei, H. et al. Timescale and evolution of the intracontinental Tianchi volcanic shield and ignimbrite-forming eruption, Changbaishan, Northeast China. *Lithos* **96**, 315–324 (2007).
27. Miyake, F., Nagaya, K., Masuda, K. & Nakamura, T. A signature of cosmic-ray increase in AD 774–775 from tree rings in Japan. *Nature* **486**, 240–242 (2012).
28. Sigl, M. et al. Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature* **523**, 543–549 (2015).
29. Chōng Inji, K. Seoul National University Library [in Chinese, 高麗史] (1451).
30. Chen, Z. Q. & Chen, Z. Identifying references to volcanic eruptions in Chinese historical records. *Geol. Soc. Lond. Spec. Publ.* **510**, 271–289 (2021).
31. Tsuboi, K. 坪井九馬三. Kohfukuiji Chronicle. Handwritten copy from National Diet Library Collection. <https://dl.ndl.go.jp/>. (1908).
32. Fujiwara Tadahira [藤原忠平] 880–949, Teishin kō ki (Journal of Fujiwara Tadahira) [貞信公記]. in *Dai Nihon kokiroku* [大日本古記録], Vol. 1 (eds Tokyo Daigaku Shiryō Hensanjō) (Iwanami Shoten, 1952–1986). Handwritten copy from National Diet Library Collection. <https://dl.ndl.go.jp/>.
33. Horn, S. & Schmincke, H.-U. Volatile emission during the eruption of Baitoushan Volcano (China/North Korea) ca. 969 AD. *Bull. Volcanol.* **61**, 537–555 (2000).
34. Jin, B. & Zhang, X. Y. *Researching Volcanic Geology in Changbai Mt. (in Chinese)* (Northeast Korean Nationality Education Press, 1994).
35. Yang, Q. F. *The Distribution of Pyroclastic Flow on Changbaishan Tianchi Volcano*. Project Report of Seismological Bureau of Jilin Province (in Chinese) (2005).
36. Yang, Q. et al. The Millennium eruption of Changbaishan Tianchi Volcano is VEI 6, not 7. *Bull. Volcanol.* **83**, 74 (2021).
37. Zhao, B., Xu, J., Yu, H. & Chen, Z. Pyroclastic density current facies in the millennium eruption of Tianchi volcano northeast china: insights from topography, stratigraphy, granulometry and petrography. *Front. Earth Sci.* **8**, 72–91 (2020).
38. Mingari, L., Costa, A., Macedonio, G., & Folch, A. Reconstructing tephra fall deposits via ensemble-based data assimilation techniques. *Geosci. Model Dev.* **16**, 3459–3478 (2023).
39. Blockley, S. P. E. et al. A new and less destructive laboratory procedure for the physical separation of distal glass tephra shards from sediments. *Quat. Sci. Rev.* **24**, 1952–1960 (2005).
40. McLean, D. et al. Integrating the Holocene tephrostratigraphy for East Asia using a high-resolution cryptotephra study from Lake Suigetsu (SG14 core), Central Japan. *Quat. Sci. Rev.* **183**, 36–58 (2018).
41. Hersbach, H. et al. ERA5 hourly data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). <https://doi.org/10.24381/cds.adbb2d47> (2018).
42. Costa, A. et al. Tephra fallout hazard assessment at Campi Flegrei caldera (Italy). *Bull. Volcanol.* **71**, 259–273 (2009).
43. Armienti, P., Macedonio, G. & Pareschi, M. A numerical model for the simulation of tephra transport and deposition: applications to May 18, 1980 Mt.St. Helens eruption. *J. Geophys. Res.* **93**, 6463–6476 (1988).
44. Bonadonna, C., Biass, S. & Costa, A. Physical characterization of explosive volcanic eruptions based on tephra deposits: propagation of uncertainties and sensitivity analysis. *J. Volcanol. Geotherm. Res.* **296**, 80–100 (2015).
45. Constantinescu, R. et al. Uncertainty quantification of eruption source parameters estimated from tephra fall deposits. *Geophys. Res. Lett.* **49**, e2021GL097425 (2022).
46. Bonasia, R., Macedonio, G., Costa, A., Mele, D. & Sulpizio, R. Numerical inversion and analysis of tephra fallout deposits from the 472 AD sub-Plinian eruption at vesuvius (Italy) through a new best-fit procedure. *J. Volcanol. Geotherm. Res.* **189**, 238–246 (2010).
47. Volentik, A. C. M., Bonadonna, C., Connor, C. B., Connor, L. J. & Rosi, M. Modeling tephra dispersal in absence of wind: Insights from the climactic phase of the 2450 BP Plinian eruption of Pululagua volcano (Ecuador). *J. Volcanol. Geotherm. Res.* **193**, 117–136 (2010).
48. Mingari, L. et al. Data assimilation of volcanic aerosol observations using FALL3D + PDAF. *Atmos. Chem. Phys.* **22**, 1773–1792 (2022).
49. Bonadonna, C. & Costa, A. Plume height, volume, and classification of volcanic explosive eruptions based on the Weibull function. *Bull. Volcanol.* **75**, 1–19 (2013).
50. Holasek, R., Self, S. & Woods, A. Satellite observations and interpretation of the 1991 Mount Pinatubo eruption plumes. *J. Geophys. Res.* **101**, 27,635–27,655 (1996).
51. Costa, A., Folch, A. & Macedonio, G. Density-driven transport in the umbrella region of volcanic clouds: Implications for tephra dispersion models. *Geophys. Res. Lett.* **40**, 4823–4827 (2013).
52. Scaillet, B. & Oppenheimer, C. Reassessment of the sulfur and halogen emissions from the Millennium Eruption of Changbaishan (Paektu) volcano. *J. Volcanol. Geotherm. Res.* **442**, 107909 (2023).
53. Gerlach, T. M., Westrich, H. R. & Symonds, R. B. Preeruption vapor in magma of the climactic Mount Pinatubo eruption: source of the giant stratospheric sulfur dioxide cloud. in *Fire and Mud: Eruptions and Lahars of Mount Pinatubo, Philippines* (eds Newhall, N. G. & Punongbayan, R. S.) 415–433 (University of Washington Press, 1996).
54. Sigl, M., Toohey, M., McConnell, J. R., Cole-Dai, J. & Severi, M. Volcanic stratospheric sulfur injections and aerosol optical depth during the Holocene (past 11 500 years) from a bipolar ice-core array. *Earth Syst. Sci. Data* **14**, 3167–3196 (2022).
55. Kravitz, B. & Robock, A. Climate effects of high-latitude volcanic eruptions: Role of the time of year. *J. Geophys. Res.* **116**, D01105 (2011).
56. Staunton-Sykes, J. et al. Co-emission of volcanic sulfur and halogens amplifies volcanic effective radiative forcing. *Atmos. Chem. Phys.* **21**, 9009–9029 (2021).
57. Wade, D. C. et al. Reconciling the climate and ozone response to the 1257 CE Mount Samalás eruption. *Proc. Natl Acad. Sci.* **117**, 26651–26659 (2020).
58. Prata, A. T., Mingari, L., Folch, A., Macedonio, G. & Costa, A. FALL3D-8.0: a computational model for atmospheric transport and deposition of particles, aerosols and radionuclides—part 2: model validation. *Geosci. Model Dev.* **14**, 409–436 (2021).
59. Costa, A. et al. Quantifying volcanic ash dispersal and impact from Campanian Ignimbrite super-eruption. *Geophys. Res. Lett.* **39**, L10310 (2012).
60. Macedonio, G., Costa, A. & Longo, A. A computer model for volcanic ash fallout and assessment of subsequent hazard. *Comput. Geosci.* **31**, 837–845 (2005).

61. Pfeiffer, T., Costa, A. & Macedonio, G. A model for the numerical simulation of tephra fall deposits. *J. Volcanol. Geotherm. Res.* **140**, 237–294 (2005).
62. Costa, A., Pioli, L. & Bonadonna, C. Assessing tephra total grain-size distribution: insights from field data analysis. *Earth Planet. Sci. Lett.* **443**, 90–107 (2016).
63. Marti, A., Folch, A., Costa, A. & Engwell, S. Reconstructing the plinian and coignimbrite sources of large volcanic eruptions: a novel approach for the Campanian Ignimbrite. *Sci. Rep.* **6**, 21220 (2016).
64. Yi, J. et al. Modeling the multi-level plumbing system of the Changbaishan caldera from geochemical, mineralogical, Sr-Nd isotopic and integrated geophysical data. *Geosci. Front.* **12**, 101171 (2021).
65. Bonadonna, C. & Phillips, J. C. Sedimentation from strong volcanic plumes. *J. Geophys. Res.* **108**, 2340 (2003).
66. Mele, D. et al. Total grain size distribution of components of fallout deposits and implications for magma fragmentation mechanisms: examples from Campi Flegrei caldera (Italy). *Bull. Volcanol.* **82**, 31 (2020).
67. Costa, A., Folch, A. & Macedonio, G. A model for wet aggregation of ash particles in volcanic plumes and clouds: I. Theoretical formulation. *J. Geophys. Res.* **115**, B09201 (2010).
68. Cornell, W., Carey, S. & Sigurdsson, H. Computer simulation and transport of the Campanian Y–5 ash. *J. Volcanol. Geotherm. Res.* **17**, 89–109 (1983).
69. McKay, M. D., Beckman, R. J. & Conover, W. J. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21**, 239–245 (1979).
70. Suzuki, T. A theoretical model for dispersion of tephra. In *Arc Volcanism: Physics and Tectonics*, (eds Shimozuru, D. & Yokoyama, I.) pp 93–113 (Terra Scientific Publishing Company (TERRAPUB), Tokyo, 1983).
71. Degruyter, W. & Bonadonna, C. Improving on mass flow rate estimates of volcanic eruptions. *Geophys. Res. Lett.* **39**, L16308 (2012).

Acknowledgements

A.C., G.M. and L.M. thank the Istituto Nazionale di Geofisica e Vulcanologia, Italy, grant Progetto INGV Pianeta Dinamico (code CUP D53J19000170001) funded by Italian Ministry MIUR (“Fondo Finalizzato al rilancio degli investimenti delle amministrazioni centrali dello Stato e allo sviluppo del Paese”, legge 145/2018). This work has been partially funded by the Centre of Excellence for Exascale in Solid Earth (ChEESSE; EuroHPC Grant Agreement No. 101093038), the project “A Digital Twin for GEO-physical extremes” (DT-GEO; Horizon Europe Grant Agreement No. 101058129), and the Meteorological/Earthquake See-At Technology Development Research Grant KMI2018-02710. We acknowledge European Centre for Medium-range Weather Forecast (ECMWF) for furnishing the meteorological re-analysis ERA-5 data (no permissions were required). We thank two anonymous reviewers and Céline Vidal for constructive criticism and feedback that improved this manuscript.

Author contributions

Conceptualisation: A.C. and V.C.S.; Methodology: A.C., L.M., and V.C.S.; Software: L.M., A.F., G.M., and A.C.; Resources: L.M., V.C.S., D.M., J.L., and S.H.Y.; Writing—original draft: A.C., L.M., and V.C.S.; Writing—review and editing: A.C., L.M., V.C.S., G.M., D.M., A.F., J.L., and S.H.Y.; Visualisation: L.M.; Supervision: A.C. and S.H.Y.; Funding acquisition: A.C. and A.F. All authors have read and approved the final version of the manuscript.

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/s43247-023-01162-0>.

Correspondence and requests for materials should be addressed to Antonio Costa.

Peer review information *Communications Earth & Environment* thanks Céline Vidal and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Joe Aslin and Carolina Ortiz Guerrero. A peer review file is available

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024