RESEARCH ARTICLE

VOLCANICA

Estimation and uncertainty quantification of magma interaction times using statistical emulation

Luca Insolia*[α], Stéphane Guerrier[α,β], Chiara P. Montagna[γ], Maria-Pia Victoria-Feser[α,δ], and Luca Caricchi[ε]

[α] Geneva School of Economics and Management, University of Geneva, Switzerland.
[β] Faculty of Science, University of Geneva, Switzerland.
[γ] Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Pisa, Italy.
[δ] Department of Statistical Sciences, University of Bologna, Italy.
[ε] Department of Earth Sciences, University of Geneva, Switzerland.

Submitted: 2023-01-20 | Accepted: 2024-06-06 | Published: 2024-08-27

Editor: K. Chamberlain | Typesetter: J. Farquharson

Abstract

Evolution of volcanic plumbing systems towards eruptions of different styles and sizes largely depends on processes at crustal depths that are outside our observational capabilities. These processes can be modelled and the outputs of the simulations can be compared with the chemistry of the erupted products, geophysical and geodetic data to retrieve information on the architecture of the plumbing system and the processes leading to eruption. The interaction between magmas with different physical and chemical properties often precedes volcanic eruptions. Thus, sophisticated numerical models have been developed that describe in detail the dynamics of interacting magmas, specifically aimed at evaluating pre-eruptive magma mingling and mixing timescales. However, our ability to explore the parameter space in order to match petrological and geophysical observations is limited by the extremely high computational costs of these multiphase, multicomponent computational fluid dynamics simulations. To overcome these limitations, we present a statistical emulator that is able to reproduce the numerical simulations results providing the temporal evolution of the distribution of magma chemistry as a function of a set of input parameters such as magma densities and reservoir shapes. The whole rock composition of volcanic rocks is one of the most common measurable parameter collected for eruptions. The statistical emulator can be used to invert the observed distribution of whole rock chemistry to determine the duration of interaction between magmas preceding an eruption and identify the best matching input parameters of the numerical model. Importantly, the statistical emulator intrinsically includes error propagation, thus providing confidence intervals on predicted interaction timescales on the base of the intrinsic uncertainty of the input parameters of the numerical simulations.

Keywords: Volcanology; Petrology; Magma Dynamics; Statistical Emulation.

1 Introduction

Magmatic processes preceding volcanic eruptions occur at depths of kilometres to tens of kilometres and are therefore not directly accessible: the only direct information about pre-eruptive magma storage conditions is provided by the volcanic rock record. Thus, textural [Cashman and Marsh 1988; Higgins 2000; Zellmer et al. 2020], rheological [Lejeune and Richet 1995; Hess and Dingwell 1996; Lavallée et al. 2007; Pistone et al. 2012; Heap et al. 2020], and chemical [Marsh 1981; Blundy and Cashman 2001; Davidson et al. 2007; Putirka 2008; Edmonds et al. 2010; Humphreys et al. 2012; Cooper and Kent 2014; Neave et al. 2014; Cassidy et al. 2016; Stock et al. 2018; Edmonds et al. 2019; Maclennan 2019; Zellmer et al. 2020] characterisation of volcanic rocks is central to our understanding of the processes leading to volcanic eruptions. However, to retrieve quantitative information on pre-eruptive processes from the analysis of the rock record, it is fundamental first to compare the chemistry of minerals, rocks, and glasses with those produced experimentally [Nimis 1995; Putirka 2008; Petrelli et al. 2020] and then to compare the analyses with the results of experiments or numerical modelling [e.g. Annen et al. 2006; Caricchi et al. 2014; Perugini et al. 2015; Jackson et al. 2018]. We will focus here on the comparison between volcanic rock record and numerical modelling results.

Physical modelling is a fundamental tool to link observable quantities at the Earth surface and magmatic processes at depth [e.g. Bagagli et al. 2017; La Spina et al. 2022]. Modelling efforts have been directed both at tackling the complex thermochemistry of multiphase magmas [e.g. Gualda et al. 2012; Bohrson et al. 2014; Keller and Suckale 2019; Rummel et al. 2020], in order to be able to replicate observed geochemical patterns in erupted products (crystals, melt inclusions [e.g. Caricchi et al. 2018; Wei et al. 2022]) as well as emitted gas chemistry and fluxes [e.g. Chiodini et al. 2012], and at replicating observed geophysical signals such as ground deformation, gravity anomaly, and seismicity recorded at active volcanoes [e.g. Poland and de Zeeuw-van Dalfsen 2021; Matoza and Roman 2022]. Inclusion of all relevant physics is often impossible, given the wide variety of length and time scales that characterise volcanic systems at large; therefore, different approximations are typically used to target specific questions: homogeneous reservoirs are often assumed in order to model ground deformation [e.g. Mogi 1958; Zhong et al. 2019]; magma motion is usually neglected when modelling crystal growth patterns [e.g. Iovine et al. 2017; Petrone et al. 2018]. Necessity of simplifying assumptions is also motivated by the very high computational costs of running large-scale, detailed

* luca.insolia@unige.ch

simulations of the overall magmatic system evolution [Garg and Papale 2022]. This also largely limits our ability to explore the wide range of initial conditions and input parameters that characterise magmatic systems, such as temperatures and pressures/depths of melt reservoirs, as well as their size, shape, and melt chemistry, and volatile content. Moreover, as properties of magma reservoirs at depth cannot be measured, models' initial conditions are typically characterised by large uncertainties which are—in turn—difficult to define.

Due to the necessary assumptions and uncertainties associated with the range of some of the input parameters of physical models, it is often impossible to quantitatively assess the impact of initial assumptions on the final results of the models. This would require the realisation of millions of simulations to systematically investigate the full range of input parameters. Additionally, the use of complex physical models during volcanic unrest can be unpractical as the realisation of simulations can be time-intensive and the integration of continuous streams of data not straightforward [Selva et al. 2012].

This is where the combination of physical modelling and statistical emulation becomes essential. Using the outputs of simulated physical states and the corresponding values for the theoretical parameters—i.e. simulated data produced by a simulator given a set of input parameters—a Statistical Emulator (SE) allows one to predict the physical model output given a specific input, along with an associated measure of the uncertainty on this prediction. More precisely, an SE is a probabilistic model that links any output of a simulator, often summary statistics of the simulated data, produced by a theoretical physical model (i.e. the simulator), to the parameter values used to produce the simulation results. Once trained on a series of simulated data, the SE can be used on observed data to calibrate the possible values of some of the input parameters of the underlying physical process. Statistical emulation is a very popular method in environmental sciences. While Sacks et al. [1989], Currin et al. [1991], Kennedy and O'Hagan [2001], Santner et al. [2003], and Higdon et al. [2008] use a Gaussian process to model the uncertainty, this requires long computational times and a large number of SEs; in practice, a (multivariate) regression model is often preferred (see e.g. Salter and Williamson [2016] for a discussion and references). Notably, different SEs that rely on various statistical assumptions are regularly proposed, and they depend both on the physical model and the available data; see e.g. Mahmood et al. [2015], Guillas et al. [2018], and Yang et al. [2020] to cite just a few.

In this paper we present results from a numerical model of magma mingling, tailored for the Campi Flegrei magmatic system (Italy) [Montagna et al. 2022]. We show how silicate melt composition evolves as two parental magmas interact in a relatively shallow reservoir at 3 km depth, approaching compositional homogeneity [Montagna and Papale 2018], for different setups characterised by varying reservoir shape and magmatic volatile contents. An SE is then trained on these results; we present the high accuracy it achieves in reproducing simulated results, even at large down-sampling, and the error quantification associated with this method.

2 Magma dynamics models

The very existence of magma within the Earth's crust requires prolonged input of mass and energy (i.e. heat) from the mantle source [Glazner et al. 2004; Annen et al. 2006; Karakas et al. 2017]. Thus, the progressive assembly of volcanic plumbing systems naturally results in the interaction between resident, and generally colder, magma and hot, generally more primitive magma rising through the crust. Abundant field evidence exists for this process, both in the plutonic and volcanic rock records [Blundy and Sparks 1992; Perugini and Poli 2005; Morgavi et al. 2016; Ridolfi et al. 2016; Morgavi et al. 2017], and in some cases such interaction can culminate in a volcanic eruption [Caricchi et al. 2021]. Examples of magma input triggering volcanic activity can be identified at Campi Flegrei, Santorini (Greece), Yellowstone Caldera (USA), and Long Valley Caldera (USA) [Wark et al. 2007; Druitt et al. 2012; Morgavi et al. 2016; Forni et al. 2018], and approaches have been proposed to determine the duration of magma interaction preceding an eruption [Perugini et al. 2015]. Here, we simulate the dynamical evolution of a melt-dominated magmatic reservoir subject to the injection of deeply-sourced, more primitive, more volatile-rich (buoyant) magma [Montagna et al. 2015; 2022], focusing on the quantification of the time scales of magmatic convection and mingling.

2.1 Physical model setup

Numerical simulations of magma chamber replenishment were performed using the finite-element C++ code GALES [Longo et al. 2012a; Garg et al. 2018a; b]. The physico-mathematical model describes the space-time dynamics of a multicomponent mixture consisting of silicate melt in thermodynamical equilibrium with a gas phase at the local conditions of pressure, temperature and composition [Longo et al. 2012b; Papale et al. 2017; Garg et al. 2019]. The conservation equations for the mass of single components and momentum of the whole mixture are solved together with the gas-liquid thermodynamic equilibrium model [Papale et al. 2006] and the constitutive equations for multiphase, multicomponent mixture properties density and viscosity [Reid et al. 1977; Ishii and Zuber 1979; Lange 1997; Giordano et al. 2008].

We explore buoyant magma mixing and its timescales based on the archetypal case of the Campi Flegrei volcanic system, where occurrences of interacting magmas have been widely testified [Tonarini et al. 2009; Arienzo et al. 2010; Fourmentraux et al. 2012; Forni et al. 2018]. We model the injection of CO_2-rich, shoshonitic magma coming from a deep reservoir into a shallower, much smaller one, containing more evolved and partially degassed phonolitic magma [Mangiacapra et al. 2008; Arienzo et al. 2009; Di Renzo et al. 2011]. The interaction dynamics is solely driven by buoyancy, stemming from the initial unstable density contrast between the shoshonitic and phonolitic end-members [Longo et al. 2023].

Different geometries and volatile contents of the shallow chamber were considered to account for a range of possible conditions at Campi Flegrei. Specifically, the shallow reservoir is oblate (Simulations 1 and 4 in Table 1), circular (Simulation 3 in Table 1), or prolate (Simulations 2 and 5 in Table 1); the surface area is kept constant. Total (dissolved and excess

fluid) volatile contents in the phonolitic magma are 0.3 wt.% CO_2 and 2.5 wt.% H_2O for Simulations 1, 2, and 3 in Table 1, while for Simulations 4 and 5 they are set to 0.1 wt.% CO_2 and 1 wt.% H_2O [Mangiacapra et al. 2008]. Figure 1 shows the simulated domain and summarises the conditions for the numerical simulations; the deep chamber contains a shoshonitic magma with 2 wt.% H_2O and 1 wt.% CO_2.

The different settings in Figure 1 can be represented by two meaningful quantities, the initial density difference at the interface among the two end-member magmas (shallow phonolite and deep shoshonite) $\Delta\rho$, and the aspect ratio of the shallow reservoir a. Table 1 shows the values of these parameters for the five simulated scenarios. The initial density contrast is due to different volatile contents in the shallow reservoir, as well as different phase partitioning in volatiles. The latter depends on pressure, in turn a function of the depth of the interface which changes according to reservoir shape.

Table 1: Values of the input parameters for each simulated scenario.

	Aspect ratio a	Density difference $\Delta\rho$ [kg m^{-3}]
Simulation 1	0.5	35
Simulation 2	2.0	20
Simulation 3	1.0	30
Simulation 4	0.5	160
Simulation 5	2.0	130

2.2 Magma mingling

As the two magmas come into contact at time 0, a gravitational Rayleigh-Taylor instability [Chandrasekhar 2013] develops at the interface because of the density difference $\Delta\rho$ between the denser magma in the shallow reservoir and the volatile-rich and less dense injected magma. Figure 2 shows the space-time evolution of the composition within the shallower, initially phonolitic reservoir where shoshonitic magma is injected.

Plumes of light magma rise into the shallow reservoir and trigger convective patterns enhancing magma mixing, while a portion of the degassed phonolitic magma initially hosted in the shallow reservoir sinks through the feeding dyke. With time, less dense shoshonitic magma tends to accumulate at the top of the system, and a density-stratified magma chamber develops, more so when convection is more efficient, i.e. at small aspect ratios and large initial density contrasts (Figure 2) [Montagna et al. 2015]. The space-time evolution of composition within the reservoir depends both on its aspect ratio a and on the initial density contrast $\Delta\rho$ (Figure 2), but in all cases it testifies for the progressive homogenisation of magma chemistry [Montagna et al. 2015]. The time duration and efficiency of the interaction only depend on the initial conditions, i.e. $\Delta\rho$ and a [Papale et al. 2017].

During the magmatic interaction described above, an eruption can occur, sampling the magma from the reservoir. Here we assess the capacity of the distribution of erupted magma chemistry to return information on the duration of interaction preceding the eruption. This sort of information is essential for determining, from the study of past eruptions triggered by the interaction between magmas, the amount of time available between potential monitoring signs of magma input into a relatively shallow reservoir and an eruption [Bagagli et al. 2017]. Instead of performing a large number of numerical simulations, we resort to statistical emulation. We use the distribution of magma chemistry as a function of time obtained from the numerical simulations to fit the SE. The final target of this approach is to build a model which, on the basis of the distribution of erupted magma chemistry, provides the timescales of magmatic interaction that preceded the eruption, and the associated uncertainties.

Magmatic composition in the model is defined as the proportion of one end-member (shoshonite) with respect to the total, thus it varies between 0 and 1. End-member compositions are in turn defined by the ten major oxides (SiO_2, TiO_2, AlO_2, Fe_2O_3, FeO, MnO, MgO, CaO, Na_2O, K_2O [Montagna et al. 2015]), plus volatiles (main components water and CO_2 only). The timescales of the dynamics described by these simulations are very short compared to typical diffusive timescales for oxides in magmas at our resolution (minimum grid size is ~1 m [Zhang and Gan 2022]), therefore the approximation of mingling holds and the modelled compositions compare with whole rock datasets. The oxides compositions of the two end-member magmatic components, as well as their volatile contents partitioned between liquid and exsolved phases, are in turn used in the model to evaluate volatile phase partitioning, and density and viscosity of the fluid, thus they have a significant impact on the evolution of the system [Longo et al. 2012a; Montagna et al. 2015; Longo et al. 2023].

3 A Statistical Emulator

An SE is a probabilistic model that links the output of a simulator ($\mathbf{D} \in \mathcal{D} \subset \mathbb{R}^{K_1}$), or a suitable function of the simulated data ($\mathbf{Y} = \mathbf{Y}(\mathbf{D}) \in \mathcal{Y} \subseteq \mathbb{R}^{K_2}$), with $K_2 \ll K_1$, produced by a theoretical (complex) model (simulator), to the parameter values ($\mathbf{x} \in \mathcal{X} \subseteq \mathbb{R}^{K_3}$) used to produce the simulation results. Very generally, one considers theoretical models that are complex enough so that simulating one possible outcome, i.e. the outcome obtained with one (set of) value(s) for the parameters governing the theoretical model (i.e. $\mathbf{x}_i \in \mathcal{X} \subseteq \mathbb{R}^{K_3}$), is computationally very demanding. In these cases, the simulator cannot simulate information for all sets of possible parameters $\mathbf{x}_i$, but instead can simulate from a finite sample, say $\mathbf{x}_i$ for $i = 1, \ldots, m$, of carefully chosen combinations of values for the $\mathbf{x}_i$'s.

In the context of time-dependent outputs, the set of parameters $\mathbf{x}_i$ also contains the time points $t \in \{t_1, \ldots, t_{n_i}\}$, that can be conveniently separated into n_i subsets $\mathbf{x}_{it}$, i.e. one for each time point. The function of the simulated data, e.g. J summary statistics, can be denoted as Y_{ijt}, and in full generality, an SE postulates a relationship of the form

$$Y_{ijt} = f_j(\mathbf{x}_{it}, \boldsymbol{\theta}_j) + \varepsilon_{ijt}, \tag{1}$$

where $\boldsymbol{\theta}_j$, for $j = 1, \ldots, J$, are the SE's parameters that are allowed to differ according to the different summary statistics. The emulators f_j can be the same functions for all

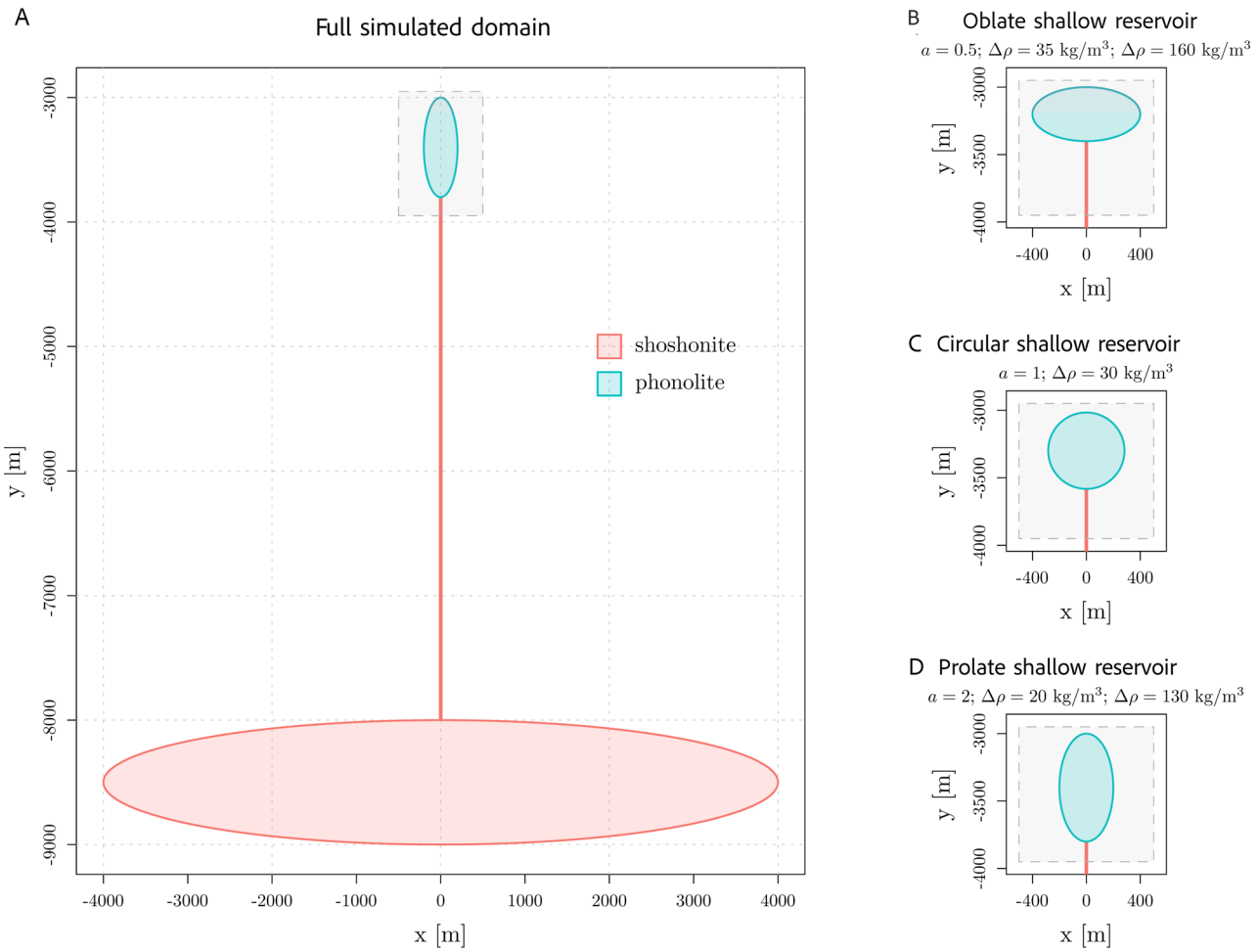

Figure 1: Summary of the five simulated settings. The upper chamber is either elliptical, with semi-axes of 200 and 400 m, or circular, with radius of 283 m to keep the same surface area. The dyke, which at time 0 hosts shoshonite like the deep chamber, is 20 m wide. The deep reservoir is elliptical, with semi-axes of 4.0 and 0.5 km. [A] full simulated domain, with prolate shallow reservoir; grey rectangle indicates the region magnified in [B], [C] and [D]. [B] Simulations 1 and 4 with oblate shallow reservoir. [C] Simulation 3 with circular shallow reservoir. [D] Simulations 2 and 5 with prolate shallow reservoir.

j, but they differ in the values of the parameters $\boldsymbol{\theta}_j$. The stochastic term ε_{ijt} allows to model what is not captured by the emulator and is defined after Equation 2. Thus, in vector form, we can consider $\mathbf{Y}_{ij} = \boldsymbol{f}_j(\mathbf{x}_i, \boldsymbol{\theta}_j) + \boldsymbol{\varepsilon}_{ij} = \left[f_j(\mathbf{x}_{it_1}, \boldsymbol{\theta}_j), \ldots, f_j(\mathbf{x}_{it_{n_i}}, \boldsymbol{\theta}_j)\right]^T + \boldsymbol{\varepsilon}_{ij}$, or, in a more compact form:

$$\begin{aligned}\mathbf{Y}_i &= \boldsymbol{f}(\mathbf{x}_i, \boldsymbol{\theta}) + \boldsymbol{\varepsilon}_i \\ &= \Big[f_1(\mathbf{x}_{it_1}, \boldsymbol{\theta}_1), \ldots, f_1(\mathbf{x}_{it_{n_i}}, \boldsymbol{\theta}_1), \ldots, \\ &\qquad \ldots, f_J(\mathbf{x}_{it_1}, \boldsymbol{\theta}_J), \ldots, f_J(\mathbf{x}_{it_{n_i}}, \boldsymbol{\theta}_J)\Big]^T + \boldsymbol{\varepsilon}_i. \end{aligned} \quad (2)$$

One also supposes, very generally, that $\boldsymbol{\varepsilon}_i = [\varepsilon_{ijt}]_{j=1,\ldots,J t=t_1,\ldots,t_{n_i}}$ is a stochastic process such that $\boldsymbol{\varepsilon}_i \sim \mathcal{F}\{\mathbf{0}, \boldsymbol{\Sigma}\}$, where $\mathcal{F}$ denotes some probability distribution in $\mathbb{R}^{J \times n_i}$ with mean $\mathbf{0}$ and covariance $\boldsymbol{\Sigma}$, where $\Sigma_{kl} = \sigma_k^2$ for $k = l$ and its off-diagonal terms can be non-zero. The vector of parameters $\boldsymbol{\theta} = [\boldsymbol{\theta}_1^T \ \ldots \boldsymbol{\theta}_J^T]^T \in \boldsymbol{\Theta} \subseteq \mathbb{R}^p$ is unknown, and needs to be estimated given a series of simulations $\mathbf{D} = (\mathbf{D}_1, \ldots, \mathbf{D}_m)$ from which J summary statistics are computed at n_i time points to obtain the $\mathbf{Y}_i$'s. The noise $\boldsymbol{\varepsilon}_i$ captures the approximation error of the link functions f_j in Equation 2, which can be made as complex as possible, but will never match exactly the true underlying theoretical process. The noise is random, and we suppose that its (probability) distribution is fully specified by its first two moments—i.e. for the i-th simulation, $\mathbb{E}[\boldsymbol{\varepsilon}_i] = \mathbf{0}$ and $\mathrm{cov}(\boldsymbol{\varepsilon}_i) = \boldsymbol{\Sigma}$. Importantly, the components of $\boldsymbol{\Sigma}$ can be estimated to obtain an estimated covariance matrix $\hat{\boldsymbol{\Sigma}}$, which can be used to compute confidence intervals for the predictions provided by the SE.

3.1 Application to magma dynamics

We consider here an SE for the magma dynamics model presented in Section 2. The salient features of the SE, as well as the construction of its inputs, are highlighted in the flow chart of Figure 3. The input parameters $\mathbf{x}_{it}$, for the $m = 5$ simulations described in Table 1, are the aspect ratio of the

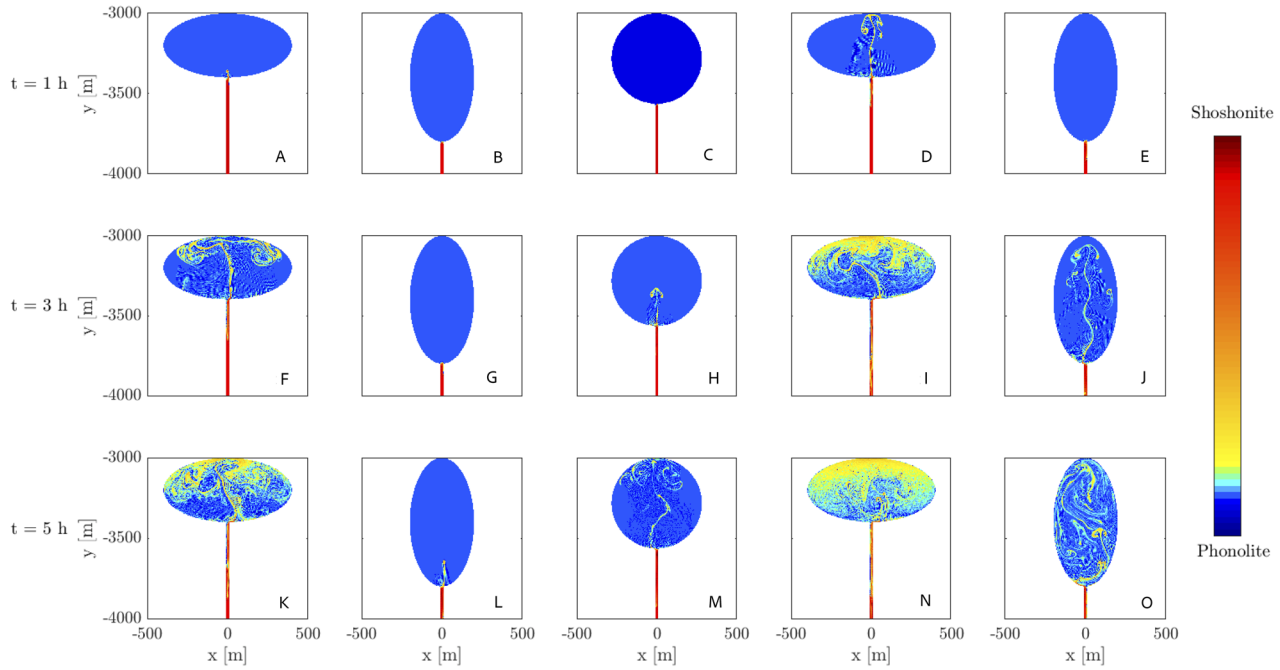

Figure 2: Space-time evolution of composition in the upper region of the simulated domain, for the different scenarios. Rows represent different times, columns represent different simulations.

shallow reservoir a_i, the initial density difference between the two end-member magmas $\Delta\rho_i$, and the interaction time $t \in \{t_1, \ldots, t_{n_i}\}$, so that $\mathbf{x}_{it} = [a_i, \Delta\rho_i, t]^T$. The output from the simulator is $\mathbf{D}_i$, an $r_i \times n_i$ matrix containing the silicate melt compositions expressed as the proportions of the two end-member magmas, which are measured across r_i locations of the shallow reservoir at n_i interaction times. From this simulator output, summary statistics are computed to reduce the output dimension, and we have chosen to consider $J = 5$ quantiles (namely, 0.05, 0.25, 0.50, 0.75, and 0.95) computed from the distribution of the compositions over the r_i locations, with $J \ll r_i$. Namely, Y_{ijt} in Equation 1 corresponds to the j-th row (i.e. j-th quantile) and t-th column of $\mathbf{Y}_i$ in Figure 3. A statistical model as in Equation 1 is then fitted to the summary statistics to obtain an estimated value $\hat{\boldsymbol{\theta}}$. The statistical model we consider is Equation 1 with $f_j(\mathbf{x}_{it}, \boldsymbol{\theta}_j) = f_j(t, \Delta\rho_i, a_i, \boldsymbol{\theta}_j)$ and

$$\begin{aligned} f_j(t, \Delta\rho_i, a_i, \boldsymbol{\theta}_j) = {} & \beta_{1j} + \beta_{2j} t + \beta_{3j} \Delta\rho_i + \beta_{4j} a_i \\ & + \beta_{5j} t \Delta\rho_i + \beta_{6j} t a_i + \beta_{7j} \Delta\rho_i a_i \\ & + \beta_{8j} t \Delta\rho_i a_i + \beta_{9j} t^2 a_i + \beta_{10j} t^3 a_i \\ & + \beta_{11j} t^4 a_i + \beta_{12j} t^2 \Delta\rho_i a_i \\ & + \beta_{13j} t^3 \Delta\rho_i a_i + \beta_{14} t^4 \Delta\rho_i a_i. \end{aligned} \quad (3)$$

Therefore, the (unknown) parameter vector $\boldsymbol{\theta} = [\boldsymbol{\theta}_1^T, \ldots, \boldsymbol{\theta}_J^T]^T$, with $\boldsymbol{\theta}_j = [\beta_{1j}, \ldots, \beta_{14j}, \sigma_j^2]^T$, has dimension $p = 15 \cdot J = 75$. To estimate the parameters in Equation 3, namely $\hat{\boldsymbol{\theta}}_j = [\hat{\beta}_{1j}, \ldots, \hat{\beta}_{14j}, \hat{\sigma}_j^2]^T$, we consider the Ordinary Least Squares (OLS) estimator for a standard linear regression model. More precisely, the data were split into J subsets and the corresponding Y_{ijt} were regressed on $t, \Delta\rho_i, a_i$ and combinations of them according to Equation 3, for each j separately. Importantly, in order to reduce the computational burden, we considered a uniform random subsample encompassing 30 % of the simulated interaction times t. Indeed, the total number of time points was, on average across the $m = 5$ simulations, approximately $n_i = 2200$, so that 30 % of the simulated interaction times represents an average of $n_i^* \approx 650$ points per simulation, which corresponds to a regression model with approximately $n_i^* \times m = 650 \times 5 = 3250$ responses. We remark that while the choice of summary statistics and statistical model is arbitrary, with the possibility of considering alternative options, the SE provides satisfactory results with the current formulation (see Section 3.2). Then, the (unknown) conditional expectation $\mathbb{E}\left[\mathbf{Y}_{ij} \middle| t, a_i, \Delta\rho_i, \boldsymbol{\theta}_j\right] = f_j(t, \Delta\rho_i, a_i, \boldsymbol{\theta}_j)$ can be estimated by $f_j(t, \Delta\rho_i, a_i, \hat{\boldsymbol{\theta}}_j)$. The latter can be used to predict t, or indeed also a or $\Delta\rho$, given a set of outcomes $\mathbf{y} = [y_1, \ldots, y_J]^T$, which are usually obtained by field measurements. In our case the measurements are the whole rock chemistry of erupted volcanic rocks. Thus, given an observed $\hat{\mathbf{y}}$ (field measurements of whole rock chemistry of erupted rocks), and as precise as possible information about a (from geophysics) and $\Delta\rho$ (which can be calculated from the chemistry of the two end-member magmas and existing models [see e.g. Lange 1997]), a statistical emulation prediction for the interaction time t can be obtained through

$$\hat{t} = \arg\min_t \left\| \hat{\mathbf{y}} - \boldsymbol{f}(t, a, \Delta\rho, \hat{\boldsymbol{\theta}}) \right\|^2_{\hat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\theta}})}, \quad (4)$$

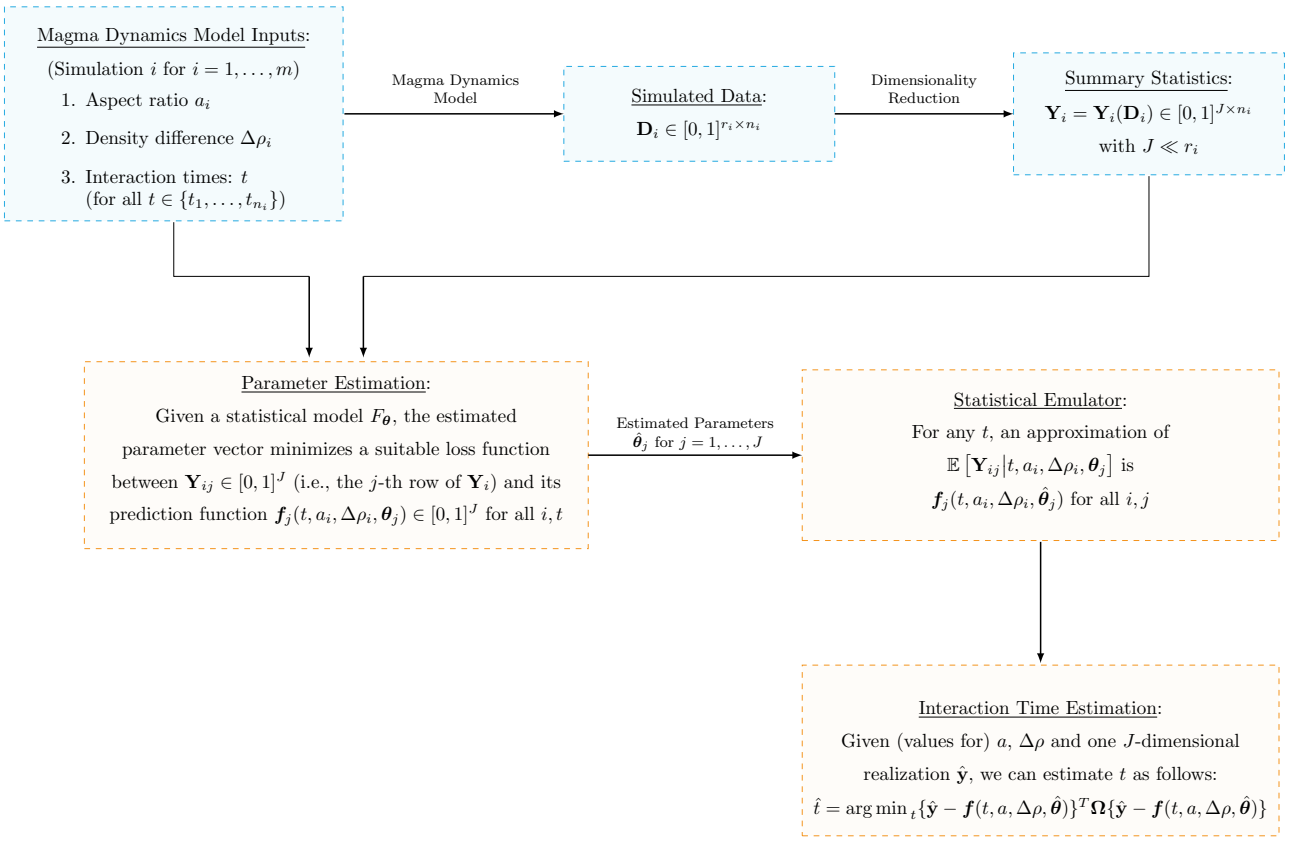

Figure 3: Flow chart from the simulator output to the statistical simulator and the prediction computation for the magma dynamics model in Section 2.

where $\|\mathbf{x}\|_A^2 = \mathbf{x}^T \boldsymbol{A} \mathbf{x}$, and $\hat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\theta}})$ is a suitably chosen positive definite weighting matrix. In particular, $\hat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\theta}})$ should be as close as possible to the inverse covariance matrix of the model's errors. In our case study we simply choose $\hat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\theta}}) = \text{diag}(1/\hat{\sigma}_{p_1}^2, \ldots, 1/\hat{\sigma}_{p_J}^2)$, where $\text{diag}(\cdot)$ denotes a diagonal matrix and $\hat{\sigma}_{p_j}$ (for $j = 1, \ldots, J$) is the standard error of the predicted mean for the j-th quantile. The prediction $\hat{t}$ defined in Equation 4 is a classical minimum distance estimator with well-known statistical properties [see e.g. Newey and McFadden 1994]. Under plausible conditions, this estimator is consistent for t and, under stronger conditions, asymptotically normally distributed, which implies that we can use this property to associate prediction errors when assessing error propagation (see Section 4).

3.2 Prediction Accuracy

To evaluate the prediction accuracy of the SE, we consider $\hat{\mathbf{y}}$ which is obtained by running the simulator at known a, $\Delta\rho$, and t. Then, we can compare the predicted interaction time $\hat{t}$ based on Equation 4 to the known interaction time t, say the *true* interaction time, hence providing a way to evaluate the (in sample) prediction accuracy.

Figure 4A compares predicted ($\hat{t}$) versus the true (t) interaction time (both measured in hours) for Simulations 2–4 of Table 1; Simulations 1 and 5 exhibit comparable results and are consequently excluded. The resulting predictions across different simulation settings are very close to the true ones, which are represented by a dashed black line. This is further highlighted in Figure 4B which reports the associated absolute prediction errors, i.e. $|\hat{t} - t|$. Overall, absolute errors are lower than 0.8 hours across all settings and typically lower than 0.2 hours. However, a noticeable increase in prediction errors is present only for Simulation 2 as interaction time approaches 5 hours. This might be due to the fact that the setting of Simulation 2 produces outputs that take much longer than the others to homogenise, as both lower density contrasts and prolate reservoir shape hinder mingling efficiency [Montagna et al. 2015], so that, at larger times, the output is notably different to the output in the other settings. This feature should be introduced in the SE, in order to produce better predictions, but this is left for further research. This prediction accuracy exercise highlights the subtle adjustments that are needed when comparing predictions to actual data. Moreover, it should be emphasised that larger prediction errors could be partially explained also by poorer performance of the SE itself in some parts of the parameter space (see Section 5.2 for further details).

Since the SE is built up using all information in the shallow reservoir, we also performed an emulation study to investigate two aspects of the SE; here we keep the SE parameters as estimated in Section 3.1. First, we note that the whole simulated data for the estimation of the parameters of the SE (i.e. $\hat{\boldsymbol{\theta}}_j$

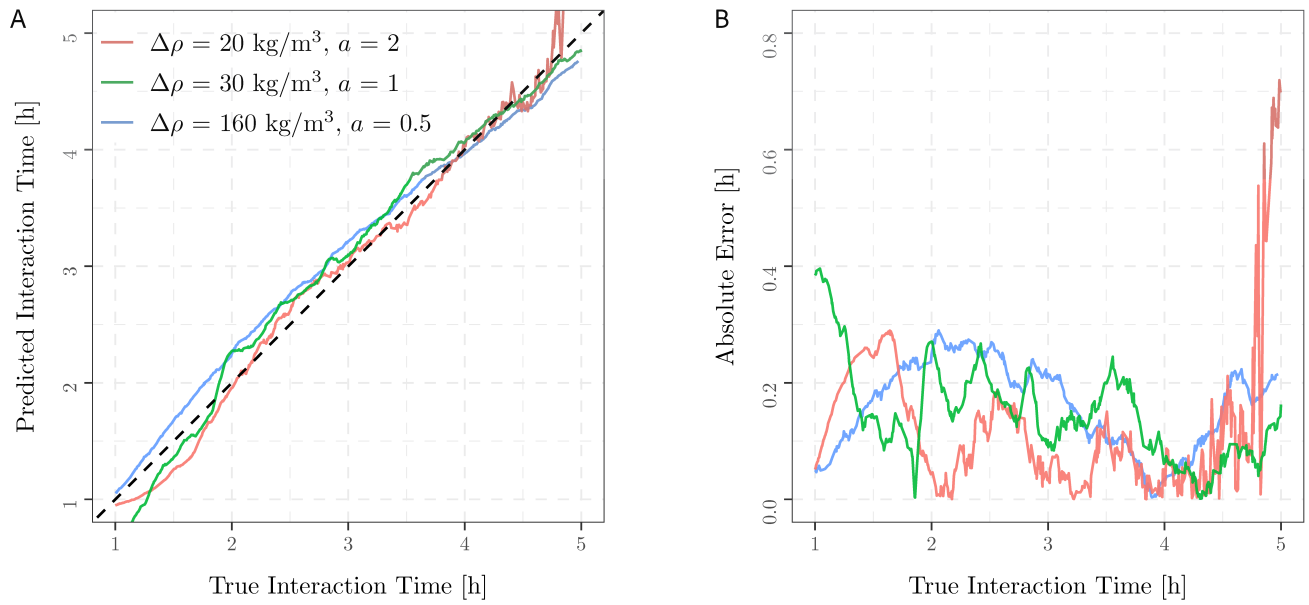

Figure 4: Panel [A] predicted (solid lines) versus true (dashed line) interaction times (in hours) across Simulations 2–4 of Table 1, computed using Equation 4. Panel [B] absolute prediction errors associated with each simulation setting.

for $j = 1, \ldots, J$) might not, *a priori*, be necessary. Therefore, we also considered uniformly subsampling the simulator's outputs for given aspect ratios, density differences, and time, i.e. the rows of $\mathbf{D}_i$ in Figure 3. The subsamples represent respectively 60, 30, and 10 % of the simulated outcomes, on which the quantiles of interest are then computed. Additionally, as an eruption could extract magma from progressively deeper portions of the magmatic reservoir, we assessed the accuracy of the SE in predicting the times considering the lower and upper half of the magma reservoir, separately. The predicted versus true interaction times (in hours) across Simulations 2–4, computed using Equation 4, are presented in Figure 5, for the different emulation settings. Panels 5A–5C assess the accuracy of the SE on the whole shallow reservoir, while panels 5D–5F and panels 5G–5I highlight its accuracy in respectively, the upper half and lower half parts of the shallow reservoir. Within each panel, the accuracy is assessed using different sampling proportions. In particular, Figure 5A–C shows that, across the considered simulation settings, subsampling from the whole shallow reservoir does not significantly impact predictive accuracy, with the exception of Simulation 2 (Figure 5A), which is characterised by a slightly longer homogenisation time, resulting in a more spread distribution. However, the remaining panels of Figure 5 highlight the presence of estimation biases even under a full sampling scheme. Such biases are more marked for Simulation 4, reported in Figure 5C, 5F, and 5H, and they can be partially explained by the nature of this simulation. These biases can be generally associated with chamber stratification. Indeed, while the whole system is more or less going towards homogeneity, it does so in a way that the upper half shallow chamber contains more of one of the two end-member magmas, and the lower half shallow chamber contains more of the other. Having said that, it should be stressed that these biases are not due to an inaccurate adjustment of the SE, but instead constitute additional information for field data collection. For instance, if the available deposits from eruptions show a transition of chemistry from the bottom to the top of the deposit, then one would adapt the SE to separate the simulator outputs from the upper and lower half, so that better predictions can be built. In other words, when chemical gradients in erupted deposits can provide additional information on the modality of magma extraction from the subvolcanic reservoir, the SE can be adapted accordingly. Indeed, due to the non-homogeneous nature of reservoir processes, different regions of the system may record different homogenisation timescales, as opposed to one common history for all.

4 Error propagation

The (fixed) parameters a, $\Delta\rho$ and $\hat{\mathbf{y}}$ in Equation 4 are subject to measurement error, $\boldsymbol{f}(\cdot)$ is subject to approximation error, and the (fixed) parameters $\hat{\boldsymbol{\theta}}$ are subject to estimation error. The last two types of errors can be deduced from the properties of the estimators using model Equation 2. We however found out that the approximation error—due to the use of a statistical model $\boldsymbol{f}(\cdot)$ to approximate the true physical mechanism—is small for the considered application (see Figure 4), and the estimation error for $\hat{\boldsymbol{\theta}}$ is negligible.

Measurement errors for a, $\Delta\rho$, and the elements of $\hat{\mathbf{y}}$ can be inferred from scientists' experience. All sources of error can then be included in the statistical emulation procedure to obtain a set of possible values for $\hat{t}$ in Equation 4, from which confidence intervals can be built. This can be achieved through simulations. Indeed, one can consider that the errors are symmetrically distributed around the actual value of the parameters, so that, for each parameter, we can simulate an error drawn from a pre-specified distribution, such as the normal distribution. This is equivalent to simulating possible values for the different parameters, from a normal distribution,

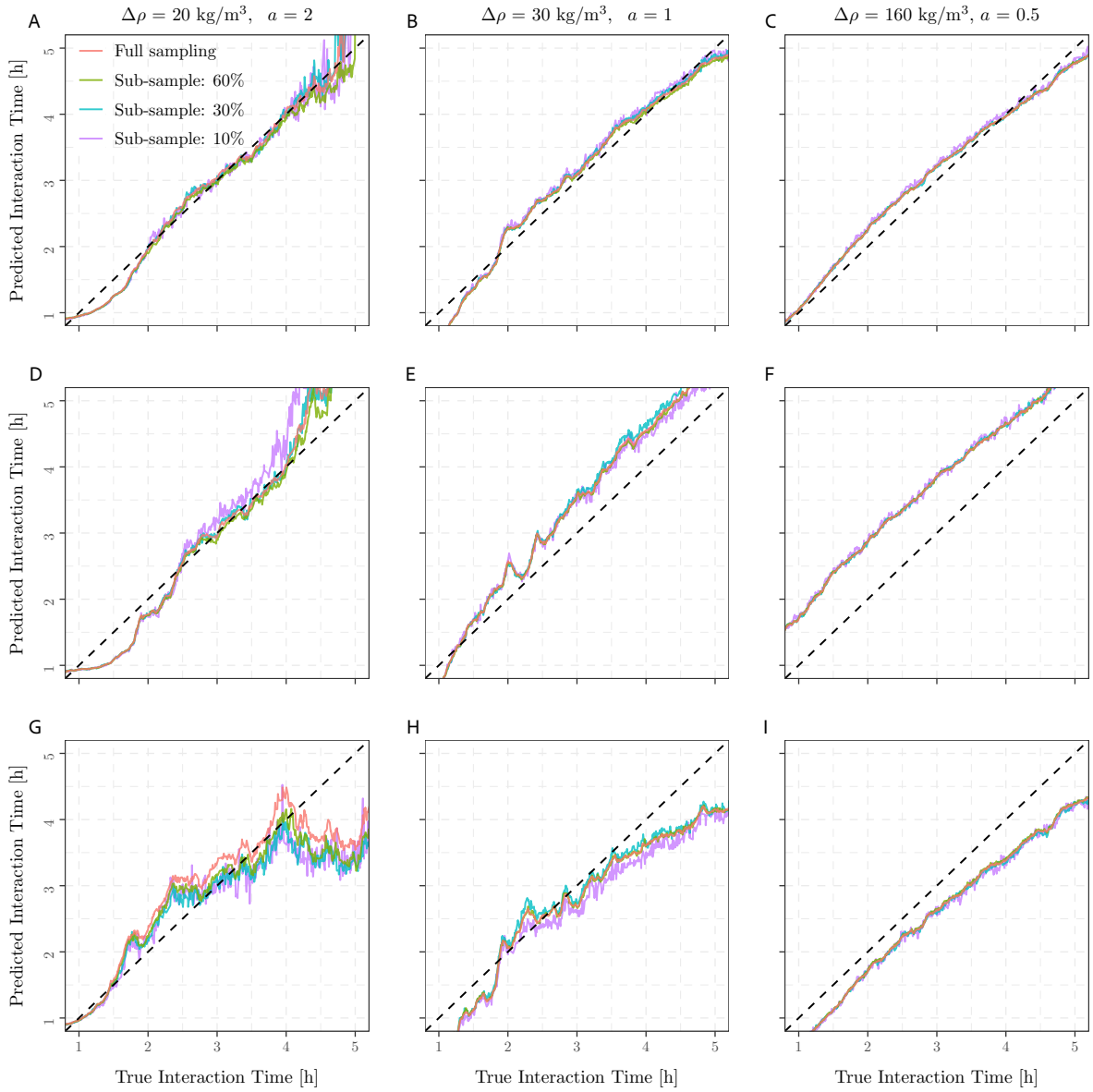

Figure 5: Predicted (solid lines) versus true (dashed lines) interaction times (in hours) across Simulations 2–4 of Table 1, computed using Equation 4. The predicted interaction times are computed on the whole shallow reservoir (panels [A]–[C]), the upper half part (panels [D]–[F]) and the lower half part (panels [G]–[I]) of the shallow reservoir. Within each panel, the predicted interaction times are computed on different (uniform) subsamples.

centred at the true value, with a pre-specified variance. For the latter, based on domain knowledge, a standard error of 10 % of the actual value is usually expected. Let the elements $\hat{y}^*_{hj}$ of $\hat{\mathfrak{y}}^*_h$, a^*_h, and $\Delta\rho^*_h$, for $h = 1, \ldots, H$, be (independent) random realisations drawn from normal distributions, then

$$\hat{t}^*_h = \arg\min_t \left\| \hat{\mathfrak{y}}^*_h - \hat{\boldsymbol{f}}(t, a^*_h, \Delta\rho^*_h, \hat{\boldsymbol{\theta}}) \right\|^2_{\hat{\boldsymbol{\Omega}}(\hat{\boldsymbol{\theta}})} \tag{5}$$

represents a suitable distribution for $\hat{t}$ that takes into account all (significant) sources of errors. Thus, a confidence interval at the $1-\alpha$ confidence level for t is given by $(\hat{t}^*_{(l)}, \hat{t}^*_{(u)})$, where $l = \lfloor(\alpha/2)H\rfloor$ and $u = \lceil(1-\alpha/2)H\rceil$, with $x_{(i)}$ being the i-th order statistics, and with $\lfloor x\rfloor$, respectively $\lceil x\rceil$, being the integer smaller than or equal, respectively larger than or equal, to x.

Figure 6 presents 95 % confidence intervals for predicted interaction times (in hours), across Simulations 2–4 (from left to right), when some parameters are fixed and others are subject to random error. Specifically, in Figure 6A–6C, both a and $\Delta\rho$ are subject to random error while the elements of $\hat{\mathfrak{y}}$ are fixed, in Figure 6D–6F, a and $\Delta\rho$ are fixed while the elements of $\hat{\mathfrak{y}}$ are subject to random error, and in Figure 6G–6I all parameters are subject to random error. Overall, predicted interaction times are close to the true ones across all simulation settings, and we note that the estimation variability increases as we assess for all input parameters (Figure 6G–6I). Confidence intervals are generally smaller than 20 %, reach-

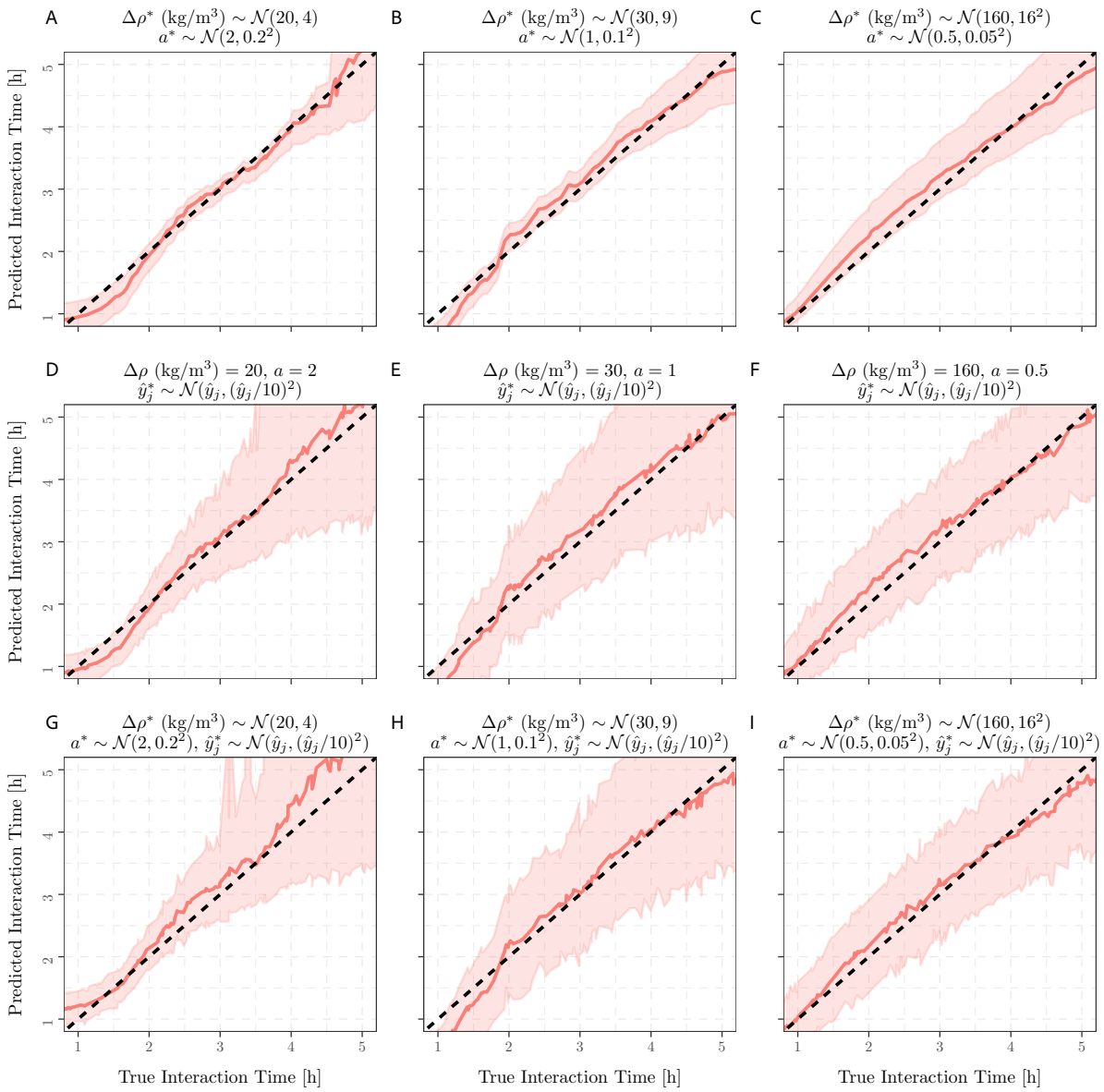

Figure 6: Sensitivity of the SE across Simulations 2–4 of Table 1, where input parameters $\Delta\rho^*$ and a^* (panels [A]–[C]), $\hat{y}_j^*$ (panels [D]–[F]), $\Delta\rho^*$, a^*, and $\hat{y}_j^*$ (panels [G]–[I]) are drawn from normal distributions centred at their true values ($\Delta\rho, a, \hat{y}_j$) with standard deviations equal to 10 % of such true values. Average predicted (solid lines) versus true (dashed lines) interaction times (in hours), as well as 95 % confidence intervals, are computed using H = 200 in Equation 5. Each setting is restricted to a (uniform) subsample encompassing 25 % of the locations sampled from the shallow reservoir.

ing up to 30 %. Given the large uncertainties that characterise our understanding of inaccessible magmatic plumbing systems [Sigurdsson et al. 2015], these confidence intervals are rather narrow and allow us to provide meaningful estimates of pre-eruptive magmatic interaction times.

5 Discussion

The results of our analysis show that the SE is capable of reproducing the results of the numerical model with small uncertainty when both the contrast in density and the shape of the reservoir in which the interaction between magma occurs are known (Figure 4). This is also true when we consider subsamples of the simulated domain. Interestingly, for some of the simulations, differences emerge when considering the upper and lower half of the simulated domain (Figure 4E, 4F, 4H, 4I). This indicates that for the specific input parameters of these runs (Table 1), the interaction will generate a zoned magma reservoir with chemical differences between its lower and upper portions. This is of interest, as it shows the range of model input parameters under which the interaction between two magmas leads to chemical heterogeneities that can be identified in the deposits of the eruption (e.g. chemical difference along the stratigraphy of deposit).

The contrast in density between two magmas can be calculated from existing models providing magma density as a function of its chemistry and the pressure and temperature

conditions at which the interaction was occurring [e.g. Lange 1997]). The temperature and pressure at which the interaction occurs can be estimated using thermobarometry [Nimis and Ulmer 1998; Neave and Putirka 2017; Petrelli et al. 2020; Jorgenson et al. 2022]. Clearly, temperature and pressure estimates, together with estimates of the volatile content of magmas, which impact their density, are all associated with uncertainties. Additionally, the geometry of the reservoir could be defined using geophysics, but this would not be possible when studying past eruptions. Thus, to quantify the duration of the pre-eruptive interaction between magmas, and fundamentally, the associated uncertainty, it is essential to propagate all uncertainties associated with the input parameters of the model. This is when the advantages of statistical emulation become evident as the alternative would be to perform a large number of time-consuming numerical simulations varying the input parameters within a range considered reasonable by the experts. The results of runs performed using the SE and considering a normal distribution with 10 % uncertainty for input parameters such as a and $\Delta\rho$ and the measurable parameters in $\hat{\mathbf{y}}$ (i.e. the distribution of chemistry of erupted magma), show that the uncertainty in the estimated interaction time preceding an eruption increases with the duration of interaction, but does not exceed 25–30 % of the total interaction time (Figure 6).

As a byproduct, the SE provides insights on the sensitivity of model results to specific input parameters. Successive runs of the simulator must thus target specific regions of the parameter space that yield most representative results. Moreover, the SE results obtained with large undersampling of the simulated outputs are practically indistinguishable from those obtained with the whole datasets (Figure 4). Eliminating unnecessary outputs would largely decrease computational costs of large simulations, contributing to an overall leaner and more manageable workflow.

5.1 Application to petrological dataset

To test the validity of the SE, we have performed a proof-of-concept inversion on a petrological dataset from the Agnano-Monte Spina eruption at Campi Flegrei, on which the numerical model is loosely based [Arienzo et al. 2010]. The products of this eruption were selected because of existing evidence that magma mixing ultimately triggered this event and it is considered as the reference for future large-size explosive eruptions at Campi Flegrei [D'Antonio et al. 1999; Di Vito et al. 1999]. We used Al_2O_3 weight fraction in melt inclusions as a proxy for composition C. We express composition as weighted sum of phonolitic and shoshonitic (indexed by P and S, respectively) end-members: $C = \alpha C_P + (1 - \alpha C_S)$, where α is phonolite weight fraction. The compositional distribution obtained from the dataset is then fed into the SE to obtain the pre-eruptive interaction times, given a distribution of the two parameters a and $\Delta\rho$. The distributions for a and $\Delta\rho$ have been chosen, respectively, as skewed towards sill-like geometries and uniform. The obtained interaction times range between 5 and 8 hours (Figure 7), which is a timescale loosely compatible with what has been proposed for the recharge-to-eruption times at Campi Flegrei (hours to days) [Perugini et al. 2010; Arzilli et al. 2016; Astbury et al. 2018]. In the simulations we consider all analyses collected along the stratigraphy together, which return a unimodal distribution (Figure 7B). Some chemical zoning was present in the reservoir as testified by the increase of mafic component toward the top of the stratigraphic sequence [Arienzo et al. 2010]. However, analyses collected for the same eruptive unit would not be sufficient to produce a statistically representative distribution of whole rock compositions to compare with the SE. Because with the progression of mixing, the compositional distribution of magma in the reservoir evolves from bimodal to unimodal, our results tend to overestimate the timescale of mixing. These considerations provide a framework for the collection of samples on which to apply the approach we propose. More precise constraints of the timescales of mixing would require the collection of a statistically representative number of samples for each eruptive unit and possibly on the largest possible number of outcrops at different locations. Regarding the case at hand, the extremely short timescale of magma interaction that preceded the Agnano-Monte Spina eruption provides solid and important constraints that can help in assessing volcanic hazards.

5.2 Limitations and further development

Some limitations of our modelling approach and potential avenues for further research, as highlighted by one of the anonymous reviewers, deserve additional consideration. Firstly, Figure 4 highlights the potential presence of some systematic biases for the considered SE, particularly in the range between 2 and 3.2 hours, where predicted interaction times consistently overestimate the true values. It also suggests that the model may not be fully reliable at the boundary (e.g. at larger interaction times) where predicted interaction times can be sensitive to slight changes in the true time. These discrepancies may arise from several sources. In particular, the (function of the) simulated data lie in the interval $(0, 1)$; we use a regression model for data in $(-\infty, \infty)$, meaning that a beta regression would be more appropriate, for example. However, in most settings of practical relevance, we obtained predictions (as well as their associated prediction intervals) between 0 and 1 (results not shown here). Another source of inaccuracy could be attributed to unaccounted correlations in the simulated data that can result into unaccounted correlations in the model's residuals. In the settings of practical relevance, we found that the residuals from our linear model are fairly well-behaved, although they exhibit some "mild" form of long-range dependency (results not shown here). Additionally and in general, the SE should also be adapted to the potential presence of non-linear relationships and heteroskedasticity. Moreover, the fitting method should also be adapted in the cases where outlying observations might be present.

We acknowledge the risk of model misspecification in general, which for the cases studied in this paper, appear to be quite mild. Indeed, as illustrated in Figure 4B, the prediction error is quite small for practical purposes. Moreover, we remark that simpler models, such as the one in use, often have a tendency to generalise better than more complex methods. While our approach maintains consistency against misspecified models, future research may explore more suitable and

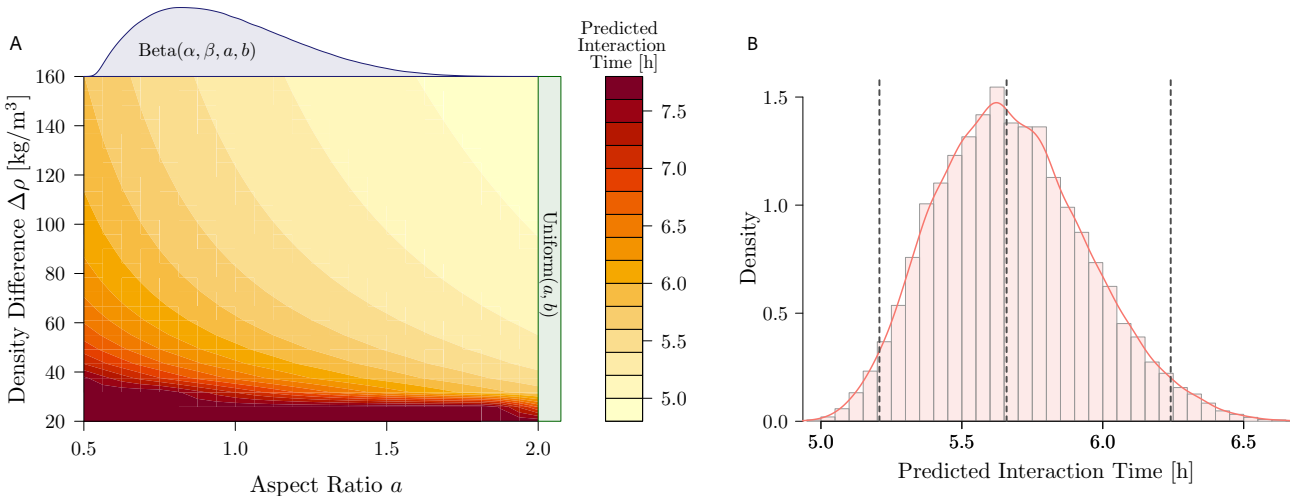

Figure 7: [A] magmatic interaction times as obtained from the SE inversion of a compositional dataset from products of the Agnano-Monte Spina eruption at Campi Flegrei across a grid of a and $\Delta\rho$ inputs. Interaction times recovered from the model match those provided in the literature for the specific eruption. [B] distribution of predicted interaction times (in hours), as well as 95 % confidence intervals and median (dashed vertical lines) based on H = 10,000 replications in Equation 5. The sensitivity of the SE was assessed at random input parameters $\Delta\rho^*$ and a^*, which are respectively drawn from a uniform and a beta distribution with shape parameters α = 2 and β = 5 (represented as light green and blue densities in panel [A], respectively) and $\hat{y}_j^*$ which is drawn from a normal distribution centred at its true values $\hat{y}_j$ with standard deviations equal to 10 % of such true values.

flexible approaches such as semi-parametric methods. In our modelling exercise, we mainly focused on point estimation and, as highlighted in Figure 6, uncertainties in $\hat{\boldsymbol{\theta}}$ are dominated by input-related factors. The overall performance of the considered SE is satisfactory for the considered case study, leading to good predictive power and generalisability.

We also remark that a very limited input grid was used in our experiments, and we are currently exploring the use of a finer grid to achieve better results. Overall, our study serves as a proof-of-concept, recognising limitations and paving the way for future enhancements in methodology and data collection.

6 Conclusions

The Earth Sciences often focus on phenomena that occurred in the past or at inaccessible depths, which inhibits our capacity to define with confidence the range of input parameters of numerical models targeting the quantification of geological processes. Additionally, because of the complexity of some physical models, the realisation of a large number of simulations—which would allow one to assess how uncertainty in the input parameter affects the final results—is prohibitive. Here, we use a specific application to volcanology that allows us to highlight the multiple advantages of combining physical modelling and statistical emulation. We show how to determine the duration of magma interaction preceding a volcanic eruption from the analysis of the distribution of the erupted rock chemistry. This particular case is appropriate for Campi Flegrei (Italy) [Bagagli et al. 2017; Morgavi et al. 2017; Montagna et al. 2022]. The approach we present here could be applied to past eruptions of Campi Flegrei to determine the duration of magma interaction preceding any volcanic eruption throughout its eruptive history. The workflow would require:

i. The analysis of the whole rock chemistry of a statistically representative set of samples of each eruption.

ii. The calibration of SE on physical models realised over a wide range of geometries, magma conditions and duration.

iii. The comparison between the measured and calculated distribution.

The results of our analysis show that statistical emulation provides an excellent means to quantify the uncertainty in duration of magma interaction preceding a volcanic eruption from the distribution of chemical composition of the erupted rocks.

Author Contributions

All authors contributed to the conceptualisation and methodology of the paper. CPM provided the raw simulated data from physical modelling. LI and SG developed the R code and performed the statistical analyses. All authors contributed to the writing and editing of the manuscript.

Acknowledgements

LI and SG acknowledge the financial support of the Swiss National Science Foundation (professorships grant n. 176843) as well as Swiss Innovation Agency Innosuisse (grants n. 37308.1 IP-ENG and n. 53622.1 IP-ENG). MPVF acknowledges the financial support of the Swiss National Science Foundation (grant n. 182684). LC acknowledges the financial support of the Swiss National Science Foundation (grant n. 200021_184632) and has received additional funding from the European Union's Horizon 2020 research and innovation actions under grant agreement No 731070.

Data Availability

The data analysed in this paper, as well as documented R code to process the data and replicate our results, are openly available at: https://github.com/LucaIns/magma. The code used for the magma dynamics simulations is available at: https://gitlab.com/dgmaths9/gales.

Copyright Notice

References

Annen, C., J. D. Blundy, and R. S. J. Sparks (2006). "The genesis of intermediate and silicic magmas in deep crustal hot zones". *Journal of Petrology* 47(3), pages 505–539. DOI: 10.1093/petrology/egi084.

Arienzo, I., R. Moretti, L. Civetta, G. Orsi, and P. Papale (2010). "The feeding system of Agnano-Monte Spina eruption (Campi Flegrei, Italy): Dragging the past into present activity and future scenarios". *Chemical Geology* 270(1-4), pages 135–147. DOI: 10.1016/j.chemgeo.2009.11.012.

Arienzo, I., L. Civetta, A. Heumann, G. Wörner, and G. Orsi (2009). "Isotopic evidence for open system processes within the Campanian Ignimbrite (Campi Flegrei–Italy) magma chamber". *Bulletin of Volcanology* 71(3), pages 285–300. DOI: 10.1007/s00445-008-0223-0.

Arzilli, F., M. Piochi, A. Mormone, C. Agostini, and M. R. Carroll (2016). "Constraining pre-eruptive magma conditions and unrest timescales during the Monte Nuovo eruption (1538 ad; Campi Flegrei, Southern Italy): integrating textural and CSD results from experimental and natural trachyphonolites". *Bulletin of Volcanology* 78(10), page 72. DOI: 10.1007/s00445-016-1062-z.

Astbury, R. L., M. Petrelli, T. Ubide, M. J. Stock, I. Arienzo, M. D'Antonio, and D. Perugini (2018). "Tracking plumbing system dynamics at the Campi Flegrei caldera, Italy: High-resolution trace element mapping of the Astroni crystal cargo". *Lithos* 318-319, pages 464–477. DOI: 10.1016/j.lithos.2018.08.033.

Bagagli, M., C. P. Montagna, P. Papale, and A. Longo (2017). "Signature of magmatic processes in strainmeter records at Campi Flegrei (Italy)". *Geophysical Research Letters* 44(2), pages 718–725. DOI: 10.1002/2016GL071875.

Blundy, J. and K. Cashman (2001). "Ascent-driven crystallisation of dacite magmas at Mount St Helens, 1980-1986". *Contribution to Mineralogy and Petrology* 140(6), pages 631–650. DOI: 10.1007/s004100000219.

Blundy, J. D. and R. S. J. Sparks (1992). "Petrogenesis of mafic inclusions in granitoids of the Adamello Massif, Italy". *Journal of Petrology* 33(5), pages 1039–1104. DOI: 10.1093/petrology/33.5.1039.

Bohrson, W. A., F. J. Spera, M. S. Ghiorso, G. A. Brown, J. B. Creamer, and A. Mayfield (2014). "Thermodynamic model for energy-constrained open-system evolution of crustal magma bodies undergoing simultaneous recharge, assimilation and crystallization: The magma chamber simulator". *Journal of Petrology* 55(9), pages 1685–1717. DOI: 10.1093/petrology/egu036.

Caricchi, L., T. E. Sheldrake, and J. D. Blundy (2018). "Modulation of magmatic processes by CO_2 flushing". *Earth and Planetary Science Letters* 491, pages 160–171. DOI: 10.1016/j.epsl.2018.03.042.

Caricchi, L., G. Simpson, and U. Schaltegger (2014). "Zircons reveal magma fluxes in the Earth's crust". *Nature* 511(7510), pages 457–461. DOI: 10.1038/nature13532.

Caricchi, L., M. Townsend, E. Rivalta, and A. Namiki (2021). "The build-up and triggers of volcanic eruptions". *Nature Reviews Earth & Environment* 2(7), pages 458–476. DOI: 10.1038/s43017-021-00174-8.

Cashman, K. V. and B. D. Marsh (1988). "Crystal size distribution (CSD) in rocks and the kinetics and dynamics of crystallization II: Makaopuhi lava lake". *Contribution to Mineralogy and Petrology* 99(3), pages 292–305. DOI: 10.1007/BF00375363.

Cassidy, M., J. M. Castro, C. Helo, V. R. Troll, F. M. Deegan, D. Muir, D. A. Neave, and S. P. Mueller (2016). "Volatile dilution during magma injections and implications for volcano explosivity". *Geology* 44(12), pages 1027–1030. DOI: 10.1130/G38411.1.

Chandrasekhar, S. (2013). *Hydrodynamic and Hydromagnetic Stability*. Courier Dover Publications. ISBN: 0486319202.

Chiodini, G., S. Caliro, P. De Martino, R. Avino, and F. Gherardi (2012). "Early signals of new volcanic unrest at Campi Flegrei caldera? Insights from geochemical data and physical simulations". *Geology* 40(10), pages 943–946. DOI: 10.1130/G33251.1.

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

Currin, C., T. Mitchell, M. Morris, and D. Ylvisaker (1991). "Bayesian prediction of deterministic functions, with applications to the design and analysis of computer experiments". *Journal of the American Statistical Association* 86(416), pages 953–963. DOI: 10.1080/01621459.1991.10475138.

D'Antonio, M., L. Civetta, G. Orsi, L. Pappalardo, M. Piochi, A. Carandente, S. De Vita, M. Di Vito, and R. Isaia (1999). "The present state of the magmatic system of the Campi Flegrei caldera based on a reconstruction of its behavior in the past 12 ka". *Journal of Volcanology and Geothermal Research* 91(2-4), pages 247–268. DOI: 10.1016/S0377-0273(99)00038-4.

Davidson, J., D. Morgan, B. Charlier, R. Harlou, and J. Hora (2007). "Microsampling and isotopic analysis of igneous rocks: Implications for the study of magmatic systems". *Annual Review of Earth and Planetary Sciences* 35(1), pages 273–311. DOI: 10.1146/annurev.earth.35.031306.140211.

Di Renzo, V., I. Arienzo, L. Civetta, M. D'Antonio, S. Tonarini, M. A. Di Vito, and G. Orsi (2011). "The magmatic feeding system of the Campi Flegrei caldera: Architecture and temporal evolution". *Chemical Geology* 281(3-4), pages 227–241. DOI: 10.1016/j.chemgeo.2010.12.010.

Di Vito, M., R. Isaia, G. Orsi, J. d. Southon, S. De Vita, M. D'Antonio, L. Pappalardo, and M. Piochi (1999). "Volcanism and deformation since 12,000 years at the Campi Flegrei caldera (Italy)". *Journal of Volcanology and Geothermal Research* 91(2-4), pages 221–246. DOI: 0.1016/S0377-0273(99)00037-2.

Druitt, T. H., F. Costa, E. Deloule, M. Dungan, and B. Scaillet (2012). "Decadal to monthly timescales of magma transfer and reservoir growth at a caldera volcano". *Nature* 482(7383), pages 77–80. DOI: 10.1038/nature10706.

Edmonds, M., A. Aiuppa, M. Humphreys, R. Moretti, G. Giudice, R. S. Martin, R. A. Herd, and T. Christopher (2010). "Excess volatiles supplied by mingling of mafic magma at an andesite arc volcano". *Geochemistry, Geophysics, Geosystems* 11(4). DOI: https://doi.org/10.1029/2009GC002781.

Edmonds, M., K. V. Cashman, M. Holness, and M. Jackson (2019). "Architecture and dynamics of magma reservoirs". *Philosophical Transactions of the Royal Society A* 377(2139), page 20180298. DOI: 10.1098/rsta.2018.0298.

Forni, F., W. Degruyter, O. Bachmann, G. D. Astis, and S. Mollo (2018). "Long-term magmatic evolution reveals the beginning of a new caldera cycle at Campi Flegrei". *Science Advances* 4(11), eaat9401. DOI: 10.1126/sciadv.aat9401.

Fourmentraux, C., N. Métrich, A. Bertagnini, and M. Rosi (2012). "Crystal fractionation, magma step ascent, and syn-eruptive mingling: The Averno 2 eruption (Phlegraean Fields, Italy)". *Contributions to Mineralogy and Petrology* 163(6), pages 1121–1137. DOI: 10.1007/s00410-012-0720-1.

Garg, D., A. Longo, and P. Papale (2018a). "Computation of compressible and incompressible flows with a space–time stabilized finite element method". *Computers & Mathematics with Applications* 75(12), pages 4272–4285. DOI: 10.1016/j.camwa.2018.03.028.

– (2018b). "Modeling free surface flows using stabilized finite element method". *Mathematical Problems in Engineering* 2018, pages 1–9. DOI: 10.1155/2018/6154251.

Garg, D. and P. Papale (2022). "High-performance computing of 3D magma dynamics, and comparison with 2D simulation results". *Frontiers in Earth Science* 9. DOI: 10.3389/feart.2021.760773.

Garg, D., P. Papale, S. Colucci, and A. Longo (2019). "Long-lived compositional heterogeneities in magma chambers, and implications for volcanic hazard". *Scientific Reports* 9(1) (3321). DOI: 10.1038/s41598-019-40160-1.

Giordano, D., J. Russell, and D. Dingwell (2008). "Viscosity of magmatic liquids: A model". *Earth and Planetary Science Letters* 271(1), pages 123–134. DOI: 10.1016/j.epsl.2008.03.038.

Glazner, A. F., J. M. Bartley, D. S. Coleman, W. Gray, and R. Z. Taylor (2004). "Are plutons assembled over millions of years by amalgamation from small magma chambers?" *GSA Today* 14(4-5), pages 4–11. DOI: 10.17615/nspy-zk53.

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

Guillas, S., A. Sarri, S. J. Day, X. Liu, and F. Dias (2018). "Functional emulation of high resolution tsunami modelling over Cascadia". *The Annals of Applied Statistics* 12(4), pages 2023–2053. DOI: 10.1214/18-AOAS1142.

Heap, M. J., M. Villeneuve, F. Albino, J. I. Farquharson, E. Brothelande, F. Amelung, J.-L. Got, and P. Baud (2020). "Towards more realistic values of elastic moduli for volcano modelling". *Journal of Volcanology and Geothermal Research* 390, page 106684. DOI: 10.1016/j.jvolgeores.2019.106684.

Hess, K.-U. and D. B. Dingwell (1996). "Viscosities of hydrous leucogranitic melts: A non-Arrhenian model". *American Mineralogist* 81, pages 1297–1300. DOI: 10.2138/am-1996-9-1031.

Higdon, D., J. Gattiker, B. Williams, and M. Rightley (2008). "Computer model calibration using high-dimensional output". *Journal of the American Statistical Association* 103(482), pages 570–583. DOI: 10.1198/016214507000000888.

Higgins, M. D. (2000). "Measurement of crystal size distributions". *American Mineralogist* 85(9), pages 1105–1116. DOI: 10.2138/am-2000-8-901.

Humphreys, M. C. S., M. Edmonds, M. Plail, J. Barclay, D. Parkes, and T. Christopher (2012). "A new method to quantify the real supply of mafic components to a hybrid andesite". *Contribution to Mineralogy and Petrology* 165(1), pages 191–215. DOI: 10.1007/s00410-012-0805-x.

Iovine, R. S., L. Fedele, F. C. Mazzeo, I. Arienzo, A. Cavallo, G. Wörner, G. Orsi, L. Civetta, and M. D'Antonio (2017). "Timescales of magmatic processes prior to the ~4.7 ka Agnano-Monte Spina eruption (Campi Flegrei caldera, Southern Italy) based on diffusion chronometry from sanidine phenocrysts". *Bulletin of Volcanology* 79(2). DOI: 10.1007/s00445-017-1101-4.

Ishii, M. and N. Zuber (1979). "Drag coefficient and relative velocity in bubbly, droplet or particulate flows". *AIChE Journal* 25(5), pages 843–855. DOI: https://doi.org/10.1002/aic.690250513.

Jackson, M. D., J. Blundy, and R. S. J. Sparks (2018). "Chemical differentiation, cold storage and remobilization of magma in the Earth's crust". *Nature* 564, pages 405–409. DOI: 10.1038/s41586-018-0746-2.

Jorgenson, C., O. Higgins, M. Petrelli, F. Bégué, and L. Caricchi (2022). "A machine learning-based approach to clinopyroxene thermobarometry: Model optimization and distribution for use in Earth sciences". *Journal of Geophysical Research: Solid Earth* 127(4), e2021JB022904. DOI: https://doi.org/10.1029/2021JB022904.

Karakas, O., W. Degruyter, O. Bachmann, and J. Dufek (2017). "Lifetime and size of shallow magma bodies controlled by crustal-scale magmatism". *Nature Geoscience* 10, pages 446–450. DOI: 10.1038/ngeo2959.

Keller, T. and J. Suckale (2019). "A continuum model of multi-phase reactive transport in igneous systems". *Geophysical Journal International* 219(1), pages 185–222. DOI: 10.1093/gji/ggz287.

Kennedy, M. C. and A. O'Hagan (2001). "Bayesian calibration of computer models". *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 63(3), pages 425–464. DOI: https://doi.org/10.1111/1467-9868.00294.

La Spina, G., F. Arzilli, M. R. Burton, M. Polacci, and A. B. Clarke (2022). "Role of volatiles in highly explosive basaltic eruptions". *Communications Earth & Environment* 3(1), page 156. DOI: 10.1038/s43247-022-00479-6.

Lange, R. A. (1997). "A revised model for the density and thermal expansivity of K_2O-Na_2O-CaO-MgO-Al_2O_3-SiO_2 liquids from 700 to 1900 K: extension to crustal magmatic temperatures". *Contribution to Mineralogy and Petrology* 130(1), pages 1–11. DOI: 10.1007/s004100050345.

Lavallée, Y., K.-U. Hess, B. Cordonnier, and D. Bruce Dingwell (2007). "Non-Newtonian rheological law for highly crystalline dome lavas". *Geology* 35(9), pages 843–846. DOI: 10.1130/G23594A.1.

Lejeune, A. M. and P. Richet (1995). "Rheology of crystal-bearing silicate melts: An experimental study at high viscosities". *Journal of Geophysical Research-Solid Earth* 100(B3), pages 4215–4229. DOI: 10.1130/G23594A.1.

Longo, A., M. Barsanti, A. Cassioli, and P. Papale (2012a). "A finite element Galerkin/least-squares method for computation of multicomponent compressible–incompressible flows". *Computers & Fluids* 67, pages 57–71. DOI: 10.1016/j.compfluid.2012.07.008.

Longo, A., D. Garg, P. Papale, and C. P. Montagna (2023). "Dynamics of magma chamber replenishment under buoyancy and pressure forces". *Journal of Geophysical Research: Solid Earth* 128(1), e2022JB025316. DOI: 10.1029/2022JB025316.

Longo, A., P. Papale, M. Vassalli, G. Saccorotti, C. P. Montagna, A. Cassioli, S. Giudice, and E. Boschi (2012b). "Magma convection and mixing dynamics as a source of Ultra-Long-Period oscillations". *Bulletin of Volcanology* 74(4), pages 873–880. DOI: 10.1007/s00445-011-0570-0.

Maclennan, J. (2019). "Mafic tiers and transient mushes: evidence from Iceland". *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 377(2139), page 20180021. DOI: 10.1098/rsta.2018.0021.

Mahmood, A., R. L. Wolpert, and E. B. Pitman (2015). "A physics-based emulator for the simulation of geophysical mass flows". *Uncertainty Quantification* 3(1), pages 562–585. DOI: 10.1137/130909445.

Mangiacapra, A., R. Moretti, M. Rutherford, L. Civetta, G. Orsi, and P. Papale (2008). "The deep magmatic system of the Campi Flegrei caldera (Italy)". *Geophysical Research Letters* 35(21), page L21304. DOI: 10.1029/2008GL035550.

Marsh, B. D. (1981). "On the crystallinity, probability of occurrence, and rheology of lava and magma". *Contribution to Mineralogy and Petrology* 78(1), pages 85–98. DOI: 10.1007/BF00371146.

Matoza, R. S. and D. C. Roman (2022). "One hundred years of advances in volcano seismology and acoustics". *Bulletin of Volcanology* 84(9), page 86. DOI: 10.1007/s00445-022-01586-0.

Mogi, K. (1958). "Relations between the Eruptions of Various Volcanoes and the Deformations of the Ground Surfaces around them". *Bulletin of the Earthquake Research Institute* 36, pages 99–134.

Montagna, C. P. and P. Papale (2018). *Time scales of shallow magma chamber replenishment at Campi Flegrei caldera.* EarthArXiv https://osf.io/bze9w. DOI: 10.17605/OSF.IO/KGT5P.

Montagna, C. P., P. Papale, and A. Longo (2015). "Timescales of mingling in shallow magmatic reservoirs". *Geological Society, London, Special Publications* 422, SP422–6. DOI: 10.1144/SP422.6.

– (2022). "Magma Chamber Dynamics at the Campi Flegrei Caldera, Italy". *Campi Flegrei*. Edited by G. Orsi, M. D'Antonio, and L. Civetta. Series Title: Active Volcanoes of the World. Berlin, Heidelberg: Springer Berlin Heidelberg, pages 201–217. ISBN: 978-3-642-37059-5. DOI: 10.1007/978-3-642-37060-1_7.

Morgavi, D., I. Arienzo, C. P. Montagna, D. Perugini, and D. B. Dingwell (2017). "Magma Mixing: History and Dynamics of an Eruption Trigger". *Volcanic Unrest : From Science to Society*. Edited by J. Gottsmann, J. Neuberg, and B. Scheu. Cham: Springer International Publishing, pages 123–137. DOI: 10.1007/11157_2017_30.

Morgavi, D., F. Arzilli, C. Pritchard, D. Perugini, L. Mancini, P. Larson, and D. B. Dingwell (2016). "The Grizzly Lake complex (Yellowstone Volcano, USA): Mixing between basalt and rhyolite unraveled by microanalysis and X-ray microtomography". *Lithos* 260, pages 457–474. DOI: 10.1016/j.lithos.2016.03.026.

Neave, D. A., J. Maclennan, M. E. Hartley, M. Edmonds, and T. Thordarson (2014). "Crystal storage and transfer in basaltic systems: The Skuggafjöll eruption, Iceland". *Journal of Petrology* 55(12), pages 2311–2346. DOI: 10.1093/petrology/egu058.

Neave, D. A. and K. D. Putirka (2017). "A new clinopyroxene-liquid barometer, and implications for magma storage pressures under Icelandic rift zones". *American Mineralogist* 102(4), pages 777–794. DOI: 10.2138/am-2017-5968.

Newey, W. K. and D. McFadden (1994). "Large sample estimation and hypothesis testing". *Handbook of Econometrics* 4, pages 2111–2245. DOI: 10.1016/S1573-4412(05)80005-4.

Nimis, P. (1995). "A Clinopyroxene Geobarometer for Basaltic Systems Based on Crystal-Structure Modeling". *Contribution to Mineralogy and Petrology* 121(2), pages 115–125. DOI: 10.1007/s004100050093.

Nimis, P. and P. Ulmer (1998). "Clinopyroxene geobarometry of magmatic rocks Part 1: An expanded structural geobarometer for anhydrous and hydrous, basic and ultrabasic systems". *Contribution to Mineralogy and Petrology* 133(1-2), pages 122–135. DOI: 10.1007/s004100050442.

Papale, P., C. P. Montagna, and A. Longo (2017). "Pressure evolution in shallow magma chambers upon buoyancy-driven replenishment". *Geochemistry, Geophysics, Geosystems* 18. DOI: 10.1002/2016GC006731.

Papale, P., R. Moretti, and D. Barbato (2006). "The compositional dependence of the saturation surface of H_2O+CO_2 fluids in silicate melts". *Chemical Geology* 229(1-3), pages 78–95. DOI: 10.1016/j.chemgeo.2006.01.013.

Perugini, D., C. P. De Campos, M. Petrelli, and D. B. Dingwell (2015). "Concentration variance decay during magma mixing: a volcanic chronometer". *Scientific Reports* 5(14225). DOI: 10.1038/srep14225.

Perugini, D. and G. Poli (2005). "Viscous fingering during replenishment of felsic magma chambers by continuous inputs of mafic magmas: Field evidence and fluid-mechanics experiments". *Geology* 33(1), pages 5–8. DOI: 10.1130/G21075.1.

Perugini, D., G. Poli, M. Petrelli, C. P. de Campos, and D. B. Dingwell (2010). "Time-scales of recent Phlegrean Fields eruptions inferred from the application of a 'diffusive fractionation' model of trace elements". *Bulletin of Volcanology* 72(4), pages 431–447. DOI: 10.1007/s00445-009-0329-z.

Petrelli, M., L. Caricchi, and D. Perugini (2020). "Machine learning thermo-barometry: Application to clinopyroxene-bearing magmas". *Journal of Geophysical Research: Solid Earth* 125(9). DOI: 10.1029/2020JB020130.

Petrone, C. M., E. Braschi, L. Francalanci, M. Casalini, and S. Tommasini (2018). "Rapid mixing and short storage timescale in the magma dynamics of a steady-state volcano". *Earth and Planetary Science Letters* 492, pages 206–221. DOI: 10.1016/j.epsl.2018.03.055.

Pistone, M., L. Caricchi, P. Ulmer, L. Burlini, P. Ardia, E. Reusser, F. Marone, and L. Arbaret (2012). "Deformation experiments of bubble- and crystal-bearing magmas: Rheological and microstructural analysis". *Journal of Geophysical Research: Solid Earth* 117(B05208). DOI: 10.1029/2011JB008986.

Poland, M. P. and E. de Zeeuw-van Dalfsen (2021). "Volcano geodesy: A critical tool for assessing the state of volcanoes and their potential for hazardous eruptive activity". *Forecasting and Planning for Volcanic Hazards, Risks, and Disasters*. Edited by P. Papale. Elsevier, pages 75–115. ISBN: 978-0-12-818082-2. DOI: 10.1016/B978-0-12-818082-2.00003-2.

Putirka, K. D. (2008). "Thermometers and barometers for volcanic systems". *Reviews in Mineralogy and Geochemistry* 69(1), pages 61–120. DOI: 10.2138/rmg.2008.69.3.

Reid, R. C., J. Prausnitz, and T. Sherwood (1977). *The Properties of Gases and Liquids*. 3^{rd} edition. International Series in the Physical and Chemical Engineering Sciences. New York: McGraw Hill.

Ridolfi, F., A. Renzulli, D. Perugini, B. Cesare, R. Braga, and S. Del Moro (2016). "Unravelling the complex interaction between mantle and crustal magmas encoded in the lavas of San Vincenzo (Tuscany, Italy). Part II: Geochemical overview and modelling". *Lithos* 244, pages 233–249. DOI: 10.1016/j.lithos.2015.11.002.

Rummel, L., B. J. P. Kaus, T. S. Baumann, R. W. White, and N. Riel (2020). "Insights into the compositional evolution of crustal magmatic systems from coupled petrological-geodynamical models". *Journal of Petrology* 61(2). DOI: 10.1093/petrology/egaa029.

Sacks, J., W. J. Welch, T. J. Mitchell, and H. P. Wynn (1989). "Design and analysis of computer experiments". *Statistical Science* 4(4), pages 409–423. DOI: 10.1214/ss/1177012413.

Salter, J. M. and D. Williamson (2016). "A comparison of statistical emulation methodologies for multi-wave calibration of environmental models". *Environmetrics* 27(8), pages 507–523. DOI: 10.1002/env.2405.

Santner, T. J., B. J. Williams, and W. Notz (2003). *The Design and Analysis of Computer Experiments*. Berlin: Springer Science & Business Media. DOI: 10.1007/978-1-4757-3799-8.

Selva, J., W. Marzocchi, P. Papale, and L. Sandri (2012). "Operational eruption forecasting at high-risk volcanoes: the case of Campi Flegrei, Naples". *Journal of Applied Volcanology* 1(5). DOI: 10.1186/2191-5040-1-5.

Sigurdsson, H., B. Houghton, S. R. McNutt, H. Rymer, and J. Stix (2015). *The Encyclopedia of Volcanoes*. 2^{nd} edition. Amsterdam: Elsevier. DOI: 10.1016/C2015-0-00175-7.

Stock, M. J., M. C. Humphreys, V. C. Smith, R. Isaia, R. A. Brooker, and Pyle (2018). "Tracking volatile behaviour in sub-volcanic plumbing systems using apatite and glass: Insights into pre-eruptive processes at Campi Flegrei, Italy". *Journal of Petrology* 59(12), pages 2463–2492. DOI: 10.1093/petrology/egy020.

Tonarini, S., M. D'Antonio, M. A. Di Vito, G. Orsi, and A. Carandente (2009). "Geochemical and B–Sr–Nd isotopic evidence for mingling and mixing processes in the magmatic system that fed the Astroni volcano (4.1–3.8 ka) within the Campi Flegrei caldera (southern Italy)". *Lithos* 107(3-4), pages 135–151. DOI: 10.1016/j.lithos.2008.09.012.

Wark, D., W. Hildreth, F. Spear, D. Cherniak, and E. Watson (2007). "Pre-eruption recharge of the Bishop magma system". *Geology* 35(3), page 235. DOI: 10.1130/G23316A.1.

Wei, Z., Z. Qin, and J. Suckale (2022). "Magma mixing during conduit flow is reflected in melt-inclusion data from persistently degassing volcanoes". *Journal of Geophysical Research: Solid Earth* 127(2), e2021JB022799. DOI: 10.1029/2021JB022799.

Yang, Q., E. B. Pitman, E. Spiller, M. Bursik, and A. Bevilacqua (2020). "Novel statistical emulator construction for volcanic ash transport model Ash3d with physically motivated measures". *Proceedings of the Royal Society A* 476(2242), page 20200161. DOI: 10.1098/rspa.2020.0161.

Zellmer, G. F., Y. Iizuka, C. Lormand, and A. Moebis (2020). "Beyond crystal mushes: evidence for uptake of high-T pyroxene antecrysts from mid- to upper crustal andesites into tephras from the Central Plateau, New Zealand". *New Zealand Journal of Geology and Geophysics* 64(2-3), pages 443–455. DOI: 10.1080/00288306.2020.1848883.

Zhang, Y. and T. Gan (2022). "Diffusion in melts and magmas". *Reviews in Mineralogy and Geochemistry* 87(1), pages 283–337. DOI: 10.2138/rmg.2022.87.07.

Zhong, X., M. Dabrowski, and B. Jamtveit (2019). "Analytical solution for the stress field in elastic half-space with a spherical pressurized cavity or inclusion containing eigenstrain". *Geophysical Journal International* 216(2), pages 1100–1115. DOI: 10.1093/gji/ggy447.