Supporting Information for "Physics-based forecasts of eruptive vent locations at calderas"

L. Mantiloni [1,2], E. Rivalta [1,3], K. R. Anderson [6], T. Davis [4], L. Passarelli [5]

[1]GFZ German Research Centre for Geosciences, Section 2.1 'Physics of Earthquakes and Volcanoes', Telegrafenberg, 14473 Potsdam, Germany.

[2]University of Potsdam, Potsdam, Germany

[3]University of Bologna, Bologna, Italy

[4]U.S. Geological Survey, Moffett Field, CA, USA

[5]University of Oxford, Oxford, United Kingdom

[6]INGV Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna, Bologna, Italy

Contents of this file

Introduction

The present Supporting Information (SI) includes results from a test where stresses calculated with the the compound-stress method are compared to those retrieved by the single-mesh method (Figure S1). The mathematical expressions of the probability density functions (PDFs) used to describe dike nucleation zones in the main text are also provided, as well as details about the Markov Chain Monte Carlo (MCMC) algorithm employed in

the stress inversions. The SI also includes plots of the joint posterior PDFs of the stress parameters and ratios retrieved from the stress inversions (Figures S2, S3 and S4), as well as the distributions of the vertical coordinates of the points where SAM backtracked dike trajectories stop in the MCMC chains, B^{end}, compared to the vertical coordinates of the points of highest dike nucleation probability, z^S (Figure S5).

S1: Limitations of the compound-stress method

In Figure S1 we test the compound-stress method against the single-mesh method (blue and red contour lines, respectively). We consider two end-member cases for the Elliptic-Caldera scenario (Figure 2c in the main text), fixing d and h so that they are shallow/short ($d = -50$ m, $h = 50$ m, Figure S1a, b), or deep/tall ($d = -500$ m, $h = 500$ m, Figure S1c, d). We use intermediate $d^{ref} = -225$ m and $h^{ref} = 225$ m.

The orientations of $\vec{v}_3$ (least-compressive principal stress axis) and the gradients of σ_3 (least-compressive principal stress magnitude) in the x, y, and z directions, which control the orientation and driving pressure of the SAM penny-shaped cracks, are similar for the two methods when the caldera and resurgent dome are shallow/short (Figure S1a, b). In contrast, when d and h are large (Figure S1c, d), we see some discrepancies, especially in the region below the resurgent dome, marked by dark-red boxes in Figure S1a, c. Discrepancies for $\vec{v}_3$ orientations are largest, while we see some discrepancies in the σ_3 gradients also outside of the boxes, limited to a few kPa/m, which is small compared to the loading pressure of either resurgent domes ($\rho_r g d =\sim 1.2$ MPa for the short dome, with $\rho_r = 2500$ kg/m^3).

S2: Dike nucleation zone PDFs

We consider three types of p: a trivariate normal distribution (TND), a torus-shaped distribution (TSD), and a skewed, trivariate normal distribution (STND). They are examples of possible choices for p, each applying to different assumptions on the reservoir size and shape, as well as where dikes nucleate from. Examples of the three PDFs are illustrated in Figure 1b in the main text.

The TND PDF (e.g. Prince, 2012) is written as:

$$p^{TND}(x,y,z) = \frac{1}{(2\pi)^{\frac{3}{2}}|\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\vec{x}-\vec{\mu})^T\Sigma^{-1}(\vec{x}-\vec{\mu})\right] \tag{1}$$

where Σ is the covariance matrix, $|\Sigma|$ its determinant, $\vec{x} = (x,y,z)$ and $\vec{\mu}$ is a vector identifying the mean of the distribution, that is, the center of the reservoir, with components (μ_x, μ_y, μ_z). The matrix Σ, which we always assume diagonal, accounts for the uncertainty on the size of the reservoir. When we employ the TND, we assume that dikes may have nucleated from anywhere in the magma storage region, with higher probability closer to its center.

The TSD PDF is obtained from a bivariate normal distribution defined over the r, z cylindrical coordinates, normalized by 2π:

$$p^{TSD}(r,z) = \frac{1}{(2\pi)^{2}|\Sigma|^{\frac{1}{2}}} \exp\left[-\frac{1}{2}(\vec{q}-\vec{\mu})^T\Sigma^{-1}(\vec{q}-\vec{\mu})\right] \tag{2}$$

where Σ is the covariance matrix, $\vec{q} = (r,z)$ and $\vec{\mu}$ is the mean of the distribution, with components (μ_r, μ_z). We assume a diagonal Σ here as well. When we employ the TSD, we assign the highest probability of dike nucleation on the boundary of a sill-like reservoir.

We use the STND to assign higher probability of dike nucleation to the shallower region of a more vertically elongated reservoir. The STND PDF (O'Hagan & Leonard, 1976) is

written as:

$$p^{STND}(x,y,z) = 2p^{TND}(x,y,z)\Phi(\vec{\lambda},\vec{x}) \tag{3}$$

where $p^{TND}(x,y,z)$ is the TND described in Equation 1, and $\Phi(\vec{\lambda},\vec{x})$ is the cumulative distribution of a standard normal PDF, where the components of $\vec{x}$ are scaled by the shape parameters $\vec{\lambda} = (\lambda_x, \lambda_y, \lambda_z)$:

$$\Phi(\vec{\lambda},\vec{x}) = \frac{1}{2}\left[1 + erf\left(\frac{\vec{\lambda}\cdot\vec{x}}{\sqrt{2}}\right)\right] \tag{4}$$

We remark that such PDFs are non-truncated, as p is non-zero everywhere. While this formally means that we assign a non-zero probability of dike nucleation above the free surface, it poses no issue, since the dike pathways model we employ (SAM, Mantiloni et al., 2023) may backtrack dikes only in the subsurface.

S3: Additional information on the Markov Chain Monte Carlo (MCMC) algorithm

S3.1: Cost function of the MCMC algorithm

The stress parameter sampling procedure described in the main text relies on maximizing the joint probability density reported in Equation 8. This is equivalent to minimizing a cost function, s, defined as:

$$s = -2\sum_{k=1}^{N} \log\left(p_k\right) \tag{5}$$

(see e.g. Laine, 2008), where N is the number of vents, and each p_k corresponds to a potential dike starting point, S_k, given a set of stress parameters. s is the cost function used by the MCMC algorithm to sample the stress parameter space.

S3.2: Choice of M

In the MCMC algorithm we use in our stress inversions, M is number of steps of the MCMC chain. Following the Delayed Rejection and Adaptive Metropolis MCMC algorithm by (Haario et al., 2006), M includes an initial burn-in time, M_{burn}, to allow the MCMC to reach a stationary regime. The Delayed Rejection and Adaptive Metropolis algorithms are not applied during the burn-in time.

Our choice of M is informed by the smallest number of steps so that the posterior Probability Density Functions (PDFs) of the stress parameters do not change significantly from one inversion to another if M is increased. In our tests we found that, for most scenarios, $M = 2,000$, with a burn-in time of $M_{burn} = 200$ steps, was enough to obtain posterior PDFs consistent with those retrieved by setting M to much larger values. We deemed that $M = 20,000$ offered the best compromise between allowing the chain as much time as needed to sample the PDFs effectively and keeping the running time of the algorithm as low as possible.

We set a burn-in time $M_{burn} = M/10$ in all our stress inversions. We discard the portions of the MCMC chains corresponding to the burn-in time, and do not include them in the posterior PDFs shown in Figures 3 and 4 of the main text.

References

Haario, H., Laine, M., Mira, A., & Saksman, E. (2006). DRAM: efficient adaptive MCMC. *Statistics and Computing*, *16*(4), 339–354.

Laine, M. (2008). *Adaptive MCMC methods with applications in environmental and geophysical models*. Finnish Meteorological Institute.

Mantiloni, L., Rivalta, E., & Davis, T. (2023). Mechanical modeling of pre-eruptive magma propagation scenarios at calderas. *Journal of Geophysical Research: Solid Earth*, *128*(3), e2022JB025956.

O'Hagan, A., & Leonard, T. (1976). Bayes estimation subject to uncertainty about parameter constraints. *Biometrika*, *63*(1), 201–203.

Prince, S. J. (2012). *Computer vision: models, learning, and inference*. Cambridge University Press.

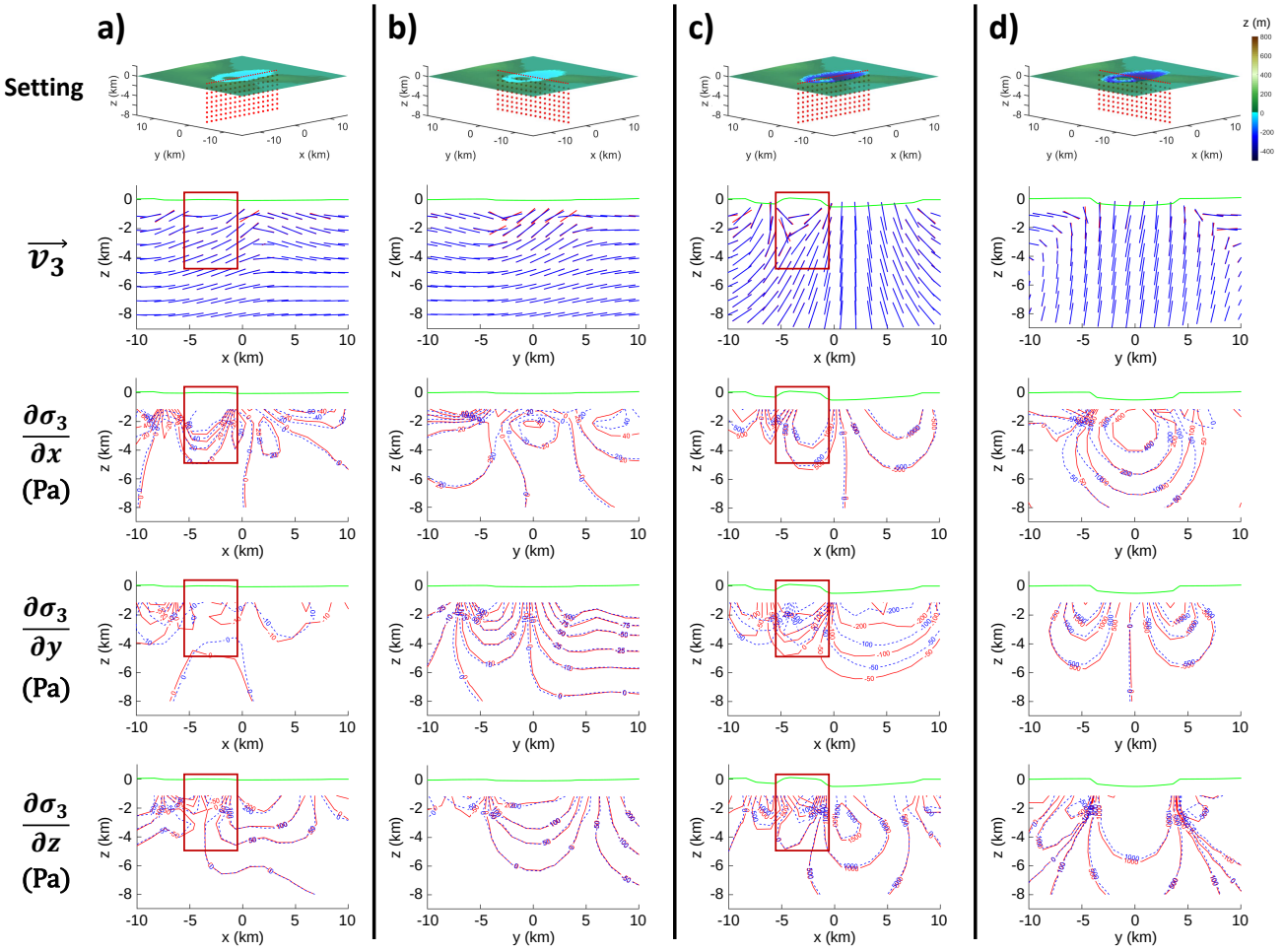

Figure S1: Comparison of $\vec{v}_3$ directions and σ_3 gradients calculated for the Elliptic-Caldera scenario (Mantiloni et al., 2023) with the single-mesh (red) and compound-stress methods (blue). Gradients are calculated by subtracting σ_3 matrices shifted by 1 km in the three directions. Such a distance is comparable to the SAM radii (c) employed in the main text. Green profiles in the stress plots are the topographic profiles shown in the first row as red dotted lines on the meshes. Observation points are spaced by 400 m along x and y and by 300 m along z. We do not include points shallower than 800 m below the mesh, as we do not propagate dikes there. All columns include, from top to bottom: view of the setting and observation points; $\vec{v}_3$ directions; σ_3 gradient along the x, y and z direction. (a) $d = -50$ m, $h = 50$ m; observation points are on the $y = 0$ plane. (b) Same as (a), observation points are on the $x = 0$ plane. (c) $d = -500$ m, $h = 500$ m; observation points are on the $y = 0$ plane. (d) Same as (c), observation points are on the $x = 0$ plane. The dark-red boxes in columns (a) and (c) mark the width of the resurgent dome (~ 5 km) and the depth interval where we expect the stress contribution of the resurgent dome to be significant.

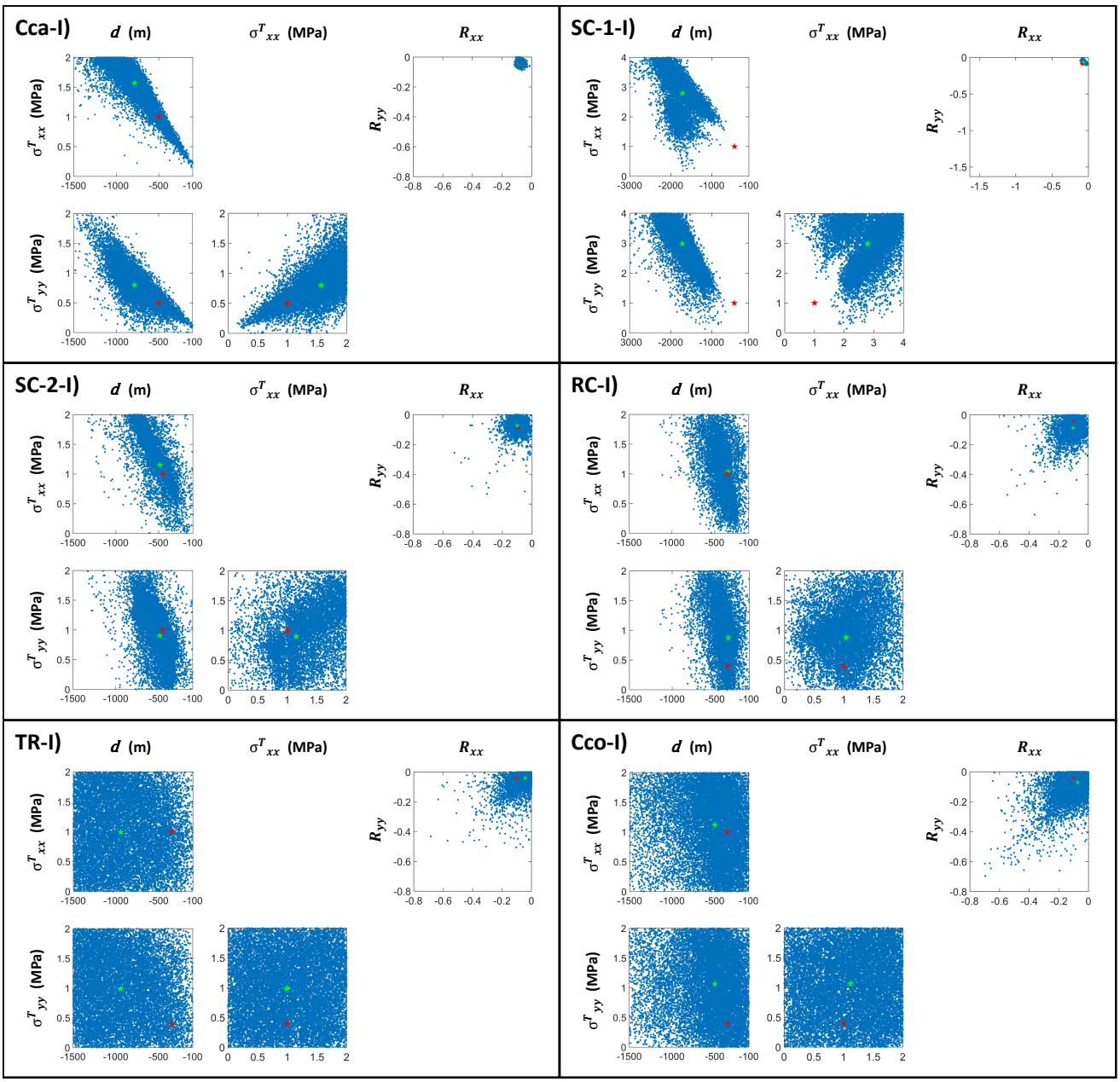

Figure S2: Joint posterior PDFs of stress parameters and ratios from the stress inversions considering three parameters (see Figure 3 in the main text). Red and green stars represent, respectively, the median and original values of the stress parameters.

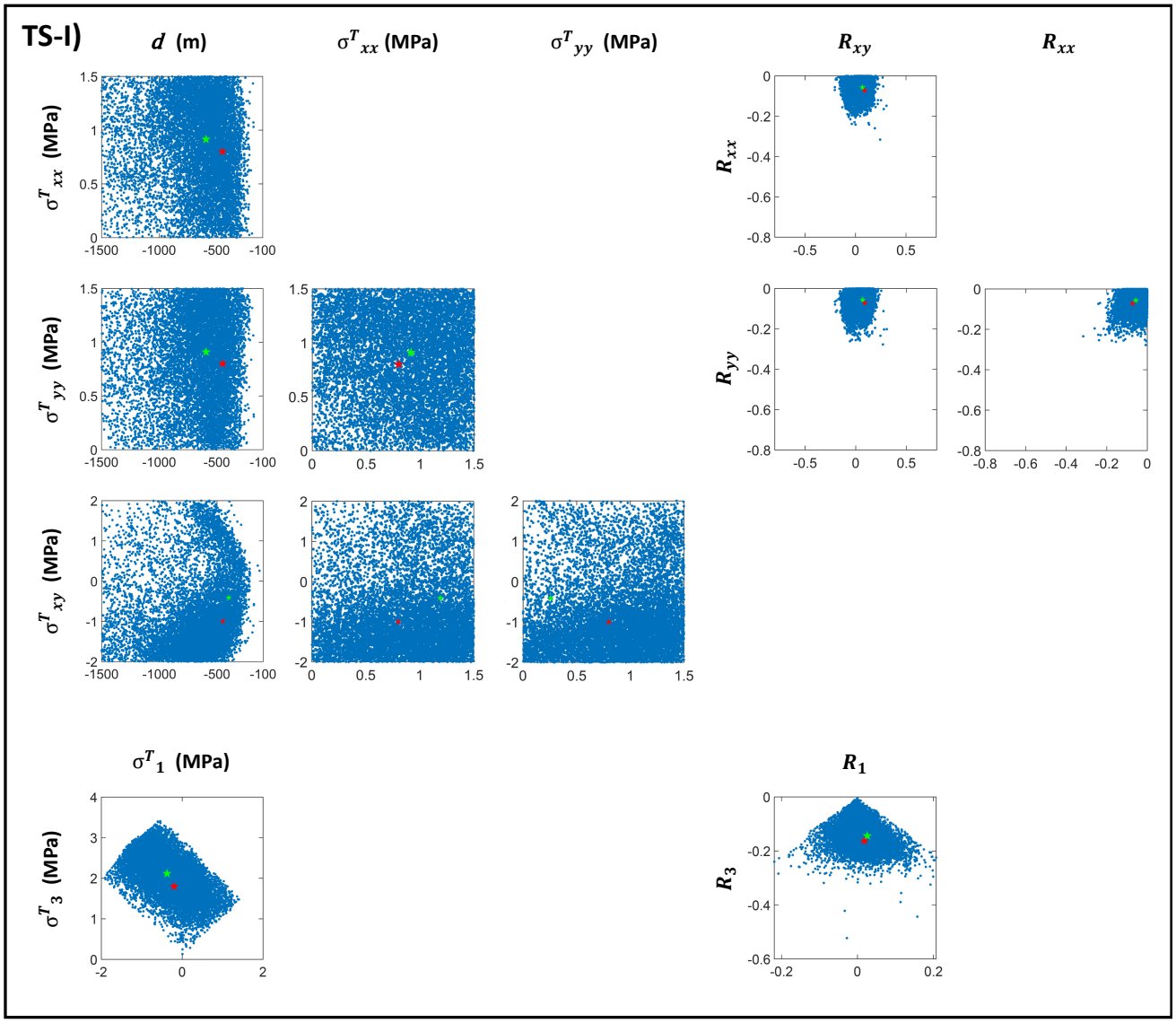

Figure S3: Joint posterior PDFs of stress parameters and ratios from the TS-I stress inversion (see Figure 4a in the main text). Red and green stars represent, respectively, the median and original values of the stress parameters.

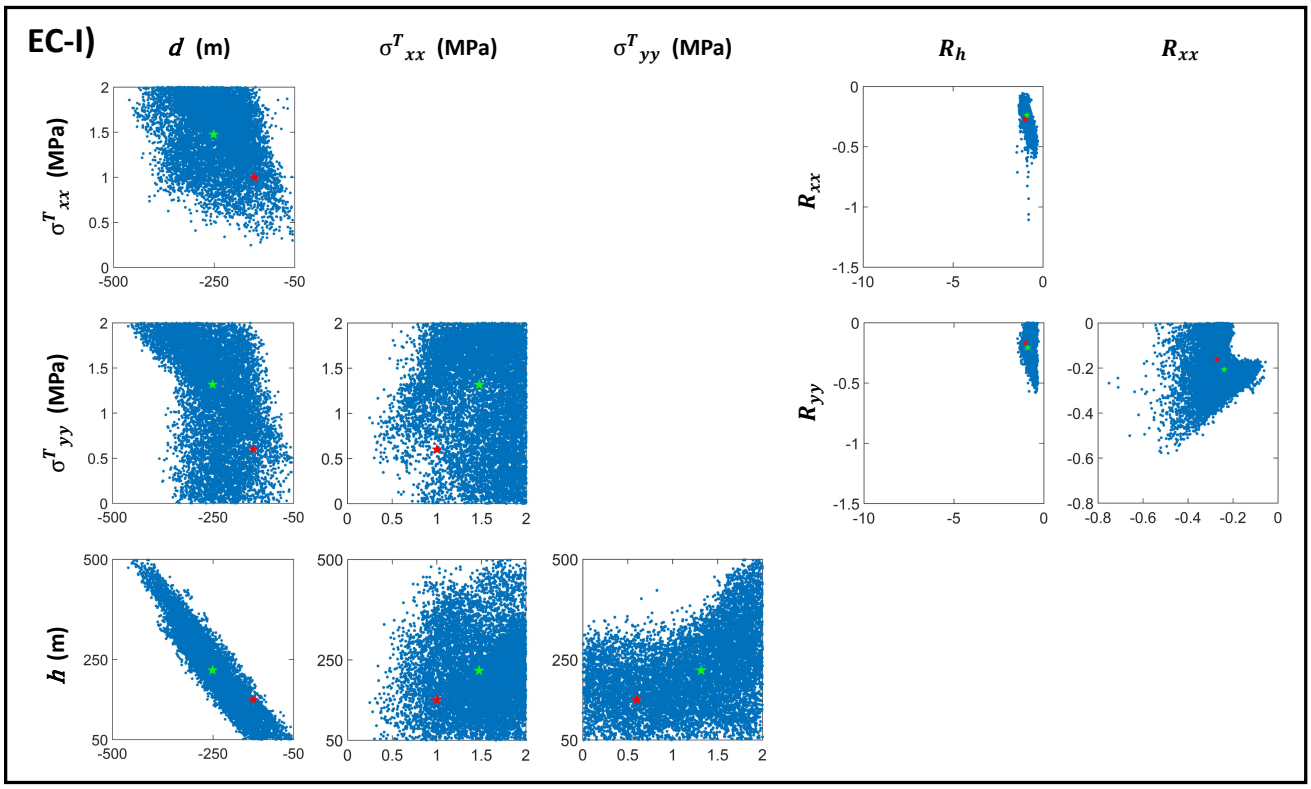

Figure S4: Joint posterior PDFs of stress parameters and ratios from the EC-I stress inversion (see Figure 4b in the main text). Red and green stars represent, respectively, the median and original values of the stress parameters.

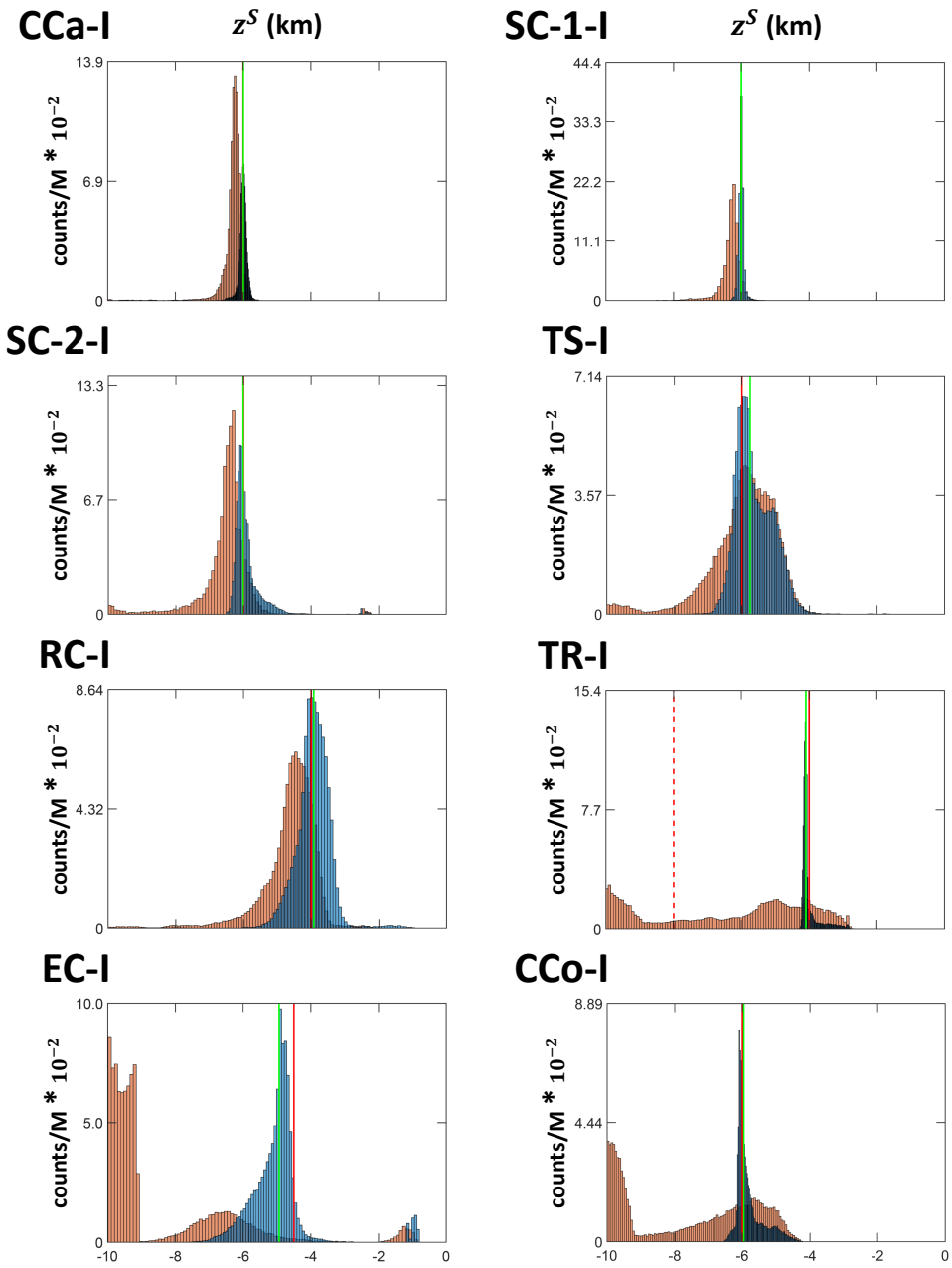

Figure S5: Posterior PDFs of the vertical coordinates of the points of maximum dike nucleation probability along SAM backtracked trajectories (z^S, blue) and distributions of the vertical coordinates where SAM trajectories stop (B^{end}, orange). Red and green lines mark, respectively, the median of the blue PDFs and the depth of the dike starting points in the original models by Mantiloni et al. (2023). In TR-I, the dashed red line marks the depth of the reservoir neglected in the stress inversion.