Supplementary Information for Deep Subsurface Water Stores Sustain Giant Sequoias

The results in the main text are derived using technical methods including full waveform inversion (FWI) and geostatistical rock physics inversion. Here we provide mathematical background, technical details, and benchmarking for these methods. The first several sections of the supplementary information present more mathematical background, technical details, and validation of the FWI methods used in this work. The latter subsections provide more details about the rock physics models and geostatistical inversion algorithms we used. We intend for these subsections to provide the information needed to make our methods and processing techniques intelligible and reproducible to the interested reader. Upon Publication of this work, we will also upload the code and data sets to an online Zenodo repository with a citable DOI.

Signal Processing Preliminaries

For context, we begin by defining the Fourier transform, $\mathcal{F}[y(t)]: \mathbb{R} \to \mathbb{C}$, and its inverse, $\mathcal{F}^{-1}$, for real-valued time series

$$\mathcal{F}[y(t)] = \hat{y}(\xi) = \int_{\mathbb{R}} y\, e^{-i2\pi\xi t}\, dt$$

$$\mathcal{F}^{-1}[\hat{y}(\xi)] = y(t) = \int_{\mathbb{R}} \hat{y}\, e^{i2\pi\xi t}\, d\xi$$

where the notation $\int_{\mathbb{R}}$ implies that we integrate over the real numbers because $t, \xi \in \mathbb{R}$. Note that here we follow common notation by having t and ξ represent time and frequency respectively. One assumption implicit to the Fourier transform is stationarity meaning that the frequency content of the signal is assumed to not change with time. For most active source seismic data, this assumption is false. In fact, analyzing the changing frequency content of data with time is often a means of gaining insight into the Earth's structure and function [1–4]

For example, it has been common practice to use the Gabor transform (also known as the short time Fourier transform) given by $G(\xi,\tau) = \mathcal{G}[y(t)] = \int_{\mathbb{R}} h\,(t-\tau) y(t) e^{-i2\pi\xi t}\, dt$ to analyze the changing frequency content of a seismogram. With respect to the Gabor transform, $h(t)$ is a windowing function that enables time-frequency analysis yet also poses certain trade-offs. For one, multiplying by $h(t)$ alters the frequency content of the input signal. Secondly, choosing a windowing function that will result in optimal time-frequency resolution is difficult. While a more localized time window may enable the elucidation of frequency content changes occurring over relatively short periods of time, it will also elide low-frequency information. Alternatively, a less localized time window will account for lower-frequency information, but may be unable to resolve short-time changes in frequency content [2].

The continuous wavelet transform (CWT) has been employed to better analyze the changing frequency content of time series data [3,5]. The CWT does not assume that the time series is stationary and can achieve good time-frequency resolution at low and high frequencies simultaneously [5]. The CWT is based on some mother wavelet $\psi(t) \in \mathbb{C}$ and is defined with respect to the scale factor $s \in \mathbb{R}^+$. For each scale factor, we consider a modified form of the mother wavelet such that it is stretched in time and scaled in amplitude

$$\psi_s(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t}{s}\right)$$

(Note the relationship between the scale factor and frequency $\xi = \xi_0/s$ where ξ_0 is the centroid frequency of the mother wavelet). While there are many possible choices of mother wavelet, in this study, we employ a complex Ricker wavelet which can be computed in the frequency domain using

$$\widehat{w}(\xi) = \frac{4\xi^2}{\sqrt{\pi}\xi_0^3}e^{-\xi^2/\xi_0^2}, \quad \forall\ \xi \geq 0.$$

Subsequently, we may specify the CWT as a function with the mapping $\mathcal{W}[y(t)]: \mathbb{R} \to \mathbb{C}$ where $y(t)$ is a real-valued time-series. Note that defining the CWT with this particular mapping implies that the mother wavelet should be complex-valued. The CWT can be defined using

$$\mathcal{W}[y(t)] = W_y(s,\tau) = \int y(t)\psi_s^*(t-\tau)\ dt$$

where * denotes the complex conjugate. In practice, it is more efficient to compute the CWT in the frequency domain [5] according to

$$\widehat{W}_y(s,\xi) = \hat{y}(\xi)\hat{\psi}_s^*(\xi).$$

Multiple formulations of the inverse CWT exist with the most simplistic theoretically being

$$\mathcal{W}^{-1}[W_y(s,\tau)] = y(t) = \mathcal{R}\left\{\frac{1}{c_\psi}\int_{\mathbb{R}^+}\int_{\mathbb{R}} W_y\,(s,\tau)\frac{1}{s^2}\psi_s(t-\tau)\ d\tau\ ds\right\}$$

where c_ψ is an admissibility constant equal to

$$c_\psi = \frac{1}{\sqrt{2\pi}}\int \frac{|\hat{\psi}(\xi)|^2}{|\xi|}\ d\xi < \infty$$

while $\mathcal{R}\{.\}$ and $\mathcal{I}\{.\}$ are operators that take the real and imaginary parts of complex numbers respectively. Note that the presence of the $\mathcal{R}\{.\}$ operator in the inverse CWT enforces the mapping $\mathcal{W}^{-1}[W_y(s,\tau)] = y(t): \mathbb{C} \to \mathbb{R}$. Another formulation of the inverse CWT can be derived by using a different mother wavelet to reconstruct the original signal from its CWT [3,5]. Using this approach, it is common practice to choose the delta wavelet [3,5] which yields

$$\mathcal{W}^{-1}[W_y(s,\tau)] = y(t) = \mathcal{R}\left\{\frac{1}{c_{\psi,\delta}}\int_{\mathbb{R}^+} W_y\,(s,\tau)s^{-1.5}\ ds\right\}$$

where the new admissibility constant becomes

$$c_{\psi,\delta} = 2\sqrt{\pi}\int \frac{\hat{\psi}^*(\xi)\hat{\delta}(\xi)}{|\xi|}\ d\xi < \infty.$$

The latter implementation of the inverse CWT because it provides a more computationally efficient means of reconstructing the signal [5].

The Time-Frequency-Phase Misfit Functional and its Adjoint Source

FWI minimizes some misfit function $\chi \in \mathbb{R}$ that can generally be written as

$$\chi = \sum_{i=1}^{N_s} \sum_{j=1}^{N_r} \sum_{k=1}^{N_c} \langle\, E[u_{ijk}(t), d_{ijk}(t)] \rangle$$

where $\langle . \rangle$ indicates integration over time and i, j, k are indices for the source, receiver, and component numbers respectively. This means that the the subscripts on the variable u_{ijk} indicate that it is the kth component of the synthetic wavefield corresponding to the ith source sampled at the jth receiver position. This is otherwise stated as $u_{ijk}(t) = \mathbf{e}_k \cdot \mathbf{u}_i(\mathbf{x}, t)\delta(\mathbf{x} - \mathbf{x}_j)$ where $\mathbf{x}$ is a coordinate vector, $\mathbf{e}_k$ is a unit vector in the direction of the kth component, $\mathbf{u}_i$ is the vector-valued wavefield excited by the ith source, and $\mathbf{x}_j$ is the location of the jth receiver position. The same goes for the variable d_{ijk} which corresponds to observed data. For the sake of notational simplicity, in the rest of this subsection, we consider the case where $N_s = N_r = N_c = 1$ which allows us to elide the subscripts and summations across components, sources, and receivers. We also note that in previous work, it has been shown that the the adjoint source corresponding to χ can be given by

$$\partial_u \chi = \langle f^\dagger \rangle$$

[2,6–9]. This being the case, as an example we may consider the misfit functional from classical FWI where χ is taken the be the L2 norm meaning that $E = \frac{1}{2}(u - d)^2 \Longrightarrow$

$$\chi_{L2} = \frac{1}{2} \int (u - d)^2 \, dt.$$

Therefore, one can use the Leibniz integration rule (also known as differentiating under the integral) to show that the corresponding adjoint source has the form

$$f_{L2}^\dagger = u - d.$$

Here we adapt the misfit functional developed in Fichtner et al. (2008) which employs the phase difference between the time-frequency representation of the observed and predicted data. Our slightly altered version of this misfit functional is given by

$$\chi = \mathcal{R}\left\{ \int \int \frac{1}{s^2} P \Delta\phi^2 \, ds \, d\tau \right\}.$$

In the above equation, the phase difference, $\Delta\phi$, can be written as

$$\Delta\phi(s, \tau) = \mathcal{I}\{\ln(W_u W_d^*)\} = \phi_u - \phi_d$$

where ϕ_u and ϕ_d are the phase responses of the synthetic and observed data respectively while $P(s, \tau)$ is an arbitrary weighting function in the s-τ domain. Following Fichtner et al. (2008) and in preparation of deriving the adjoint source, we evaluate $\partial_u \Delta\phi$ by applying the chain rule and invoking the definition of the CWT.

$$\partial_u(\phi_u - \phi_d) = \partial_{W_u} \phi_u \times \partial_u W_u = \mathcal{I}\left\{ \frac{1}{W_u} \right\} \times \int \psi_s^*(t - \tau) \, dt$$

At this point, we may also note the identity

$$\mathcal{I}\left\{\frac{1}{W_u}\right\} = \mathcal{I}\left\{\frac{W_u^*}{W_u^* W_u}\right\} = -\mathcal{I}\left\{\frac{W_u}{W_u^* W_u}\right\}.$$

Now, we differentiate χ with respect to the predicted data using the Leibniz integration and chain rules

$$\partial_u \chi = \mathcal{R}\left\{\int\int 2\frac{1}{s^2} P\Delta\phi \times \partial_u \Delta\phi \; ds \; d\tau\right\}$$

Which in turn yields

$$\langle f^\dagger \rangle = \mathcal{R}\left\{\int\int -2\frac{1}{s^2} P\Delta\phi \mathcal{I}\left\{\frac{W_u}{W_u^* W_u}\right\} \int \psi_s^*(t-\tau) \; dt \; ds \; d\tau\right\}.$$

Some concerns may arise from the above the expression for the adjoint source due to instabilities caused via dividing by the $W_u^* W_u$ term. Hence, one purpose of the weighting function, P, becomes clear - to stabilize this division via multiplication. Another use of the weighting function may be to de-noise the phase measurement $\Delta\phi$. Since the phase values of a signal may be large even when the signals' amplitude is small, it is advantageous to only take phase difference measurements where both the observed and synthetic signals are above some threshold. Fichtner et al. (2008) use a weighting term in a similar manner however, here we use a use an alternative formulation by setting

$$P := \frac{1}{2c_{\psi^*}} \Gamma(s,\tau) \frac{W_u^* W_u}{\| W_u \| + \nu}.$$

Note the now the $W_u^* W_u$ term multiplies to one with the denominator of the adjoint source while we have the $\Gamma(s,\tau)$ term mute phase measurements made using low-amplitude signals by setting

$$\Gamma := \frac{1}{1 + e^{-\epsilon C + \eta}}$$

with $C(s,\tau) = \| W_s W_d^* \| / \max\{\| W_s W_d^* \|\}$. One may observe that Γ is a sigmoid function with regards to C and we find that effective values for the parameters ϵ and η are 10^1-10^3 and 0.1-0.25 respectively. Finally, taking the previous expression for the adjoint source, substituting in our prescribed definition for P, and using the definition of the inverse CWT, we derive the following expression for the adjoint source

$$f^\dagger = -\mathcal{W}^{-1}\left[\Gamma\Delta\phi\mathcal{I}\left\{\frac{W_u}{\| W_u \| + \nu}\right\}\right].$$

We remark that $\nu \in \mathbb{R}$ is a water-level parameter to avoid division by zero (sensible values for ν range from 10^{-3}-$10^{-1} \times \max\{\| W_u \|\}$).This formulation of the adjoint source is very similar to that derived in Fichtner et al (2008), albeit with slight differences due to us using the CWT rather than the Gabor transform and due to our different choice of weighting function.

FWI Overview

Full waveform inversion involves computing numerical solutions to the differential system

$$\rho\, \partial_t^2 \mathbf{u} - \nabla \cdot \boldsymbol{\sigma} = \mathbf{f} = \mathbf{L}(\mathbf{u}, m)$$

where $\rho(\mathbf{x})$ is the density field, $\mathbf{u}(\mathbf{x}, t) \in \mathbb{R}^2$ is the displacement vector field, $\boldsymbol{\sigma}(\mathbf{x}, t) \in \mathbb{R}^{2\times 2}$ is the stress tensor field, and $\mathbf{f}(\mathbf{x}, t) \in \mathbb{R}^2$ is the force vector that excites elastic waves. Note that here, we solve the elastic wave equation in two spatial dimensions and that the stress tensor field is given by $\boldsymbol{\sigma} = \mathbf{C}: \frac{1}{2}(\nabla \mathbf{u} + \nabla \mathbf{u}^T)$, where : represents tensor multiplication and $\mathbf{C}(\mathbf{x}, m)$ is the stiffness tensor parameterized by elastic moduli $\lambda(\mathbf{x}, m)$ and $\mu(\mathbf{x}, m)$. In turn, we parameterize the elastic earth model, $m(\mathbf{x})$, using pressure-wave velocity, $Vp = \sqrt{(\lambda + 2\mu)/\rho}$, and shear-wave velocity, $Vs = \sqrt{\lambda/\rho}$. Numerous methods for deriving numerical solutions to (24) exist [10], and here we utilize the spectral element method and the open-source software specfem2d because of its ability to accurately model the free surface boundary condition on irregular topography [11].

Past works such as [6,12,13] have shown that in using the adjoint method, the gradient of the misfit functional with respect to the earth model can be computed using

$$\nabla_m \chi = \sum_{i=1}^{N_S} \int \mathbf{u}_i^{\dagger} \cdot \nabla_m \mathbf{L}_i \, dt$$

where $\mathbf{u}^{\dagger}$ is the adjoint wavefield. The same numerical algorithm used to find $\mathbf{u}(m)$ can be used to compute $\mathbf{u}^{\dagger}$, albeit with the alterations that the adjoint wavefield is excited by the adjoint source, $\mathbf{f}^{\dagger}$, and needs to be solved propagating backwards in time [6,12,14]. In using the adjoint method to compute $\nabla_m \chi$, we may iteratively update the earth model using a gradient descent algorithm via

$$m^{i+1} = m^i - \alpha \mathcal{P} \nabla_m \chi$$

where α is the step length, which can be computed using a backtracking line search [15], and $\mathcal{P}(\mathbf{x})$ is a preconditioner. The preconditioner can be used to improve the convergence of the gradient descent algorithm and one common choice is to set

$$\mathcal{P}^{-1} := \sum_{i=1}^{N_S} \int \partial_t^2 \mathbf{u}_i \cdot \partial_t^2 \mathbf{u}_i \, dt.$$

This formulation of $\mathcal{P}$ helps to counterbalance the effects of having higher wavefield amplitudes in certain parts of the model (e.g., near sources or in low velocity zones) [15,16].

FWI Workflow

For practical applications of FWI, data preprocessing is usually a crucial step which may include filtering and muting of the record sections. Here, we followed other work implementing FWI for CZ study by designing muting windows that would allow the inversion to focus on body waves [17,18]. We also employed a Butterworth bandpass filter to implement a classical frequency continuation multiscale strategy [19]. We sequentially inverted data in frequency bands of 5-20 Hz, 5-35 Hz, 5-50 Hz, and 5-65 Hz which helped improve convergence and avoid local minima. In order to iteratively update the earth model, we computed gradients of the misfit function with respect to Vp and Vs using the adjoint method. We did not update density and parameterized all earth models with a homogeneous $\rho(\mathbf{x})$ field [20–22]. For inversions at each frequency band, we smoothed the gradients by convolving them with two-dimensional Gaussian functions [16,20,23,24]. As we increased the frequency content of the data, we decreased the radii of the Gaussian kernels, using radii of 20 m, 15 m, 10 m, and 5 m respectively for each of the frequency bands.

Before inverting the data, we estimated unique STFs at each shot location using the following equation developed by [25]

$$\hat{f}_i(\xi) = \sum_{j=1}^{N_r}\sum_{k=1}^{N_c} \frac{\hat{u}_{ijk}\hat{d}_{ijk}^{*}}{\hat{u}_{ijk}\hat{u}_{ijk}^{*} + \gamma}$$

where γ is a small real-valued constant to avoid division by zero. Note that the above equation implies that we estimate the STFs in the frequency domain and use both the horizontal and vertical components, meaning the $N_c = 2$. In order to simulate the synthetic data used for the source estimation, we use a 30 Hz first derivative of a Gaussian function. Then, before Fourier transforming the observed and synthetic data, we applied muting around the first arrivals. Furthermore, after obtaining the STFs in the frequency domain and applying the inverse Fourier transform, we multiplied the STFs by a time domain preconditioning function to localize the energy around the expected shot time.

Synthetic FWI Tests

To verify our FWI approach, we implement a synthetic test using the acquisition geometry of the hillslope line. The Vp and Vs initial models are the same used for the real-data inversion. For the synthetic test, we create target models by adding velocity anomalies to the initial models. The results from the synthetic test show to what extent we are able to recover the known velocity models. In the upper 30 m of the models, where the data coverage is highest and the wavelengths are relatively small, we are able to recover the target velocity models and Vp:Vs ratio model well. Below 30 m, there are more errors in the recovered velocity models which are exasperated in the recovered Vp:Vs ratio models. Note that our analysis in the main text focuses on features in the upper 30 m of the subsurface where the velocity models are better constrained.

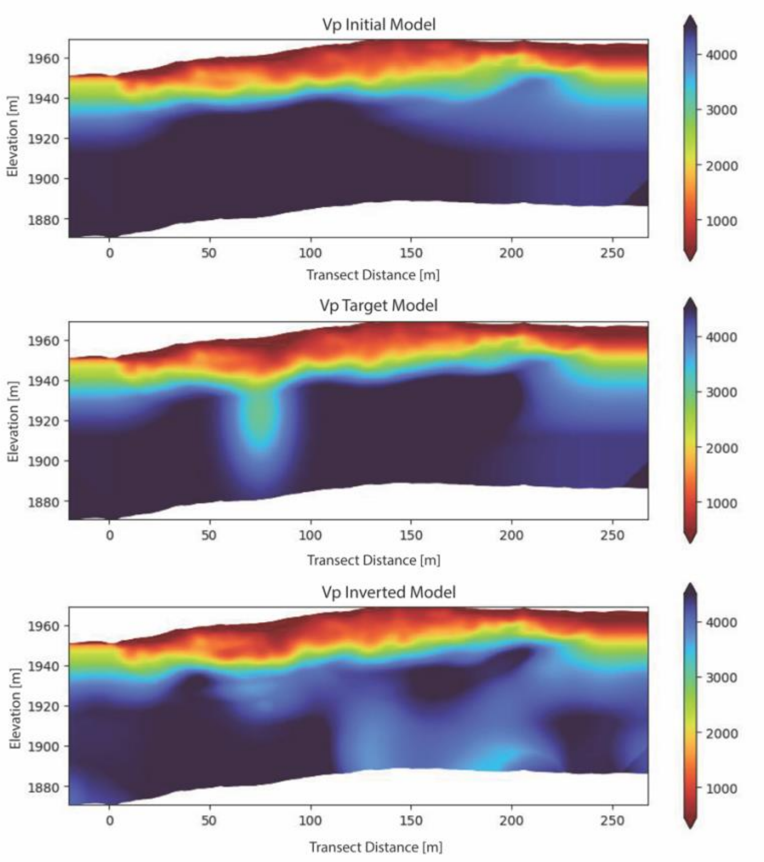

Figure S1: The initial (top row), target (middle row), and recovered (bottom row) Vp models used for the synthetic FWI test.

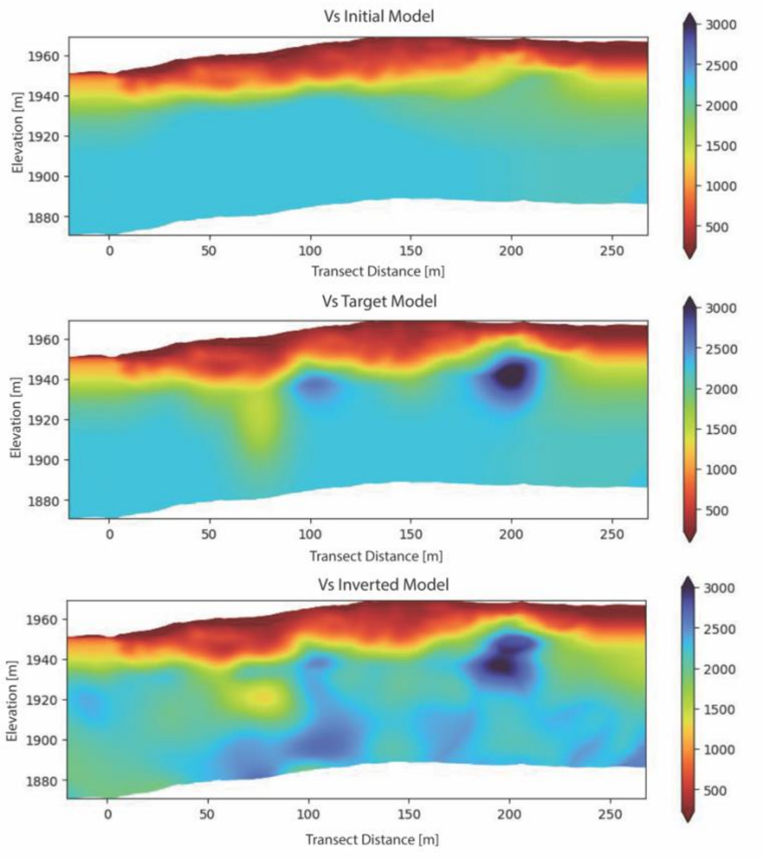

Figure S2: The initial (top row), target (middle row), and recovered (bottom row) Vs models used for the synthetic FWI test.

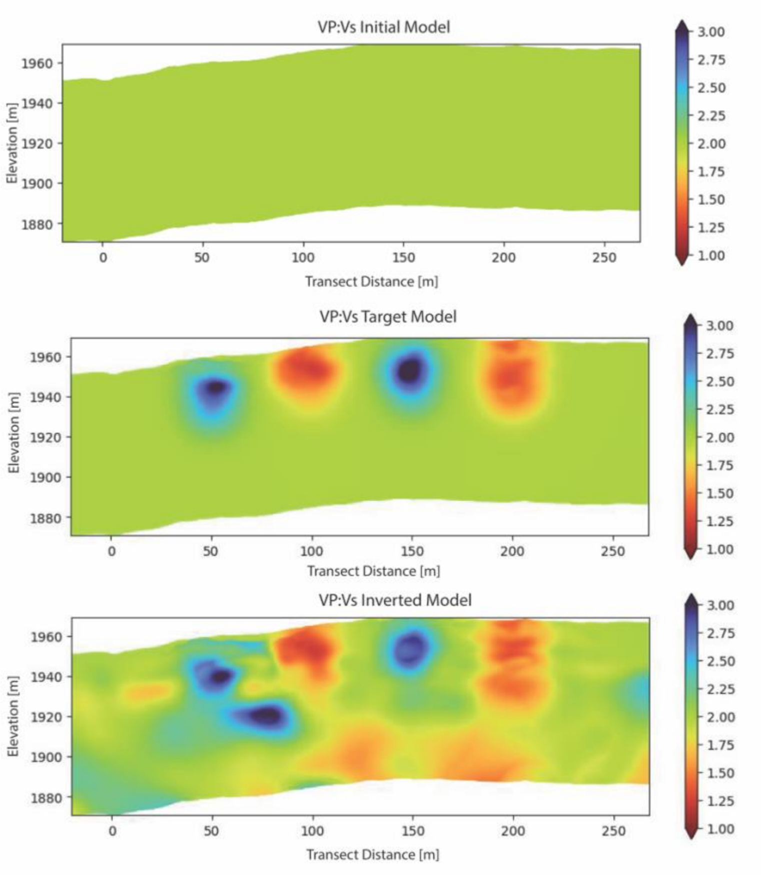

Figure S3: The initial (top row), target (middle row), and recovered (bottom row) Vp:Vs ratio models used for the synthetic FWI test. Note that the "Inverted" Vp:Vs ratio model is not directly constrained by the inversion, but instead, is computed after FWI with the final Vp and Vs models.

FWI Results Data Comparison

Below, we show several example shot gathers of preprocessed (with muting, trace normalization, and 5-65 Hz bandpass filtering) field and synthetic data. The synthetic data are generated from the final FWI velocity models (inverted from the field data). We show one representative shot gather for each component from each seismic profile. For the ridgeline, we show both a horizontal and vertical component shot gather. For the hillslope and drainage profiles, we show only vertical component shot gathers because this was the only component of seismic data collected and inverted. Also note that for all the inversions, we use a maximum source, receiver offset of 100 m, which is why many of the traces are muted in the ridgeline profile shot gathers. For the hillslope and drainage lines, fewer traces are muted because the lines are shorter and have a smaller receiver spacing.

Figure S4:

Figure S5:

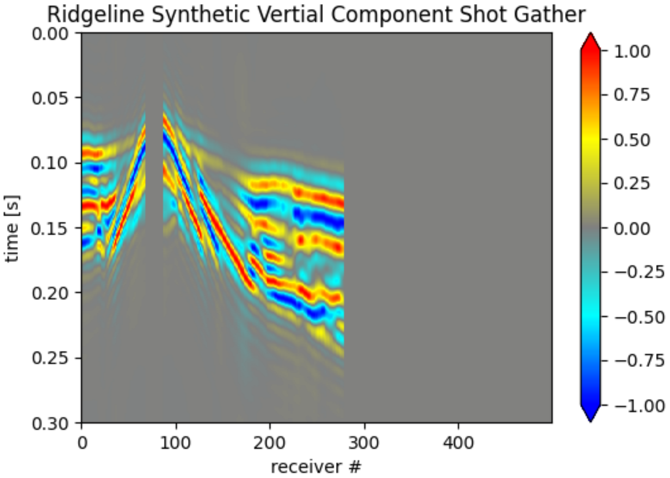

Figure S6:

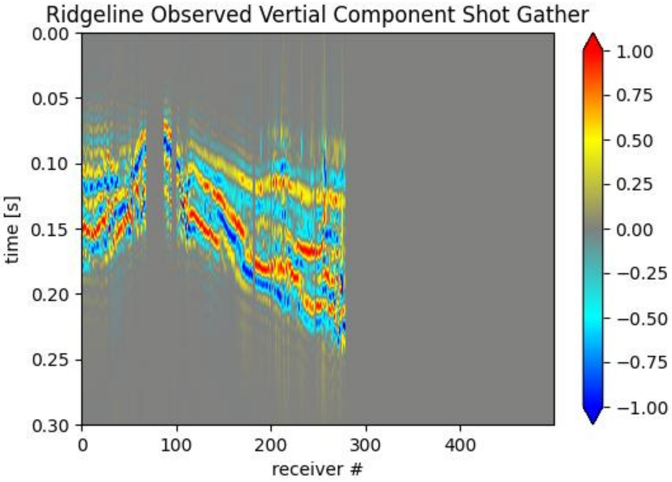

Figure S7:

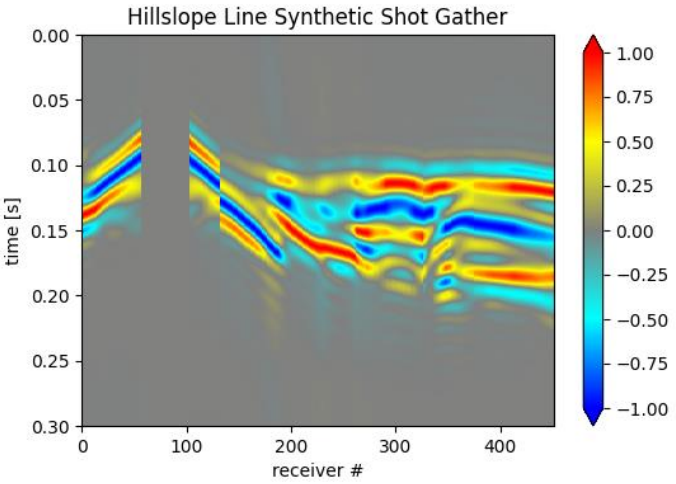

Figure S8:

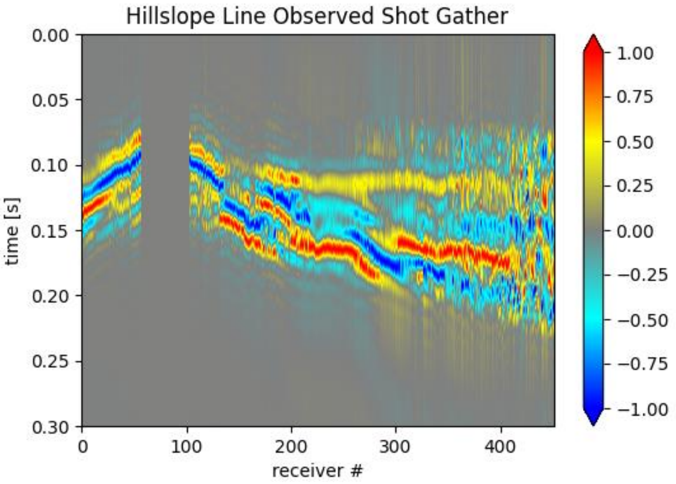

Figure S9:

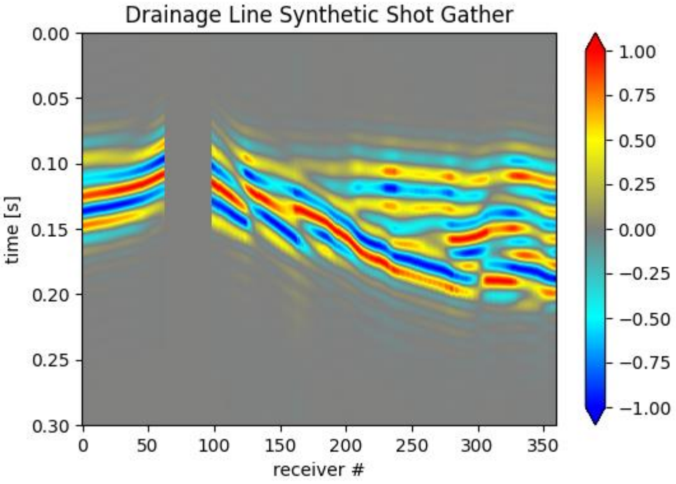

Figure S10:

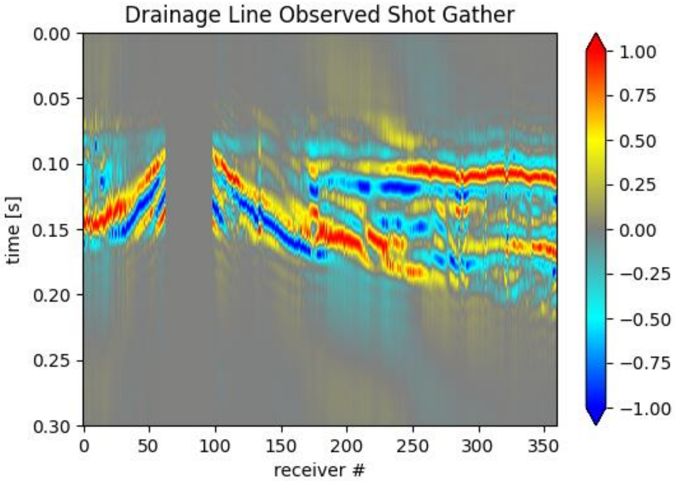

Figure S11:

Rock Physics Model

The rock physics model links geophysical variables p-wave and s-wave velocity to hydrological variables of porosity, φ, and water saturation, s_w. The first order control on seismic velocity in the critical zone is porosity, which can range from ~0.6 is soils to less than 0.01 in unweathered bedrock [26,27]. For the more weathered soils and saprolite, Dvorkin's sand model [28] applies well, while Berryman's inclusion model applies well to low-porosity bedrock [29] but neither rock physics model can accommodate the full range of critical zone materials. For this reason, we rely on an empirical biexponential model that approximates Dvorkin's at high porosity and Berryman's at low porosity. The biexponential model estimates totally unsaturated (meaning that $s_w = 0$) p-wave and s-wave velocity using

$$Vp_{unsat} = a_p exp\left(-b_p\varphi\right) + c_p\left(1 - exp\left(-b_p\varphi\right)\right)$$

$$Vs_{unsat} = a_s exp(-b_s\varphi) + c_s\left(1 - exp(-b_s\varphi)\right)$$

where a_p, b_p, a_d, b_d are empirical constants that can tuned to account for the observed variation in seismic velocity. The totally unsaturated velocities only depend on porosity and latter we use Gassman's equation to account for the effect of variable saturation. The density of the totally unsaturated material can be computed as a linear function of porosity

$$\rho_{unsat} \;=\; \rho_{solid}(1-\varphi) \;+\; \rho_{air}\varphi$$

where ρ_{solid} is the density of the solid phase and ρ_{air} is the density of the gas phase. Using the totally unsaturated values of density and s-wave velocity, the shear modulus of the material can be computed

$$\mu = \rho_{unsat} Vs_{unsat}^2.$$

Note that the shear modulus is only a function of porosity and will not change with saturation, thus the *unsat* subscript is not needed here. Given the shear modulus, the bulk modulus for totally unsaturated material can be calculated using

$$K_{unsat} = \rho_{unsat} Vp_{unsat}^2 - \frac{4}{3}\mu.$$

Now, using Gassman's equation[30] we can compute the bulk moduli of the material given totally empty pores (filled with neither gas nor liquid)

$$K_{dry} = \frac{K_{unsat}\left(\frac{K_{solid}}{K_{air}}\varphi + 1 + \varphi\right) - K_{solid}}{\frac{K_{solid}}{K_{air}} + \frac{K_{unsat}}{K_{solid}} - 1 - \varphi}.$$

At this point, we wish to account for the effect of water saturation on the physical properties of the material. We can compute the effect of saturation on the bulk modulus by using another formulation of Gassman's equation[30]

$$K_{sat} = K_{dry} + \frac{\left(1 - \frac{K_{dry}}{K_{solid}}\right)^2}{\frac{\varphi}{K_{fluid}} + \frac{1-\varphi}{K_{solid}} + \frac{K_{dry}}{K_{solid}^2}}$$

where the *sat* subscript indicates some amount of water saturation ($0 < s_w < 1$). In the above equation, the bulk modulus of the fluid can be computed using the Voigt model

$$K_{fluid} = K_w s_w + K_{air}(1 - s_w)$$

where K_w is the bulk modulus of water. In a similar manner, we can compute the density of the fluid phase using a linear average of the densities of water and air

$$\rho_{fluid} = \rho_w s_w + \rho_{air}(1 - s_w)$$

and then use this figure to compute the density of the partially saturated material

$$\rho_{sat} = \rho_{solid}(1 - \varphi) + \rho_{fluid}\varphi.$$

Finally, using the saturation-adjusted bulk modulus and density, we can compute the p-wave and s-wave velocity of the partially saturated material with

$$Vp_{sat} = \sqrt{\frac{K_{sat} + \frac{4}{3}\mu}{\rho_{sat}}}$$

and

$$Vs_{sat} = \sqrt{\frac{\mu}{\rho_{sat}}}.$$

In the resulting rock physics model, both Vp and Vs will decrease as porosity increases, but Vp will increase with saturation while Vs will slightly decrease. This behavior reflects previously published modelled and observed critical zone rock physics relationships [31–34].

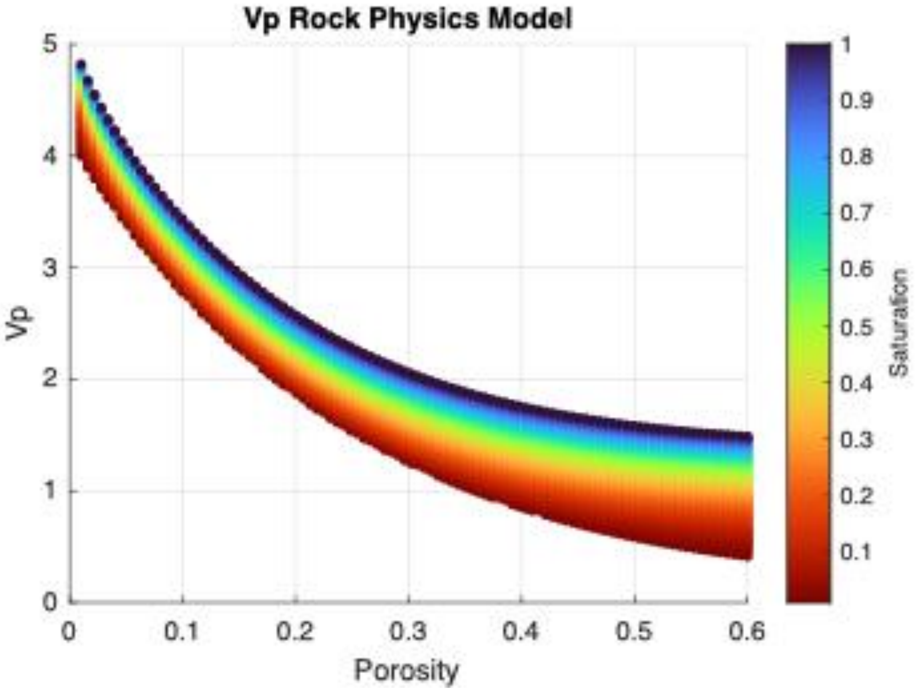

Figure S12: Behavior of the Vp rock physics model over the assumed range of hydrological parameters.

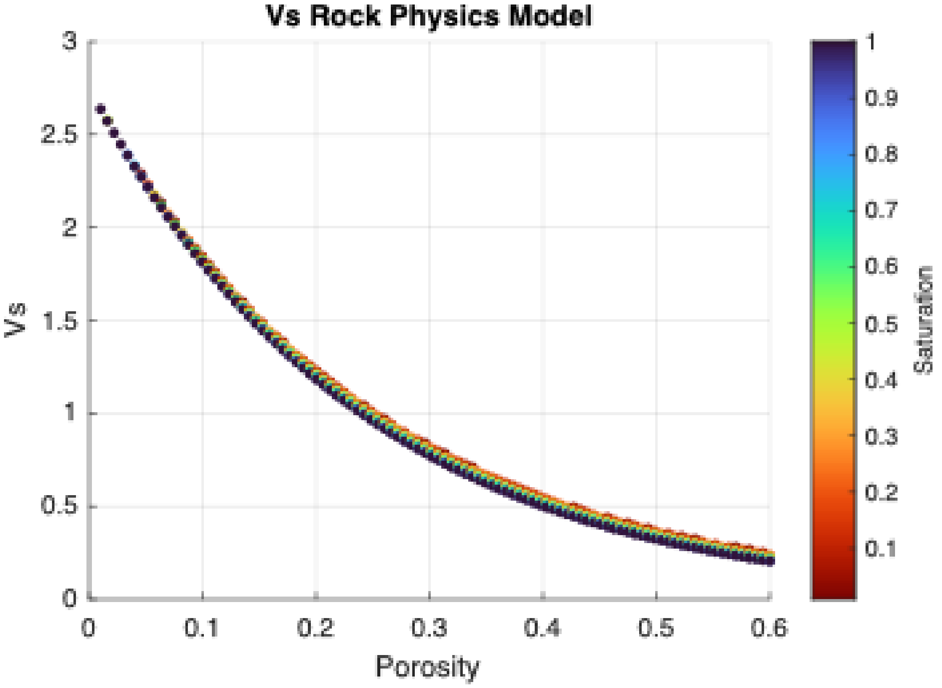

Figure S13: Behavior of the Vs rock physics model over the assumed range of hydrological parameters.

Geostatistical Rock Physics Inversion

With a rock physics model in hand, we can estimate the most probable saturation and porosity fields given the FWI-derived seismic velocities. The first step is to discretize the FWI velocity models (both Vp and Vs) in the form of a large vector

$$\boldsymbol{v} = \begin{bmatrix} \boldsymbol{Vp} \\ \boldsymbol{Vs} \end{bmatrix}.$$

Likewise, we can use the same discretization scheme to represent corresponding saturation and porosity models

$$\boldsymbol{w} = \begin{bmatrix} \boldsymbol{\varphi} \\ \boldsymbol{s_w} \end{bmatrix}.$$

Now, consider the function $g(\boldsymbol{w}) = \boldsymbol{v} : (\boldsymbol{\varphi}, \boldsymbol{s_w}) \rightarrow (\boldsymbol{Vp}, \boldsymbol{Vs})$ which implements our rock physics model. With the discretization scheme, we can utilize the ensemble smoother algorithm to estimate an optimal $\boldsymbol{w}$ from $\boldsymbol{v}$ [35]. The first step is to generate an ensemble of porosity and saturation models, denoted $\{\boldsymbol{w}_j\}_{j=1}^{N_e}$ where N_e is the number of model realizations in the ensemble (we use 5,000). We generate the initial ensemble of porosity and saturation models using the probability field simulation method [36] and impose horizontal and lateral smoothing throughout to enforce spatial regularization. We also mandate that in the initial ensemble, porosity monotonically decreases while saturation monotonically increases with depth, which reflects the fact that less weathered and more saturated materials tend to occur deeper in the critical zone. Once the initial ensemble is created, we can update each ensemble member using the following update formula[35]

$$\boldsymbol{w}_j^{i+1} = \boldsymbol{w}_j^i + \boldsymbol{\Gamma}_{w,v}^i \left(\boldsymbol{\Gamma}_{v,v}^i + \alpha \boldsymbol{\Gamma}_\sigma \right)^{-1} \left(\boldsymbol{v}^* - \boldsymbol{v}_j^i \right)$$

where i is the iteration number, $0 < \alpha < 1$ is an inflation factor, $\boldsymbol{v}_j^i = g(\boldsymbol{w}_j^i)$ is the predicted (from the rock physics model) velocity model from the jth member of the current porosity and saturation ensemble, $\boldsymbol{v}^* \sim \mathcal{N}(\boldsymbol{v}, \boldsymbol{\Gamma}_\sigma)$ is a perturbed velocity model vector with $\boldsymbol{\Gamma}_\sigma$ being the assumed covariance of the velocity model, $\boldsymbol{\Gamma}_{w,v}^i$ is the cross covariance matrix of the ensembles $\{\boldsymbol{w}_j\}_{j=1}^{N_e}$ and $\{\boldsymbol{v}_j\}_{j=1}^{N_e}$, and $\boldsymbol{\Gamma}_{v,v}^i$ is the auto covariance matrix of the ensemble of velocity models ($\{\boldsymbol{v}_j\}_{j=1}^{N_e}$).

References

1. Fichtner, A., Kennett, B. L. N., Igel, H. & Bunge, H.-P. Full seismic waveform tomography for upper-mantle structure in the Australasian region using adjoint methods. *Geophysical Journal International* **179**, 1703–1725 (2009).
2. Yong, P., Brossier, R., Métivier, L. & Virieux, J. Localized adaptive waveform inversion: theory and numerical verification. *Geophysical Journal International* **233**, 1055–1080 (2023).
3. Sinha, S., Routh, P. S., Anno, P. D. & Castagna, J. P. Spectral decomposition of seismic data with continuous-wavelet transform. *GEOPHYSICS* **70**, P19–P25 (2005).
4. Allstadt, K. E. *et al.* Seismic and acoustic signatures of surficial mass movements at volcanoes. *Journal of Volcanology and Geothermal Research* **364**, 76–106 (2018).
5. Torrence, C. & Compo, G. P. A Practical Guide to Wavelet Analysis. *Bull. Amer. Meteor. Soc.* **79**, 61–78 (1998).
6. Fichtner, A. *Full Seismic Waveform Modelling and Inversion*. (Springer Berlin Heidelberg, Berlin, Heidelberg, 2011). doi:10.1007/978-3-642-15807-0.
7. Métivier, L., Brossier, R., Kpadonou, F., Messud, J. & Pladys, A. A review of the use of optimal transport distances for high resolution seismic imaging based on the full waveform. Preprint at https://doi.org/10.48550/arXiv.2204.08514 (2022).
8. Métivier, L., Brossier, R., Mérigot, Q., Oudet, E. & Virieux, J. Measuring the misfit between seismograms using an optimal transport distance: application to full waveform inversion. *Geophys. J. Int.* **205**, 345–377 (2016).

9. Bozdağ, E., Trampert, J. & Tromp, J. Misfit functions for full waveform inversion based on instantaneous phase and envelope measurements: Misfit functions for full waveform inversion. *Geophysical Journal International* **185**, 845–870 (2011).
10. Virieux, J. & Operto, S. An overview of full-waveform inversion in exploration geophysics. *GEOPHYSICS* **74**, WCC1–WCC26 (2009).
11. Komatitsch, D. & Tromp, J. Introduction to the spectral element method for three-dimensional seismic wave propagation. *Geophys. J. Int.* **139**, 806–822 (1999).
12. Tromp, J., Tape, C. & Liu, Q. Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels: Seismic tomography, adjoint methods, time reversal and banana-doughnut kernels. *Geophysical Journal International* **160**, 195–216 (2004).
13. Liu, Q. & Tromp, J. Finite-Frequency Kernels Based on Adjoint Methods. *Bulletin of the Seismological Society of America* **96**, 2383–2397 (2006).
14. Plessix, R.-E. A review of the adjoint-state method for computing the gradient of a functional with geophysical applications. *Geophysical Journal International* **167**, 495–503 (2006).
15. Modrak, R. T., Borisov, D., Lefebvre, M. & Tromp, J. SeisFlows—Flexible waveform inversion software. *Computers & Geosciences* **115**, 88–95 (2018).
16. Modrak, R. & Tromp, J. Seismic waveform inversion best practices: regional, global and exploration test cases. *Geophys. J. Int.* **206**, 1864–1889 (2016).
17. Liu, X., Zhu, T. & Hayes, J. Critical Zone Structure by Elastic Full Waveform Inversion of Seismic Refractions in a Sandstone Catchment, Central Pennsylvania, USA. *JGR Solid Earth* **127**, e2021JB023321 (2022).

18. Eppinger, B. J., Holbrook, W. S., Liu, Z., Flinchum, B. A. & Tromp, J. 2D Near-Surface Full-Waveform Tomography Reveals Bedrock Controls on Critical Zone Architecture. *Earth and Space Science* **11**, e2023EA003248 (2024).
19. Bunks, C., Saleck, F., Zalesky, S. & Chavent, D. multiscale seismic waveform inversion. *GEOPHYSICS* (1995).
20. Köhn, D. *et al.* Comparison of time-domain SH waveform inversion strategies based on sequential low and bandpass filtered data for improved resolution in near-surface prospecting. *Journal of Applied Geophysics* **160**, 69–83 (2019).
21. Luo, Y. & Schuster, G. T. Wave-equation traveltime inversion. (1991).
22. Wang, Y., Khorrami, M., Tran, K. T. & Horhota, D. Application of ambient noise tomography for deep void detection. *Journal of Applied Geophysics* **209**, 104922 (2023).
23. Smith, J. A. *et al.* Tunnel detection at Yuma Proving Ground, Arizona, USA — Part 2: 3D full-waveform inversion experiments. *GEOPHYSICS* **84**, B107–B120 (2019).
24. Groos, L., Schäfer, M., Forbriger, T. & Bohlen, T. Application of a complete workflow for 2D elastic full-waveform inversion to recorded shallow-seismic Rayleigh waves. *GEOPHYSICS* **82**, R109–R117 (2017).
25. Aktharuzzaman, M., Anwar, S., Borisov, D. & He, J. Experimental full waveform inversion for elastic material characterization with accurate transducer modeling. *Mechanical Systems and Signal Processing* **213**, 111320 (2024).
26. Befus, K. M., Sheehan, A. F., Leopold, M., Anderson, S. P. & Anderson, R. S. Seismic Constraints on Critical Zone Architecture, Boulder Creek Watershed, Front Range, Colorado. *Vadose Zone Journal* **10**, 915–927 (2011).

27. Holbrook, W. S. *et al.* Geophysical constraints on deep weathering and water storage potential in the Southern Sierra Critical Zone Observatory. *Earth Surf Processes Landf* **39**, 366–380 (2014).
28. Dvorkin, J. & Nur, A. Elasticity of high-porosity sandstones: Theory for two North Sea data sets. *GEOPHYSICS* **61**, 1363–1370 (1996).
29. Mixture Theories for Rock Properties. 205–228 (1995) doi:10.1029/rf003p0205.
30. Mavko, G. *The Rock Physics Handbook: Tools for Seismic Analysis of Porous Media*. (Cambridge University Press, 2009).
31. Denolle, M. A. *et al.* Ambient field seismology in critical zone hydrological sciences. *Comptes Rendus. Géoscience* **357**, 425–451 (2025).
32. Shen, Z. *et al.* Fiber-optic seismic sensing of vadose zone soil moisture dynamics. *Nat Commun* **15**, (2024).
33. Solazzi, S. G., Bodet, L., Holliger, K. & Jougnot, D. Surface-Wave Dispersion in Partially Saturated Soils: The Role of Capillary Forces. *JGR Solid Earth* **126**, (2021).
34. Garambois, S. *et al.* Analysis of ballistic waves in seismic noise monitoring of water table variations in a water field site: added value from numerical modelling to data understanding. *Geophysical Journal International* **219**, 1636–1647 (2019).
35. Emerick, A. A. & Reynolds, A. C. Ensemble smoother with multiple data assimilation. *Computers & Geosciences* **55**, 3–15 (2013).
36. Srivastava, R. Reservoir characterization with probability field simulation. (1992).