Reliability Engineering and System Safety 260 (2025) 110921

Contents lists available at ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Effects of inhomogeneity and statistical and material anisotropy on THM simulations

Aqeel Afzal Chaudhry [a,*], Chao Zhang [b], Oliver G. Ernst [c], Thomas Nagel [a,d,e]

[a] Technische Universität Bergakademie Freiberg, Gustav-Zeuner-Str. 1, Freiberg 09599, Germany
[b] Technical University of Denmark, Lyngby, Denmark
[c] Chemnitz University of Technology, Chemnitz, Germany
[d] Freiberg Center for Water Research – ZeWaF, Freiberg, Germany
[e] Helmholtz Centre for Environmental Research GmbH – UFZ, Leipzig, Germany

ARTICLE INFO

Keywords:
Coupled thermo-hydro-mechanical processes
Non-isothermal Richards equation
Random fields
Anisotropy
Inhomogeneity
OpenGeoSys

ABSTRACT

When modeling the material properties of host rocks for thermo-hydro-mechanical simulations in barrier integrity investigations for deep geological disposal of radioactive waste, numerous modeling aspects must be considered. If complete information were available, the material properties would be functions of space, with inhomogeneity and anisotropy expressed by spatially varying and tensor-valued coefficients. In practice, uncertainty is present in particular related to spatial variability of physical properties. This variability can be modeled by random fields, whose realizations are functions of space. A common choice is a Gaussian random field, determined by its mean and two-point covariance function. Anisotropy can occur both in the statistical covariance structure, resulting in different correlation lengths along principal axes, and in the physical properties themselves, leading to tensor-valued random fields. In this study, we focus on both cases, considering dominant material properties such as thermal conductivity, intrinsic permeability, and Young's modulus, and present numerical simulations illustrating the effects of inhomogeneity, randomness, and anisotropy. Since spatial variability is a key feature in the analysis of in-situ data, this study quantifies the individual contribution of each of the listed effects in a well-controlled synthetic case and discusses them in the context of scale.

1. Introduction

Deep geological disposal of radioactive waste requires a comprehensive understanding of the material properties of host rocks to ensure barrier integrity and long-term safety [1–10]. However, modeling these properties poses significant challenges due to the complex nature of geological formations and the limited availability of comprehensive data [11]. In practice, complete information about material properties as functions of space is unattainable for several reasons including depth and the limited accessibility by drilling [12]. Furthermore, in certain applications not only the physical inaccessibility but also the extreme conditions such as high pressures, temperatures, and corrosive environment result in limited direct sampling [13,14]. Consequently, the cost and time required for data acquisition is usually very high. In geological disposal, extensive invasive exploration and the maintenance of an intact barrier are to a certain degree conflicting objectives. In any case, deep geological formations are complex and inhomogeneous, exhibiting variations in rock types, structure, and properties [15–19]. Inferring their structure from geotechnical and geophysical exploration remains an interpretative task and predicting their behavior over long periods is challenging due to inherent uncertainties [20–24]. Factors such as the presence of fractures, fault zones, or natural pathways for fluid migration can affect the safety and stability when it comes to barrier integrity [25].

To account for part of the uncertainty, a common approach is to model the rock medium as piecewise homogeneous, treating material properties within each homogeneous sub-region as random variables [26]. This allows for the representation of spatial variability only from one layer to another. Incorporating randomness with a more general structure can be achieved through the use of random fields, which are functions of space that are not necessarily constant [27,28].

Several studies have explored the application of random fields to represent the spatial variability of subsurface properties. Follin et al. [29] utilized random fields to model discrete fracture networks in

* Corresponding author.
E-mail addresses: aqeel.chaudhry@ifgt.tu-freiberg.de (A.A. Chaudhry), chaz@dtu.dk (C. Zhang), oernst@math.tu-chemnitz.de (O.G. Ernst), thomas.nagel@ifgt.tu-freiberg.de (T. Nagel).

https://doi.org/10.1016/j.ress.2025.110921
Received 7 October 2024; Received in revised form 30 January 2025; Accepted 14 February 2025
Available online 24 February 2025
0951-8320/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

crystalline rocks, demonstrating their use in capturing heterogeneous permeability distributions relevant to nuclear waste repositories. Joyce et al. [30] extended these applications to multi-scale groundwater flow models, highlighting the impact of spatial variability on safety assessments at sites like Forsmark, Sweden. Ernst et al. [31] studied the probability distribution of travel time of radioactive contaminants, using data from the Waste Isolation Pilot Plant (WIPP) in Carlsbad, NM, USA. These studies underline the critical role of random fields in representing the stochastic nature of geological properties, essential for accurate predictions of subsurface behavior.

Stochastic simulations using random fields have also been used to model fault networks and geological structures, contributing significantly to the understanding of fluid migration pathways in fractured rock systems. Cherpeau et al. [32] applied these methods in fault network modeling, while Wang et al. [33] developed a segmentation approach for three-dimensional geological modeling using hidden Markov random fields. Cvetkovic and Frampton [34] used Gaussian random fields to characterize the internal structure of faults and concluded that more complex representations of inhomogeneity in fracture networks might be required. These advancements underscore the importance of inhomogeneity in enhancing the predictive capabilities of models used in radioactive waste management.

Anisotropy, which refers to the direction-dependent behavior of material properties, is another crucial aspect to consider. It can occur both in the statistical covariance structure and in the local thermal, hydraulic, or mechanical properties of the rock mass itself. Anisotropy in the statistical covariance structure results in different correlation lengths along principal axes, while anisotropy in material properties leads to tensor-valued random fields with anisotropic tensors as realizations. These tensor-valued random fields can exhibit both spatial and statistical anisotropy.

Norberg et al. [35] demonstrated the significance of modeling discrete geological structures as anisotropic random fields, emphasizing their role in capturing spatial dependencies. The work of Hadgu et al. [36] on discrete fracture network models for crystalline rocks further illustrated the impact of anisotropic behaviors on flow and transport predictions. Kim and Inoue [37] analyzed anisotropic permeability through rock joints, showing how direction-dependent correlation lengths influence flow predictions. Similarly, Tsang et al. [38] explored hydrologic issues in low-seismic activity regions, using direction-dependent simulations to assess the performance of geological repositories under varying conditions. Shao et al. [39] provided insights into in situ measurements of anisotropic permeability in clay, further illustrating the variability of geological formations. These studies highlight the need for incorporating anisotropic models to accurately reflect the direction-dependent behaviors of subsurface properties.

Thermo-hydro-mechanical (THM) simulations have become a widely used approach for understanding the combined effects of thermal, hydraulic, and mechanical processes in geological media [40,41]. Studying random fields through THM simulations will provide a comprehensive understanding of their interactions and effects on related geological processes. Salager et al. [42] studied the mechanical anisotropy of Opalinus clay using THM models, illustrating the role of anisotropy in predicting deformation behavior. Barton and Quadros [43] further explored methodologies for modeling anisotropy in geological media, emphasizing the importance of directional dependencies in rock mechanics.

Geomaterials display structure and texture at various scales and how they are represented in models is a question of scale. Consider the illustration in Fig. 1 where the spatial variability of a rock property of interest is represented as a random field, usually statistically anisotropic. At each point, the values of local physical properties are themselves the result of a complex microstructure (b). As it is usually both too expensive and not necessary to explicitly resolve these microstructures with the geometry, arrangement and properties of each

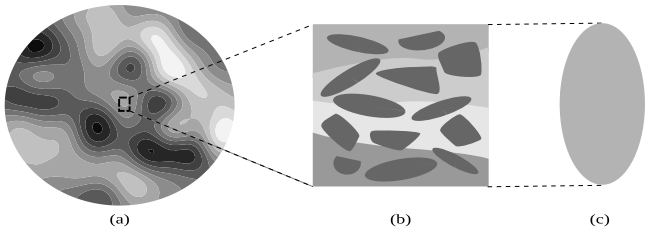

Fig. 1. Basic concept illustrating anisotropy in a random field. (a) shows a (realization of (a) random field representing spatial variability in geological properties at the modeled scale. (b) shows a zoomed-in view of a representative elementary volume, revealing layered media with patches of varying properties, illustrating spatial variability at a smaller scale. After averaging over the REV, physical properties may be represented as anisotropic point-wise quantities (c) in the model (a).

of its constituents, local averaging procedures on representative elementary volumes (REVs) are applied that result in a single anisotropic tensorial value – indicated by the ellipse in (c) – for the REV (b) representing the material points in the macroscale model (a).

In this study, we focus on both types of anisotropy mentioned above by considering at least the dominant material property of each physical process involved in THM simulations: thermal conductivity ($\boldsymbol{\lambda}$) for the thermal part, intrinsic permeability ($\mathbf{k}$) for the hydraulic part, and the principal Young's moduli (E_i) for the mechanical part. These properties play key roles in heat transfer, fluid flow, and mechanical behavior within the host rocks. It is worth mentioning that the choice of the parameters is also supported by previous studies where local and global sensitivity analyses were performed [44] to observe the most significant material parameters affecting the thermal, hydraulic and mechanical behavior of the host rock and a Design of Experiments based history-matching workflow was adapted for uncertainty quantification [45]. The parameters chosen in the current work consistently ranked among those most significant for coupled THM processes.

Thus, the objectives of the paper are to demonstrate how random fields can be introduced in THM simulations on a sound mathematical basis; to show how structure (material inhomogeneity) at different scales can be represented in such simulations as either anisotropy in pointwise properties resulting from an upscaling process that remains, here, implicit or as anisotropy in the random field statistics; and finally, to discuss implications of considering random fields in comparison to mere parameter variations in a spatially homogeneous setting on the results of THM analyses.

This study is organized as follows: Section 2 introduces the methods with Section 2.1 giving the brief summary of the governing equations and Section 2.2 giving a detailed explanation of the approximation of Gaussian random fields. Section 3 outlines the model setup, providing the framework for our simulations. In Section 4, we present a series of study cases with Section 4.1 illustrating and discussing the study cases with homogeneous, isotropic, and anisotropic conditions, and Section 4.2 showing the inhomogeneous study cases under statistically isotropic (Section 4.2.1) and anisotropic (Section 4.2.2) conditions. The paper concludes with a discussion and synthesis of our findings in Section 5, emphasizing the practical implications for the selection of safe disposal sites for radioactive waste.

2. Methods

2.1. Governing equations

We employ the Thermo-Richards-Mechanics (TRM) process [46,47] implemented in OpenGeoSys-6 (OGS-6) [48]. The process is representative of a non-isothermal porous medium with a solid phase and a liquid phase, whereas the gas phase is considered isobaric following Richards' assumption [49]. It should be noted here that this assumption simply means that the gas phase is allowed to move freely, and thus there

is no change in gas pressure [46,47]. Thus, the model is governed by three balance equations, i.e., energy balance, mass balance of the liquid phase and linear momentum balance of the mixture which are thus formulated based on three independent state variables i.e., temperature T, liquid pressure p_{LR} and solid displacement $\mathbf{u}_{\mathrm{S}}$, respectively. To avoid repetition, we will not mention all the terms appearing in the equations in this section. Only the terms we consider significant enough for this brief summary will be mentioned in this section, while the definitions of all variables and mathematical symbols are listed in Table A.4. For a detailed description of the TRM model summarized here as well as its comparison to a non-isothermal two-phase two-component flow with mechanics model ($\mathrm{TH^2M}$) [50], the reader is referred to [47]. The model used in this study has been extensively verified and validated by comparison to experimental data, see [47,51,52] and references therein. The energy balance equation is written as

$$\begin{aligned}&(\rho c_p)_{\mathrm{eff}}\frac{\mathrm{d}T}{\mathrm{d}t}+L_0\frac{\mathrm{d}\theta_{\mathrm{vap}}}{\mathrm{d}t}-\operatorname{div}\left(\boldsymbol{\lambda}_{\mathrm{eff}}\operatorname{grad}T\right)\\&\quad+\operatorname{div}\left(\frac{L_0\boldsymbol{J}_{\mathrm{G}}^{\mathrm{W}}}{\rho_{\mathrm{GR}}^{\mathrm{W}}}\right)+\operatorname{grad}T\cdot\left(c_{p\mathrm{L}}\boldsymbol{A}_{\mathrm{L}}+c_{p,\mathrm{vap}}\boldsymbol{J}_{\mathrm{G}}^{\mathrm{W}}\right)=Q_T,\end{aligned}\tag{1}$$

where $(\rho c_p)_{\mathrm{eff}}$ is the effective volumetric heat capacity of the medium and L_0 is the volumetric latent heat of vaporization.

The mass balance equation for the liquid phase is given as

$$\begin{aligned}&\rho_{\mathrm{LR}}S_{\mathrm{L}}(\alpha_{\mathrm{B}}-\phi)\beta_{p,\mathrm{SR}}\frac{\mathrm{d}p_{\mathrm{LR}}}{\mathrm{d}t}-\rho_{\mathrm{LR}}S_{\mathrm{L}}(\alpha_{\mathrm{B}}-\phi)\operatorname{tr}(\boldsymbol{\alpha}_{T,\mathrm{SR}})\frac{\mathrm{d}T}{\mathrm{d}t}\\&\quad+\phi\left((1-S_{\mathrm{L}})\frac{\mathrm{d}\rho_{\mathrm{GR}}^{\mathrm{W}}}{\mathrm{d}t}+S_{\mathrm{L}}\frac{\mathrm{d}\rho_{\mathrm{LR}}}{\mathrm{d}t}\right)\\&\quad+\left(\rho_{\mathrm{LR}}-\rho_{\mathrm{GR}}^{\mathrm{W}}\right)\left[\phi+p_{\mathrm{LR}}S_{\mathrm{L}}(\alpha_{\mathrm{B}}-\phi)\right]\frac{\mathrm{d}S_{\mathrm{L}}}{\mathrm{d}t}\\&\quad+\rho_{\mathrm{LR}}S_{\mathrm{L}}\alpha_{\mathrm{B}}\operatorname{div}\left(\frac{\mathrm{d}\boldsymbol{u}_{\mathrm{S}}}{\mathrm{d}t}\right)+\operatorname{div}\left(\boldsymbol{A}_{\mathrm{L}}^{\mathrm{W}}+\boldsymbol{J}_{\mathrm{G}}^{\mathrm{W}}\right)=Q_H,\end{aligned}\tag{2}$$

where α_{B}, $\beta_{p,\mathrm{SR}}$ and $\boldsymbol{\alpha}_{T,\mathrm{SR}}$ represent the Biot–Willis coefficient, grain compressibility of the solid and linear expansion coefficient matrix of the solid, respectively, whereas Q_H on the right hand side represents the fluid source (or sink).

The linear momentum balance of the overall mixture is given as

$$\operatorname{div}\left(\boldsymbol{\sigma}^{\mathrm{eff}}-\alpha_{\mathrm{B}}\chi(S_{\mathrm{L}})p_{\mathrm{LR}}\mathbf{I}\right)+\rho\mathbf{g}=\mathbf{0},\tag{3}$$

with

$$\dot{\boldsymbol{\sigma}}^{\mathrm{eff}}=\boldsymbol{C}:(\dot{\boldsymbol{\epsilon}}-\dot{\boldsymbol{\epsilon}}_{\mathrm{pl}}-\dot{\boldsymbol{\epsilon}}_{\mathrm{th}}-\dot{\boldsymbol{\epsilon}}_{\mathrm{sw}}),\tag{4}$$

where $\boldsymbol{C}$ is the fourth order elastic tensor while $\boldsymbol{\epsilon}$, $\boldsymbol{\epsilon}_{\mathrm{pl}}$, $\boldsymbol{\epsilon}_{\mathrm{th}}$ and $\boldsymbol{\epsilon}_{\mathrm{sw}}$ represent the total, plastic, thermal and swelling strains, respectively. The present study is restricted to linear elastic solid behavior.

The fluid density is assumed to be temperature and pressure dependent while fluid viscosity is assumed to be temperature dependent [53], and are given as

$$\rho_{\mathrm{LR}}=\frac{\rho_{\mathrm{LR}}^0}{e^{\alpha_{\mathrm{L}}(T-T_0)-c_{\mathrm{L}}(p_{\mathrm{LR}}-p_{\mathrm{LR}}^0)}},\tag{5}$$

$$\mu_{\mathrm{L}}=\left(4.2844\cdot10^{-5}+(0.157(\frac{T}{\mathrm{K}}-208.157)^2-91.296)^{-1}\right)\mathrm{Pa\,s}.\tag{6}$$

2.2. Random fields and their approximation

When uncertain quantities modeled as random variables are allowed to vary in space, the resulting mathematical object is a *stochastic process* indexed by the spatial coordinate, i.e., a family of random variables, one associated with each spatial point. A stochastic process indexed by a spatial coordinate is commonly referred to as *random field*. In the following, we model spatially varying uncertain physical quantities as random fields on a bounded domain $D\subset\mathbb{R}^2$.

2.2.1. Random fields

We denote by $Z(\mathbf{x},\omega)$ a random field indexed by $\mathbf{x}\in D$ on a probability space $(\Omega,\mathcal{A},\mathbf{P})$. For each fixed $\omega\in\Omega$, $Z(\mathbf{x},\omega)$ denotes a realization or sample of the random process, i.e., a function of $\mathbf{x}$ defined on D. The probability law of random fields can be chosen to model different types of uncertainty and to account for known physical properties of the uncertain quantities being modeled. In this study, we consider *Gaussian random fields*. These are characterized by the fact that the joint distribution of any collection $\{Z(\mathbf{x}_1,\omega),\dots,Z(\mathbf{x}_n,\omega)\}$ of Z at a finite number of locations is multivariate Gaussian. *Lognormal* random fields are obtained as the composition $\exp(Z(\mathbf{x},\omega))$ of a Gaussian random field Z with the exponential function, resulting in strictly positive realizations. A Gaussian random field Z is completely characterized by its pointwise mean and covariance functions, which we assume to exist and are defined, respectively, as

$$\begin{aligned}&\overline{Z}(\mathbf{x}):=\mathbf{E}\left[Z(\mathbf{x},\cdot)\right], && \mathbf{x}\in D,\\&\text{and}\quad c(\mathbf{x},\mathbf{y}):=\mathbf{Cov}(Z(\mathbf{x}),Z(\mathbf{y}))=\mathbf{E}\left[(Z(\mathbf{x})-\overline{Z}(\mathbf{x}))(Z(\mathbf{y})-\overline{Z}(\mathbf{y}))\right], && \mathbf{x},\mathbf{y}\in D,\end{aligned}\tag{7}$$

where $\mathbf{E}[\cdot]$ denotes mathematical expectation with respect to $\mathbf{P}$. The covariance function quantifies the strength of correlation of a random field at two distinct points $\mathbf{x},\mathbf{y}\in D$, whereas for $\mathbf{x}=\mathbf{y}$ it gives the variance $c(\mathbf{x},\mathbf{x})=\mathbf{Var}\,Z(\mathbf{x},\cdot)$. Covariance functions are said to be (wide-sense) *stationary* if $c(\mathbf{x},\mathbf{y})=c(\mathbf{x}+\mathbf{d},\mathbf{y}+\mathbf{d})$ for any shift vector $\mathbf{d}$ and *isotropic* if $c(\mathbf{x},\mathbf{y})$ is a function of only the separation distance $|\mathbf{x}-\mathbf{y}|$, where $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^2$. For an overview of commonly used covariance models see [54,55].

A particularly flexible family of isotropic covariance functions often used in the geosciences is that of the *Matérn* kernels given by

$$c(\mathbf{x},\mathbf{y})=\frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)}\left(\frac{2\sqrt{\nu}|\mathbf{x}-\mathbf{y}|}{\ell}\right)^{\nu}K_{\nu}\left(\frac{2\sqrt{\nu}|\mathbf{x}-\mathbf{y}|}{\ell}\right),\tag{8}$$

where $\sigma^2>0$ is the marginal variance, Γ is the Gamma function, ℓ is the characteristic length scale, $\nu>0$ is the smoothness parameter, and K_ν is the modified Bessel function of the second kind of order ν. The smoothness parameter ν controls the smoothness of the kernel function at the origin. Moreover, realizations of a Gaussian random field with Matérn kernel of smoothness parameter ν are $\nu-1$ times differentiable (cf. [56, Section 5.5]). Fig. 2 displays some commonly used Matérn kernel functions with different values of the smoothness parameter ν. In this study, we fix $\nu=1/2$, which corresponds to rough random fields as typically encountered with geophysical quantities.

We note that the choice of a statistical model, reflecting our assumptions about the uncertainty in the spatially varying properties, is inherently a modeling decision. Our choice of Gaussian random fields is motivated by their established use in subsurface geophysical modeling, where their flexibility allows key field characteristics, e.g., smoothness, periodicity, and stationarity, to be configured through appropriate covariance functions. While Gaussian or lognormal distributions often serve as suitable initial models, site-specific characterization requires critical evaluation of the chosen statistical model. This includes careful consideration of all available information and its limitations to ensure the model adequately reflects the nature of the remaining uncertainty. In this work, we assume stationarity and smoothness to balance model complexity, though we acknowledge that these assumptions may not always hold. The primary focus of this work is to demonstrate that explicitly incorporating spatial uncertainty into THM simulations provides a valuable tool for realistically assessing its impact on model outcomes and conclusions regarding structural integrity, rather than comparing alternative statistical models. In addition, we have made efforts to tailor the distributions, including considerations of positivity and scaling, to align with the nominal values and observed variations

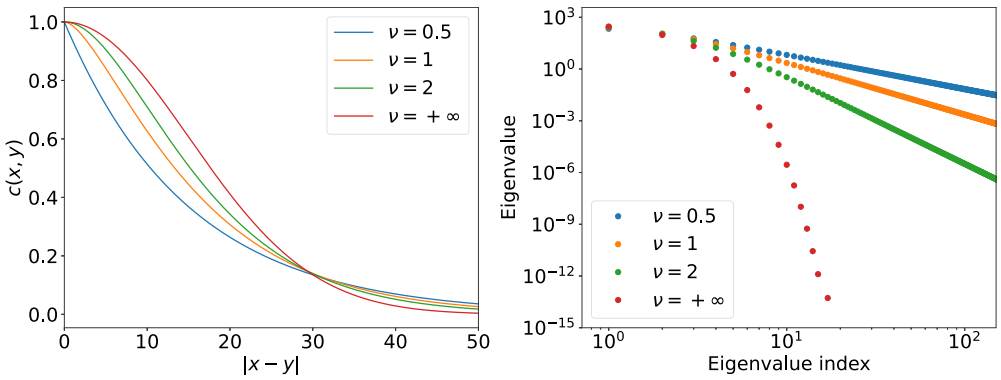

Fig. 2. Matérn family of covariance functions with different smoothness parameter ν are plotted on the left panel, while eigenvalues of their associated covariance matrices for a 1D problem are also displayed with respect to their indices on the right panel. The correlation length ℓ is fixed at 15, and the scaling factor σ is set to 1 across all covariance functions. Note that in the case of $\nu = +\infty$, the eigenvalues decay rapidly towards zero as the index increases, and here only the largest 21 eigenvalues are plotted, as the accuracy of the remaining near-zero eigenvalues is restricted by machine precision.

of the physical quantities. However, non-Gaussian models may become more appropriate when non-Gaussian characteristics are expected [57, 58].

2.2.2. Approximation of random fields

The conceptually simplest approach for approximating a Gaussian random field Z defined on D with given mean and covariance functions (7) is to fix a grid of points $\{\mathbf{x}_1, \dots, \mathbf{x}_n\} \subset D$ and approximate the multivariate normal random vector $\mathbf{Z} = \mathbf{Z}(\omega) = [Z(\mathbf{x}_1, \omega), \dots, Z(\mathbf{x}_n), (\omega)]^\top$ obtained by restricting the random field to the finite grid. The associated pointwise covariance matrix $\mathbf{C} = [c(\mathbf{x}_i, \mathbf{x}_j)]_{i,j=1}^n \in \mathbb{R}^{n\times n}$ is necessarily symmetric and positive definite, and therefore possesses a Cholesky factorization $\mathbf{C} = \mathbf{L}\mathbf{L}^\top$, which can be used to generate samples of $\mathbf{Z}$ as $\mathbf{z} = \boldsymbol{\mu} + \mathbf{L}\boldsymbol{\xi}$ where $\boldsymbol{\mu} = \mathbf{E}[\mathbf{Z}]$ and the n components of $\boldsymbol{\xi}$ are *i.i.d.* standard normal random values. Challenges arise both from the $O(n^3)$ complexity of Cholesky factorization as well as the ill-conditioning of the covariance matrix, and numerous approaches have been developed to address these (cf. [56, Chapter 7], [59–61]).

An alternative approach which performs an approximation in function space prior to discretization is based on the *Karhunen–Loève (KL) expansion*, a modal expansion for random fields with finite second moments similar to Fourier series [62], [63, §37]. It is given as

$$Z(\mathbf{x}, \omega) = \overline{Z}(\mathbf{x}) + \sum_{m=1}^{\infty} \sqrt{\lambda_m}\, f_m(\mathbf{x})\, \xi_m(\omega), \qquad \mathbf{x} \in D, \tag{9}$$

where $(\xi_m)_{m\in\mathbb{N}}$ is a sequence of uncorrelated real-valued random variables with zero mean and unit variance and (λ_m, f_m) are the eigenpairs of the second-kind Fredholm integral operator

$$\int_D c(\mathbf{x}, \mathbf{y}) f_m(\mathbf{y})\, \mathrm{d}\mathbf{y} = \lambda_m f_m(\mathbf{x}), \qquad m \in \mathbb{N}, \quad \mathbf{x} \in D, \tag{10}$$

which is the *covariance operator* $C : L^2(D) \to L^2(D)$ associated with the covariance function c of Z. Note that for a Gaussian random field Z the random coefficients ξ_m are also Gaussian, hence $\xi_m \sim \mathrm{N}(0, 1)$. In this scaling, the eigenfunctions have unit norm in $L^2(D)$. By the Hilbert–Schmidt theorem for compact self-adjoint linear operators (cf. [56, Section 1.3]), the eigenvalues are non-negative and, ordered decreasing, form a null sequence which is square summable. The KL expansion represents Z as a superposition of its covariance eigenmodes f_m, each of which is weighted by a random coefficient. Due to the decay of the eigenvalues, terms of higher index have a smaller influence on the value of Z. Moreover, a convergent approximation to Z can be obtained by truncating the KL expansion after a finite number M of terms, giving

$$Z(\mathbf{x}) \approx Z_M(\mathbf{x}) := \overline{Z}(\mathbf{x}) + \sum_{m=1}^{M} \sqrt{\lambda_m}\, f_m(\mathbf{x})\, \xi_m(\omega), \qquad \mathbf{x} \in D. \tag{11}$$

The approximation error is given by

$$\|Z - Z_M\|^2_{L^2(\Omega; L^2(D))} = \sigma^2 |D| - \sum_{m=1}^{M} \lambda_m, \tag{12}$$

where $|D|$ denotes the volume of D (a derivation is given in Appendix B.2.1), hence the L^2-error due to truncation can be ascertained given an approximation of the first M covariance eigenvalues. Truncation of the KL expansion after the M largest terms retains the dominant spatial patterns of a random field and yields a parametrization of Z in terms of M uncorrelated random variables. The truncation index can be chosen independently of the discretization of the physical model as long as all eigenfunctions are sufficiently resolved. The KL expansion has become an established tool for computational modeling of spatial variability for uncertainty quantification in engineering applications [64–67].

2.2.3. Computational realization of the Karhunen–Loève expansion

Closed-form expressions for the covariance eigenpairs (λ_m, f_m) are known only for special one-dimensional cases (cf. [68, Section 2.3], [69, Section 3.4.1]), and extending these analytical solutions to multi-dimensional settings is only possible by assuming a separable covariance structure, which leads to questionable statistical models in most engineering problems [55, Section 2.11]. A numerical approach to solving the covariance eigenproblem (10) requires first discretizing the covariance operator C and solving the resulting finite-dimensional eigenproblem. For the former, we employ a Galerkin approximation using piecewise constant functions on a triangulation of the domain D. This has the advantage of flexibility with regard to the domain shape, and can easily accommodate higher order basis functions. Conforming Galerkin discretizations in $L^2(D)$ require no continuity across element boundaries.

Assuming $D \subset \mathbb{R}^2$ has a polygonal boundary, a finite-dimensional subspace of $L^2(D)$ is obtained as the span of a set of piecewise constant basis functions $\{\phi_1, \dots, \phi_N\}$. Expanding the mth eigenfunction in this basis

$$f_m(\mathbf{x}) = \sum_{i=1}^{N} f_m^i \phi_i(\mathbf{x}), \tag{13}$$

with coefficients f_m^i; substituting the representation (13) into (10), successively multiplying by each test function ϕ_i and integrating over D, we arrive at the discrete generalized eigenvalue problem

$$\mathbf{C}\mathbf{f}_m = \lambda_m \mathbf{M}\mathbf{f}_m, \quad \mathbf{f}_m = [f_m^1, \dots, f_m^N]^\top, \tag{14}$$

where the Galerkin matrix $\mathbf{C} \in \mathbb{R}^{N\times N}$ is symmetric positive definite with entries

$$[\mathbf{C}]_{i,j} = \int_D \phi_j(\mathbf{x}) \int_D c(\mathbf{x}, \mathbf{y}) \phi_i(\mathbf{y})\, \mathrm{d}\mathbf{y}\, \mathrm{d}\mathbf{x}, \tag{15}$$

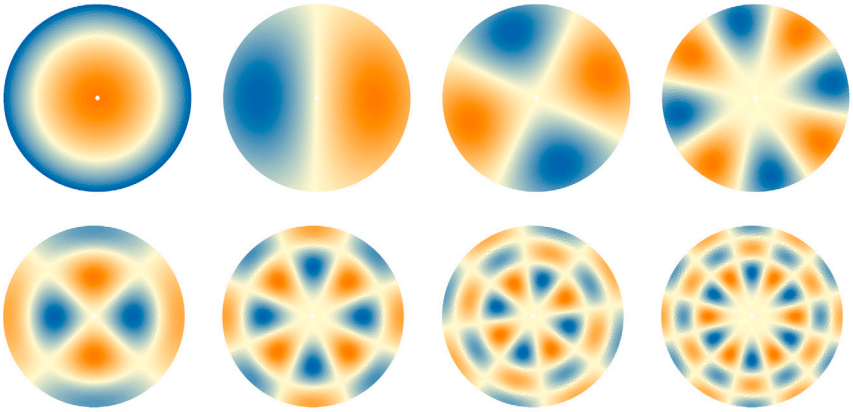

Fig. 3. Eigenfunctions 1, 2, 5, 11, 13, 25, 43, and 64 of the domain of interest, using a Matérn covariance function with smoothness parameter $\nu = 0.5$ and characteristic length $\ell = 15\,\mathrm{m}$. The outer diameter of the domain is set to 100 m; see Section 3 for more details. Orange hues indicate higher values and a colorbar is omitted as the relative magnitudes of the eigenfunctions are not of primary interest in this context.

and the Gram matrix $\mathbf{M} \in \mathbb{R}^{N\times N}$ is also symmetric positive definite with entries

$$[\mathbf{M}]_{i,j} = \int_D \phi_j(\mathbf{x})\phi_i(\mathbf{x})\,\mathrm{d}\mathbf{x}. \tag{16}$$

$\mathbf{M}$ is diagonal due to the disjoint supports of the piecewise constant basis functions.

Computational challenges arise from the fact that the Galerkin matrix $\mathbf{C}$ is dense due to the nonlocal nature of the covariance operator and the entries (15) consisting of double multivariate integrals can require expensive quadrature, in particular near the diagonal where the covariance function has lower regularity. We address issues of quadrature accuracy by applying techniques from boundary element methods [70]. Solving the dense generalized eigenvalue problem (14) with the symmetric QR algorithm would incur $O(N^3)$ computational complexity. Besides being intractable for all but the smallest problems, computing the full spectral decomposition of $\mathbf{C}$ would be wasteful insofar as the KL requires only a sufficient number of dominant eigenpairs. We therefore employ a Krylov subspace iteration based the *thick-restart Lanczos method* of Simon and Wu, which allows specifying a priori the number of desired eigenpairs and iteratively refines the associated target invariant subspace [71,72]. In addition, the algorithm requires only matrix–vector products and, in contrast to QR iteration, no dense linear algebra modifications of the matrices $\mathbf{C}$ and $\mathbf{M}$. The final crucial ingredient of our computational approach to KL expansion is a data-sparse representation of the dense Galerkin matrix $\mathbf{C}$ using the hierarchical matrix ($\mathcal{H}$-matrix) format [73,74], which reduces the complexity of constructing the Galerkin matrix as well as matrix–vector multiplication to $O(N \log N)$, and is especially effective for matrices resulting from the discretization of compact integral operators such as the covariance operator C. The $\mathcal{H}$-matrix approximation first divides a matrix into sub-blocks according to a geometric admissibility condition and then approximates those sub-blocks that are far away from the diagonal by low-rank approximations, e.g., the adaptive cross-approximation algorithm. Besides computational savings, our approach combining $\mathcal{H}$-matrix representation and thick-restart Lanczos iteration also results in substantially reduced storage complexity.

The KL computations were carried out with in-house code utilizing the *HLIBpro* hierarchical matrix package developed at the Max–Planck Institute for Mathematics in the Sciences, Leipzig [75] as well as the thick-restart Lanczos implementation in the *SLEPc* package [76]. Fig. 3 shows the first few eigenfunctions that are computed on the domain under study in this work.

2.2.4. Further modeling issues

In the study cases investigated in Section 4, we model the uncertain material host rock properties thermal conductivity $\boldsymbol{\lambda}$, intrinsic hydraulic permeability $\mathbf{k}$ and the principal Young's moduli E_i as random fields. In doing so, we consider both isotropic and anisotropic models. For the latter, we distinguish statistical and material anisotropy.

Statistical anisotropy is modeled by modifying the geometry in the pointwise covariance functions

$$c_{\mathbf{A}}(\mathbf{x},\mathbf{y}) = \mathbf{Cov}_{\mathbf{A}}(Z(\mathbf{x}), Z(\mathbf{y})) := c\left(\sqrt{(\mathbf{x}-\mathbf{y})^{\top}\mathbf{A}(\mathbf{x}-\mathbf{y})}\right) \tag{17}$$

with c an isotropic stationary covariance function and $\mathbf{A} \in \mathbb{R}^{2\times 2}$ a symmetric positive definite matrix defining the modified distance (Euclidean distance is recovered when $\mathbf{A}$ is the identity matrix). In the statistically anisotropic cases considered below, the principal axes of $\mathbf{A}$ are aligned with the bedding plane (x axis) and normal to the bedding plane (y-axis), so that $\mathbf{A}$ simplifies to the diagonal matrix $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & \gamma^2 \end{bmatrix}$, where $\gamma > 0$ is a scaling parameter known as the *anisotropy ratio* (cf. [77, Section 2.5.2]).

For material anisotropies, the realizations of the random fields are symmetric positive definite 2×2 matrices, which we again assume to be aligned and orthogonal, respectively, with the bedding plane, resulting in diagonal realizations

$$\mathbf{Z}(\mathbf{x},\omega) = \begin{bmatrix} Z_{\parallel}(\mathbf{x},\omega) & 0 \\ 0 & Z_{\perp}(\mathbf{x},\omega) \end{bmatrix}. \tag{18}$$

Our models assume the diagonal entries to be fully correlated, i.e., $Z_{\perp} = sZ_{\parallel}$, with $s > 0$ used to vary the strength of (material) anisotropy (we will come to this in Section 3). Computational models for pairs of Gaussian random fields with less than full correlation can be found in [78,79], but simulating finite correlation effects is beyond the scope of the present study.

Finally, we note that we consider random material properties with both normal and lognormal probability distributions (cf. Table 2). For each uncertain property, we have an interval in which it is believed to range as well as a nominal (Best/Mean) value inside this interval. Rather than truncating the normal and lognormal distributions, the support of which are unbounded, we have scaled the distributions in the statistically homogeneous (constant but random) in such a way that 95% of realizations lie in the given intervals (cf. Appendix B). For the random field models we used a rough scaling based on the constant mean and the global variance parameter σ^2 of the random field.

3. Model setup

We consider a benchmark-type setup used in several previous studies [44,45,47]. The domain consists of a circular two-dimensional region of diameter 100 m having a circular hole in the center with a diameter of 2.48 m. The inner hole mimics a disposal drift with an

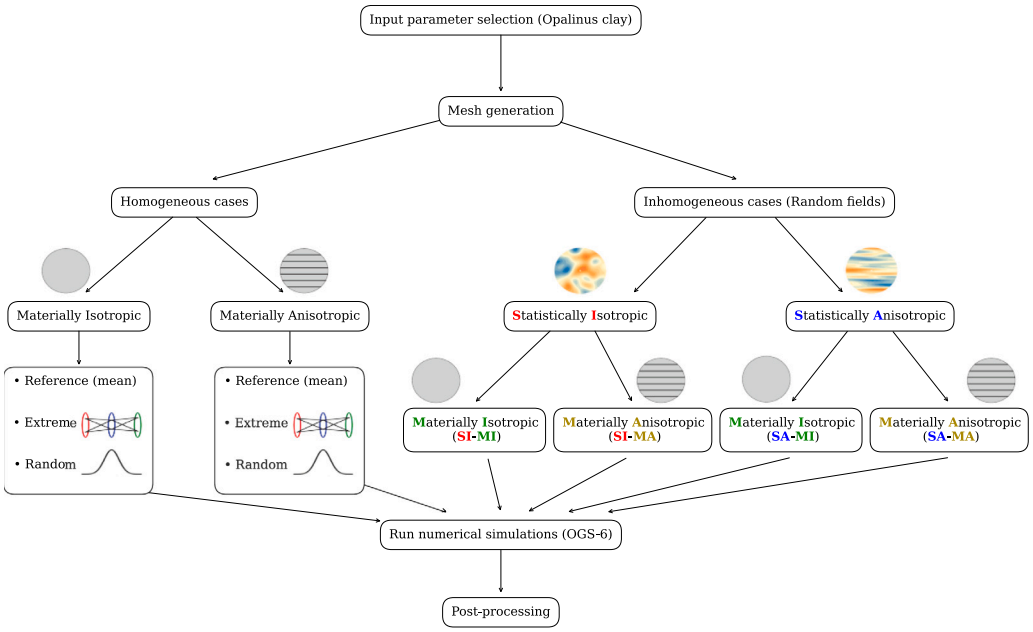

Fig. 4. Graphical representation of the study layout illustrating the steps from input parameter selection to post-processing. The diagram differentiates between homogeneous and inhomogeneous study cases, highlighting materially and statistically isotropic and anisotropic scenarios.

Table 1
Input parameters for Opalinus clay used in the analysis (excluding those used to generate random fields) [80–82]. The parameters include solid and liquid phase properties, along with mechanical properties characterized for parallel (index ∥) and normal (index ⊥) directions relative to the bedding plane.

Parameter	Symbol/Unit	Value
Solid density	$\rho/\mathrm{kg\,m^{-3}}$	2689.65
Porosity	$\phi/-$	0.13
Biot–Willis coefficient	$\alpha_B/-$	1
Specific heat capacity of solid	$c_{p,S}/\mathrm{J\,kg^{-1}\,K^{-1}}$	995
Specific heat capacity of liquid	$c_{p,L}/\mathrm{J\,kg^{-1}\,K^{-1}}$	4181.3
Thermal expansivity of solid	$\alpha_S/\mathrm{K^{-1}}$	$1.5 \cdot 10^{-5}$
Thermal expansivity of liquid	$\alpha_L/\mathrm{K^{-1}}$	$4 \cdot 10^{-4}$
Poisson's ratio (parallel, parallel)	$\nu_{\parallel\parallel}/-$	0.35
Poisson's ratio (parallel, normal)	$\nu_{\parallel\perp}/-$	0.25
Shear modulus (parallel, normal)	$G_{\parallel\perp}/\mathrm{MPa}$	1200

emplaced waste cell emitting heat, while the exterior annular domain represents the host rock. The x-axis is assumed to be the axis parallel to the bedding plane, while the y-axis is chosen to be perpendicular to the bedding plane. The initial temperature is $T_0(\mathbf{x}) = 15\,^\circ\mathrm{C}\ \forall \mathbf{x} \in D$ while the initial pore pressure is $p_0(\mathbf{x}) = 2\,\mathrm{MPa}\ \forall \mathbf{x} \in D$. At the tunnel boundary, a heat source is applied by increasing heater power from 0 W to 500 W in 30 d, to 1000 W in next 34 d and then kept constant at 1350 W for a total period of 5 years which is also the time at which all the results in this study are shown. The heater length is considered to be 4.6 m. On the tunnel boundary, pore pressure is set to be 0 thus mimicking a drained boundary. The displacement is fixed along the outer boundary. It is noteworthy that the choice of a circular geometry in our study is intentional based on the observation that using a rectangular region and constraining the normal displacement on the outer boundary introduces additional anisotropic effects which are difficult to segregate from the anisotropic effects under observation in this study. Indeed, an alternative approach could have been to increase the domain size, but it will consequently result in higher computational cost. To streamline the workflow, we used Jupyter-Notebooks. The linear triangular mesh is generated using Gmsh [83] and contains 4096 nodes and 8046 elements. The model employs a backward Euler scheme for time discretization, ensuring stability for implicit solutions of coupled processes. Time stepping is governed by an iteration-number-based adaptive strategy. The initial time step is set to 4.32×10^5 s (5 days), with a minimum of 8.64×10^2 s (0.01 days) and a maximum of 1.728×10^6 s (20 days). The time step size evolves based on the number of iterations per time step, with a multiplier factor ranging from 1.225 (for fewer iterations) to 0.625 (for higher iteration counts), enabling computational efficiency while maintaining accuracy during periods of rapid changes. A basic Newton–Raphson solver is employed to handle non-linearities in the system, with convergence criteria defined per component. The L_2-norm is used, with absolute tolerances set to 10^{-4} K for temperature, 10^{-2} Pa for pressure, and 10^{-6} m for the displacement components. The mesh resolution is selected to accurately capture the spatial variability of the random field, ensuring that the highest spatial frequencies are well-resolved relative to the mesh size. The input parameters which are not changed throughout this study are given in Table 1 while the properties of the input parameters used as random variables are given in Table 2. For the sake of simplicity, we will denote the varying input parameters i.e. λ, k and E shown in Table 2 as f ($\mathbf{f}$ in the anisotropic case) wherever no distinction is necessary. Furthermore, it is worth mentioning that throughout this work, in study cases with randomness involved, the value of an anisotropic quantity along the y-axis is kept strictly dependent on the quantity's value along the x-axis with a constant ratio s. This ratio s is computed from the best values of the available data for the input parameters in Table 2 such that $s = \mathrm{Best}(f_\perp)/\mathrm{Best}(f_\parallel)$. This ensures that the transverse ($f_\perp$) and parallel ($f_\parallel$) components of the anisotropic quantity remain fully correlated and consistent with the data. Further details of the study cases will follow in the respective section while a brief summary of these study cases is given in Table 3 as well as a workflow chart is shown in Fig. 4.

4. Study cases

As in this work, we account for both inhomogeneity and anisotropy, several different study cases were designed to isolate effects of certain modeling choices. In order to keep the discussion easy to understand, we consider it appropriate to describe the details of each study case and the corresponding results in parallel. Furthermore, due to the large volume of data and the corresponding plots generated, only selected results will be shown here. The primary variables temperature T, pore fluid pressure p and radial solid displacement u_r as well as the

Table 2
Uncertain input parameters for Opalinus clay, including their best estimates, minimal and maximal values, and distributions used to construct random variables and random fields [25,81,82]. The parameters are characterized for isotropic (iso), parallel (∥), and normal (⊥) orientations relative to the bedding plane. "Isotropic" represents averaged behavior over all orientations, "parallel" indicates measurements along the bedding plane, and "normal" refers to measurements perpendicular to it.

Parameter	Symbol/Unit	Min	Best/Mean	Max	Distribution
Thermal conductivity (isotropic)	λ_{iso}/W m^{-1} K^{-1}	1.31	1.85	2.39	Normal
Thermal conductivity (parallel)	$\lambda_{\parallel}$/W m^{-1} K^{-1}	1.7	2.4	3.1	Normal
Thermal conductivity (normal)	$\lambda_{\perp}$/W m^{-1} K^{-1}	0.92	1.3	1.68	Normal
Intrinsic permeability (isotropic)	k_{iso}/m^2	$0.6 \cdot 10^{-20}$	$2.25 \cdot 10^{-20}$	$6 \cdot 10^{-20}$	Lognormal
Intrinsic permeability (parallel)	$k_{\parallel}$/m^2	$1 \cdot 10^{-20}$	$3.75 \cdot 10^{-20}$	$10 \cdot 10^{-20}$	Lognormal
Intrinsic permeability (normal)	$k_{\perp}$/m^2	$0.2 \cdot 10^{-20}$	$0.75 \cdot 10^{-20}$	$2 \cdot 10^{-20}$	Lognormal
Young's modulus (isotropic)	E_{iso}/MPa	3000	4500	6000	Normal
Young's modulus (parallel)	$E_{\parallel}$/MPa	4000	6000	8000	Normal
Young's modulus (normal)	$E_{\perp}$/MPa	2000	3000	4000	Normal

Table 3
Summary of the study cases: in the first six homogeneous cases spatially constant values were used for the three uncertain inputs. In the two cases labeled Extreme, all eight combinations of the three extreme values Ext ∈ {Min, Max} were simulated. In the cases labeled Random, the spatially constant values of the three inputs were randomly sampled from their probability distributions. The last four cases employed (spatially varying) random fields for the three input quantities.

Study case	Runs	Material property	Correlation length	
			x	y
Homogeneous, isotropic (reference)	1	$f = \text{Best}(f_{iso})$		
Homogeneous, anisotropic (reference)	1	$\mathbf{f} = \text{diag}\left[\text{Best}(f_{\parallel}), \text{Best}(f_{\perp})\right]$		
Homogeneous, isotropic (Extreme)	8	$f = \text{Ext}(f_{iso})$		
Homogeneous, anisotropic (Extreme)	8	$\mathbf{f} = \text{diag}\left[\text{Ext}(f_{\parallel}), \text{Ext}(f_{\perp})\right]$		
Homogeneous, isotropic (Random)	10 000	$f = \text{Rand}(f_{iso})$		
Homogeneous, anisotropic (Random)	10 000	$\mathbf{f} = \text{diag}\left[\text{Rand}(f_{\parallel}), s\,\text{Rand}(f_{\parallel})\right]$		
Inhomogeneous, statistically and materially isotropic	10 000	$f = \text{RF}(f_{iso})$	15 m	15 m
Inhomogeneous, statistically isotropic, materially anisotropic	10 000	$\mathbf{f} = \text{diag}\left[\text{RF}(f_{\parallel}), s\,\text{RF}(f_{\parallel})\right]$	15 m	15 m
Inhomogeneous, statistically anisotropic, materially isotropic	10 000	$f = \text{RF}(f_{iso})$	50 m	5 m
Inhomogeneous, statistically and materially anisotropic	10 000	$\mathbf{f} = \text{diag}\left[\text{RF}(f_{\parallel}), s\,\text{RF}(f_{\parallel})\right]$	50 m	5 m

secondary variables radial Darcy velocity v_r, von Mises stress $\Delta q = \sqrt{\frac{3}{2}\boldsymbol{\sigma}'_{dev} : \boldsymbol{\sigma}'_{dev}}$ and effective hydrostatic stress $\Delta p' = -\frac{1}{3}\text{tr}\,\boldsymbol{\sigma}'$ are chosen as output variables, i.e., quantities of interest. The choice of considering these secondary variables as model outputs is justified due to their relevance in the context of the integrity criteria for evaluating geological barriers. The Darcy velocity provides information on potentially localized flow paths and consequently the direction of advective mass transport through the host rock. In addition, the relative magnitude of the von Mises stresses and the effective hydrostatic stresses can provide information about the possibility of integrity-threatening stress states that may indicate shear-induced failure. Furthermore, the negative values of effective hydrostatic stresses can provide useful insight into the possibility of tensile failure modes.

4.1. Homogeneous, isotropic and anisotropic cases

4.1.1. Reference cases

To properly delineate the effects of inhomogeneity, it is beneficial to establish a baseline based on homogeneous-isotropic and homogeneous-anisotropic cases. It is worth mentioning that these are usually the only cases that are considered within most studies. For the isotropic case, constant values $f \equiv \text{Best}(f_{iso})$ are used for the three uncertain parameters, and in the anisotropic case the constant diagonal tensors $\mathbf{f} = \text{diag}\left[\text{Best}(f_{\parallel}), \text{Best}(f_{\perp})\right]$. Fig. 5 shows the contour plot of the output variables for the homogeneous-isotropic case. We note that the computational domain used in the simulation was larger and the plots show the relevant portion of the domain extending to radius $r = 40$ m. As we remain in the heating phase, T remains at its peak at the tunnel (heater) boundary and decreases with increasing distance from the heat source. The thermal expansion caused by the heat source and the zero pressure boundary condition at the center both affect fluid flow and pressure distribution, causing pressure to peak outward from the center. The radial displacement shows convergence close to the unsupported cavity and divergence farther away, while decaying at greater distances from the center. The radial Darcy velocity v_r being related to the fluid pressure gradient can be interpreted accordingly, showing both drainage into the drift and outward fluid displacement on the far side of the pressure peak.

Slightly negative values of $\Delta p'$ can be observed in the regions with high pressures. Other regions experience strong effective pressure increases. Also, Δq reaches several MPa in some regions indicating changes in the amount of shear. The mechanical significance of these stress changes is dependent on the initial stress field and thus both site and depth dependent as well as affected by the excavation, ventilation and backfill processes which all affect the near-field stresses around the cavity. We discuss here the relative changes in these stress indicators from one study case to the next without drawing such situation-specific conclusions. In general, as the study case is homogeneous and isotropic, the results show the expected behavior, i.e., the output variables are uniform in all directions and no artificial anisotropic effects, e.g., arising from irregular boundary conditions, are present.

Fig. 6 shows the contour plots of the output variables for the homogeneous and (materially) anisotropic case for the domain trimmed at $r = 40$ m. The effect of transverse anisotropy of the input parameters can be observed in the observed quantities, as the contour lines are mostly of elliptical shape in contrast to the radially symmetric contours in the isotropic case. A closer look at the plots for p and u_r for the anisotropic case indicates that the principal axes reverse with radial distance, as the major axis of the elliptic contour lines shifts from horizontal to vertical with increasing distance from the heat source. This behavior is a result of the complex THM coupling and was also observed in our previous work [44].

Although Figs. 5 and 6 already convey significant information about the differences caused by the material anisotropy for the domain of interest, we provide the comparison plots shown in Fig. 7 for a more quantitative visualization which we will refer to as *polar plots* in the

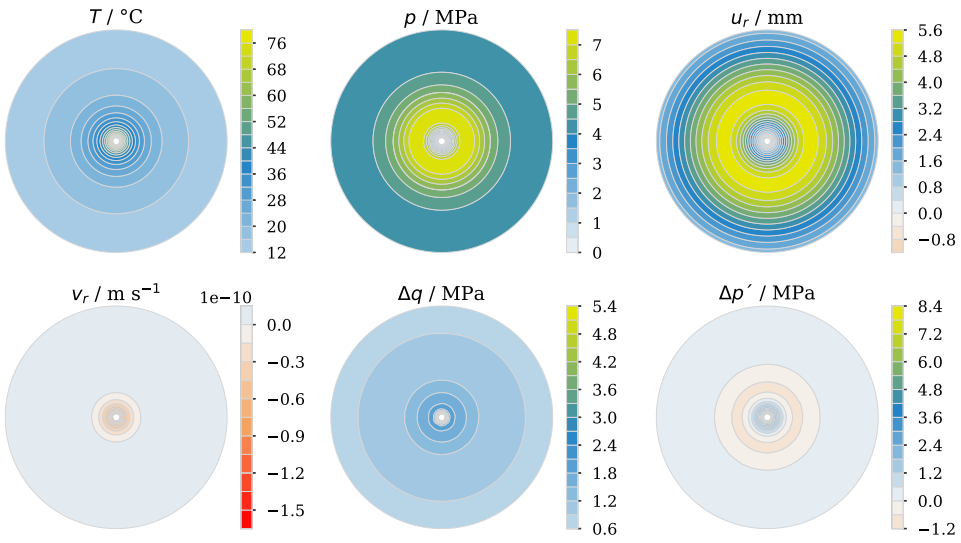

Fig. 5. Contour plot for homogeneous and isotropic case for mean values of input parameters shown up to a radius of $r = 40\,\text{m}$.

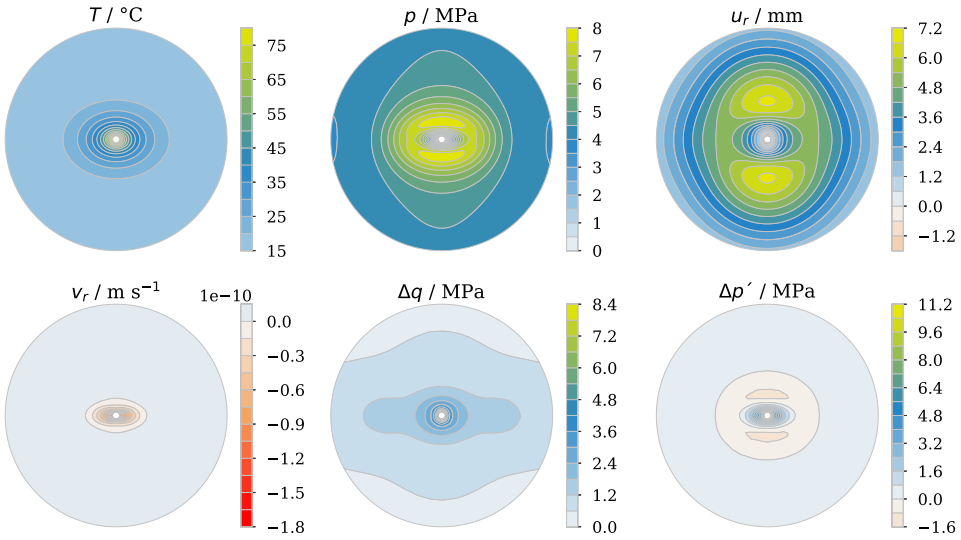

Fig. 6. Contour plot for homogeneous and (materially) anisotropic case for mean values of input parameters chosen for the corresponding case shown up to a radius of $r = 40\,\text{m}$.

following. To generate these polar plots, we fix a distance r from the heat source and plot the output variables at $(r\cos\theta, r\sin\theta)$ with $\theta \in [0°, 360°)$ in Cartesian coordinates first and then project the plot along the polar axis. In the isotropic case circular profiles are obtained in polar coordinates. It should be noted that the radius of this circle in polar coordinates does not represent the distance from the heat source, but rather the magnitude of the output variables. Consequently, the circular shape is lost in the presence of any anisotropy, which makes such a polar plot a suitable candidate to quantitatively analyze the effects of anisotropy. Fig. 7 thus shows such a polar plot for chosen output variables at distances $r = 3\,\text{m}, 8\,\text{m}, 15\,\text{m}$ for the isotropic (solid lines) and anisotropic cases (dashed lines). In case of T, the anisotropic effects are present but far less pronounced than for the other quantities. Near the heat source, p and strongly linked quantities such as v_r and $\Delta p'$ show strong anisotropic effects and even though p increases as we move away from the heat source, the anisotropic effects become less pronounced. This does not appear to be the case with u_r, though, as elastic mechanical signals propagate far instantaneously.

In the case of Δq, the difference of magnitude between isotropic and anisotropic cases is significant while the effect of anisotropy can hardly be seen near the heat source. Mechanical anisotropy causes higher anisotropy in the stress states and thus stronger shear responses. This is known also from excavations in anisotropically stressed rock masses. We observe a strong influence of anisotropy in both stress-indicators with relevance for barrier integrity, confirming that mechanical anisotropy (properties and in-situ stresses) are important considerations.

4.1.2. Extreme and random cases

As the input parameters under study are uncertain, using their best values may provide an initial indication of the magnitude of the output variables but does not capture the extent of uncertainty in the input parameters themselves. To assess the ranges obtained by different approaches, we test here two commonly used strategies and refer to them as the Extreme and Random cases. In the Extreme case, we extend the isotropic and anisotropic study cases by using selected

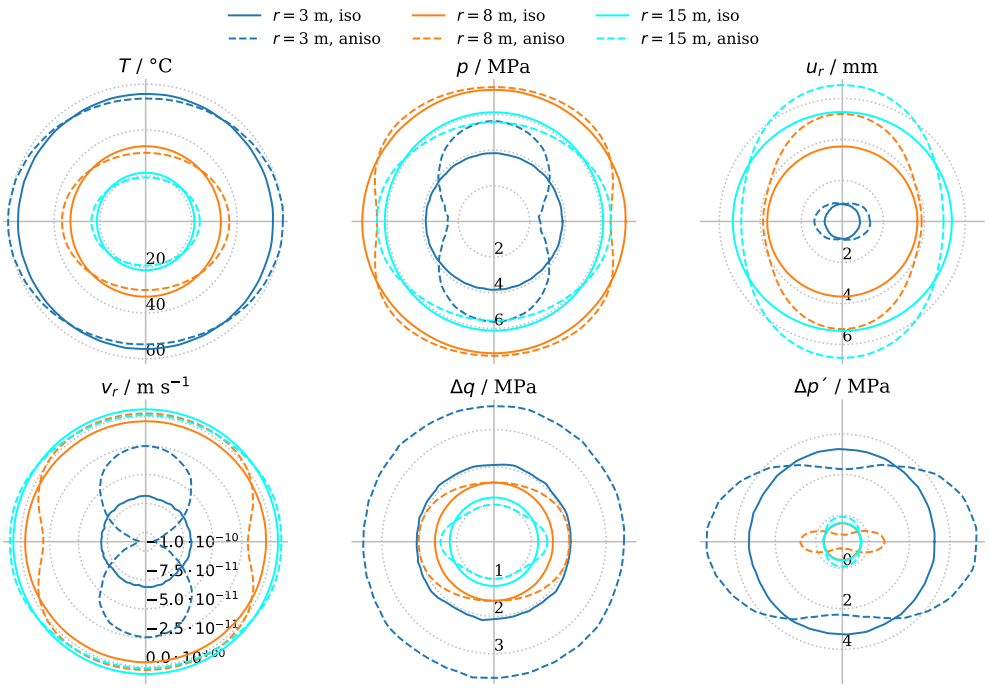

Fig. 7. Polar plots showing the output variables at different distances from the heat source for the homogeneous isotropic and homogeneous anisotropic cases.

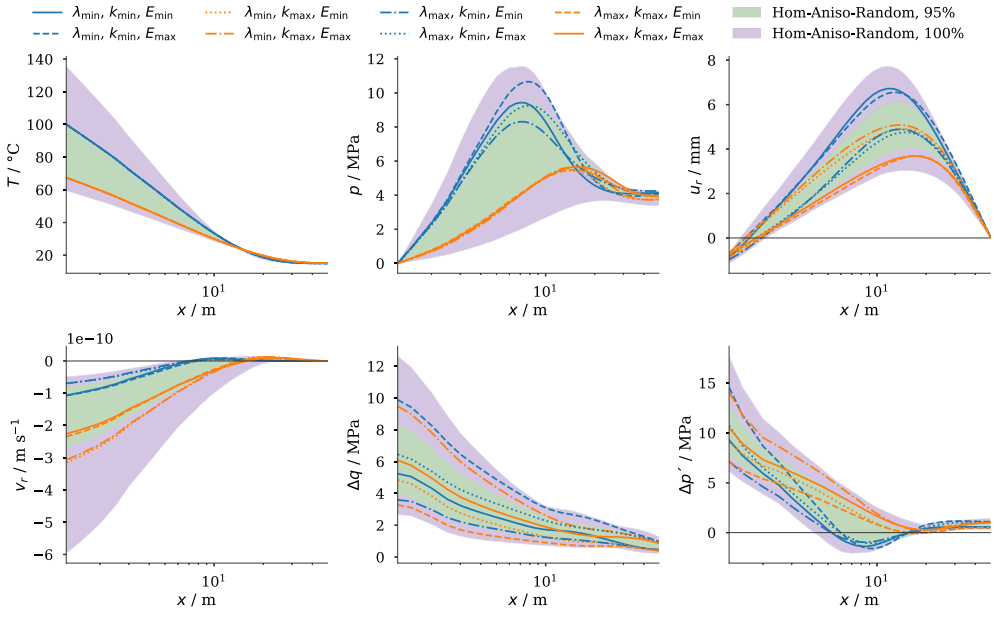

Fig. 8. Plot showing the variation of output variables along the radial distance (x-axis) from the heat source for all combinations of minimum and maximum values of input parameters in homogeneous and (materially) anisotropic cases. The x-axis represents the distance in meters on a logarithmic scale. The line plots depict the extreme cases, while the shaded regions indicate the 95% and 100% percentiles of the random cases.

combinations of extreme values (Min and Max values in Table 2) of the input parameters. For the three input parameters, this results in eight simulations each for the isotropic and anisotropic cases and is thus computationally inexpensive. For the isotropic cases the values $f \equiv \mathrm{Ext}(f_{\mathrm{iso}})$ were used and in the anisotropic cases $\mathbf{f} \equiv \mathrm{diag}\left[\mathrm{Ext}(f_{\parallel}), \mathrm{Ext}(f_{\perp})\right]$ with $\mathrm{Ext} \in \{\mathrm{Min}, \mathrm{Max}\}$. In the Random cases, as the name indicates, we treat the input parameters as random variables and sample their values from the distributions given in Table 2. We performed 10 000 simulations for each for the isotropic and anisotropic cases, respectively. Thus, in the isotropic cases, $f \equiv \mathrm{Rand}(f_{\mathrm{iso}})$ while in the anisotropic cases, $\mathbf{f} \equiv \mathrm{diag}\left[\mathrm{Rand}(f_{\parallel}), s\,\mathrm{Rand}(f_{\parallel})\right]$. The designator "Random" should not be confused with the "random field" as the former refers to homogeneous cases with (constant) values chosen randomly for each simulation while the latter only refers to the inhomogeneous cases, where each realization displays spatial variation.

Fig. 8 displays the six output variables along the x-axis for all combinations of minimum and maximum values of the three input parameters for the homogeneous and (materially) anisotropic cases. As most of the model outputs vary most strongly in the proximity of the heat source, a logarithmic scale was adopted for the spatial axis to

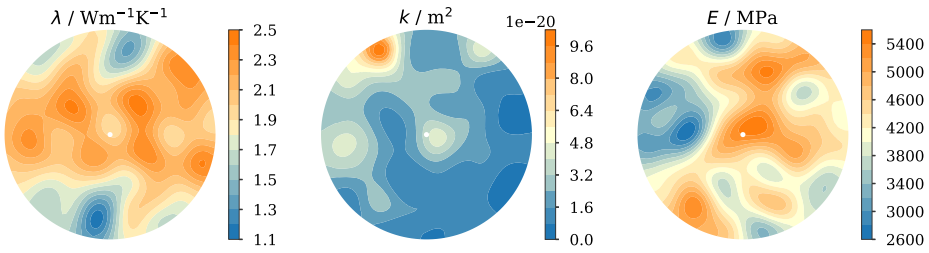

Fig. 9. One realization of statistically and materially isotropic inhomogeneous random field for input parameters.

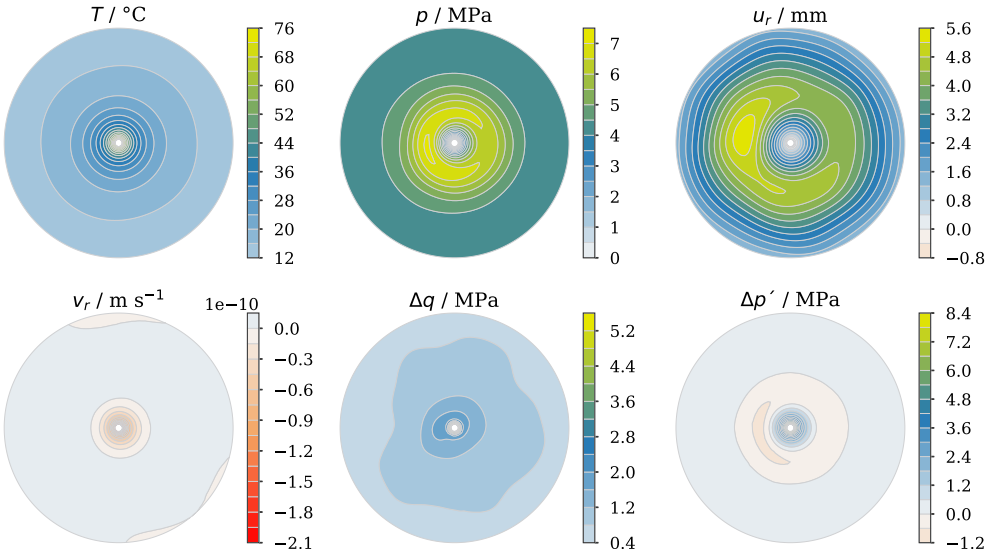

Fig. 10. Contour plot for the statistically and materially isotropic case using the random realization of the random field shown in Fig. 9 for input parameters up to a radius of $r = 40$ m.

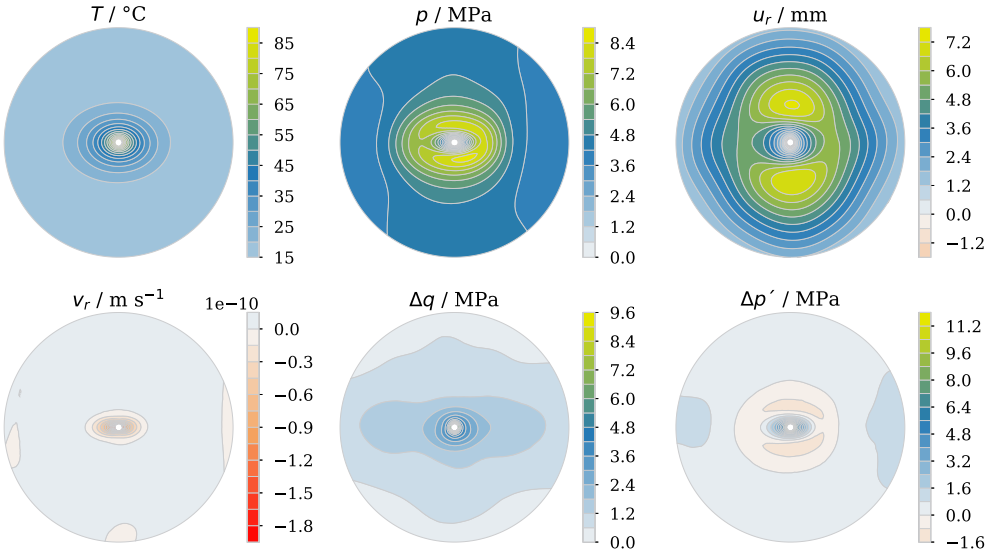

Fig. 11. Contour plot for statistically isotropic and materially anisotropic case using one random realization of a random field (not shown) for input parameters up to a radius of $r = 40$ m.

enhance visibility. The solid and dashed lines show the results for the Extreme cases, while the shaded regions indicate the 95% and 100% percentiles of the range of the six output quantities for the Random case.

We first discuss the Extreme cases and then compare with the Random cases. Considering primarily direct effects (no couplings), the input parameters under consideration are known to be inversely related to the corresponding primary process variables. For example, an increase in k should result in lower resistance to fluid flow and thus lower fluid pressure. In the case of T, we observe that all eight combinations of extremal input parameter values produce only two distinct sets of curves. A closer look at the data indicates that λ is the only significant parameter that causes a change in T. This is expected due to the fact that temperature in the THM model is mainly unilaterally coupled to

the other processes. Thus, the changes in k and E have no significant effect on T, which was also observed in our previous work [44]. As far as the nature of the relationship between T and λ is concerned, it is evident that up until around 15 m, they are inversely proportional but at greater distances faster propagation of heat becomes evident, although the magnitude of T is very small in this region. In case of p, we observe a more complex input–output relation, with k appearing to be controlling the coupling strength. For the cases with the higher value of k (orange lines), the coupling effects seem to be negligible especially in the area near the heat source, while for the cases with the lower value of k (blue lines), the coupling effects can be clearly observed, with increases in E resulting in higher values of p due to mechanical opposition to fluid expansion. An increase in λ lowers pressure peaks due to lower peak temperatures. Consequently, the highest peak value of p is observed for the case with the lowest value of λ and k and highest value of E, but notably this does not give the outer bounds for all distances from the heat source and the behavior shifts after around $r = 10$ m. In case of u_r, λ appears to be the dominating parameter due to thermal strain effects. Overall the input–output parameter dependence in case of independent variables mostly matches the observations made in our previous work [44] whereas the dependent variables were not studied before in this context and thus further improve our understanding of their behavior in terms of barrier integrity. It is, however, clear that no single combination of input parameters covers the bounds of all output variables. Furthermore, even for an individual output variable, the input parameter combination which spans the extreme values of the output quantity can be radius dependent. This simply means that all min–max perturbations should be simulated.

In the Random cases also depicted in Fig. 8, the input parameters were sampled from (non-truncated) normal and lognormal distributions, respectively, scaled in such a way that 95% of the realizations remain between the extremal values given in Table 2. This sampling approach was adopted to correspond more closely to the subsequent inhomogeneous study cases. Since this will invariably result in some physically unlikely outliers, we show the 95% percentile in addition to the total range of the output quantities as a rough indication of the physically meaningful variation. Overall, it can be observed that the bounds produced by the 95% percentiles in the Random cases are tighter than the outer bounds produced by the Extreme case for all output variables except T. In case of temperature T, which depends essentially linearly on the input quantities (cf. discussion above), the 95% quantile corresponds exactly to the extreme input cases for $\lambda_{\min}$ and $\lambda_{\max}$.

The extreme cases of all output variables except T are controlled by coupling. Thus, the sampling strategy has to select extreme-value combinations, i.e., simultaneously sample extreme values of each distribution. We generated a set of 100 000 samples and checked the extreme 200 samples for all three input parameters. None of the set of samples was close enough to the extreme set of desired values. Thus, for a coupled problem with multiple input parameters, the likelihood of sampling all combinations of extreme values naturally is low due to the extent of the parameter space and the probabilistic sampling strategy. Deliberately including Extreme cases ensures that the boundaries of the parameter space are adequately explored, a common practice in engineering studies to evaluate the system's robustness under limiting scenarios.

While we discussed these effects using plots along the x-axis for simplicity, it is evident that in anisotropic cases peak values are likely to occur along other directions, as shown previously. Thus, full-field evaluations or, at least, orthogonal axes should be studied (data not shown here).

4.2. Inhomogeneous cases

To generate inhomogeneous input variables as RFs, the covariance eigenpairs of the KL expansion are computed for the same mesh as used in the physical simulation using the approach discussed in Section 2.2. It bears mentioning that more terms in the truncated KL expansion lead to higher accuracy in the approximation of the random field, but this leads to higher resolution demands as successively higher eigenmodes become more oscillatory. Thus, a suitable compromise needs to be found between accuracy and computational cost. Based on several preliminary tests using different numbers of eigenmodes and mesh resolutions, we determined 64 eigenmodes to be a suitable choice for the mesh used in this study. In both the statistically isotropic and anisotropic cases, the retained eigenmodes were computed for a Gaussian random field with mean zero, marginal variance one and correlation length 15 m (isotropic) and 50 m and 5 m (anisotropic), respectively. The resulting KL expansions were then shifted and scaled for each input parameter according to the values given in Table 2 (cf. also Appendix B for the scaling employed). In the generation of realizations of each of the three input fields, the scalar random KL coefficients are sampled independently, resulting in statistically independent input fields. In the materially anisotropic cases, we employ independent fields for the three input variables, but the parallel and transverse components of each diagonal tensor are fully correlated and scaled according to the anisotropy ratio. The mean values are given in Table 2 while to compute the variance, we use the methodology in Appendix B in such a way that 95 % of the values remain within the intervals mentioned in Table 2. Along with the number of eigenmodes, we also specify the correlation lengths along the x- and y-axes. Here, we will refer to the cases with equal correlation lengths along both axes as the *statistically isotropic* cases, while to differing correlation lengths as the *statistically anisotropic* cases.

4.2.1. Statistically isotropic cases

In case of statistical isotropy (SI), we use an equal correlation length of 15 m along both axes. In the first case of SI, we keep the input parameters materially isotropic (MI) i.e., $f = \mathrm{RF}(f_{\mathrm{iso}})$ while in the second case, the material anisotropy (MA) is also considered such that $\mathbf{f} = \mathrm{diag}\left[\mathrm{RF}(f_{\|}), s\,\mathrm{RF}(f_{\|})\right]$. We performed simulations of 10 000 samples each for SI-MI as well as SI-MA case. Fig. 9 shows a single realization of the RFs used for the input parameters for the SI-MI case. The lognormal nature of distribution in case of k is visible in its RF as it tends more towards lower values as compared to λ and E, both of which are normally distributed. Fig. 10 shows the contour plots up to $r = 40$ m for the output variables for the inhomogeneous SI-MI case. These results correspond to the realization for which the RFs are shown in Fig. 9. In case of T, the effects of inhomogeneity are not visible in the case shown, as well as other random cases analyzed but not shown here. This is likely related to the ratio of the correlation length to the thermal diffusion length. For other output variables, the effects of inhomogeneity are more clearly visible. For example, in case of p, the area of peak value seems to be the area where the value of k (Fig. 9) is lower. Among the stress measures shown, Δq shows more irregular behavior as compared to $\Delta p'$. Furthermore, the strong dependence of u_r and $\Delta p'$ on k as in case of p can also be seen, consistent with the physical insights gained in the simpler cases above. As far as the overall peak values of the outputs are concerned, one may be interested to compare this inhomogeneous case with the homogeneous and isotropic (reference) case (Fig. 5) where mean values of the input parameters were considered. Thus, when comparing the peak values in Fig. 10 and Fig. 5, it appears as if there is not much difference in peak values in both cases. This can be due, again, to the chosen length scales but also an artifact of looking at a single realization. Thus, metrics of the ensemble should be analyzed as well.

Fig. 11 shows the contour plots of all outputs up to $r = 40$ m for the statistically isotropic and materially anisotropic (SI-MA) case for a selected RF realization. In case of T, although the effect of inhomogeneity is still not visible, the effect of material anisotropy can be observed as clearly as in the homogeneous cases. For other output variables we see the effects of both inhomogeneity and material anisotropy, although

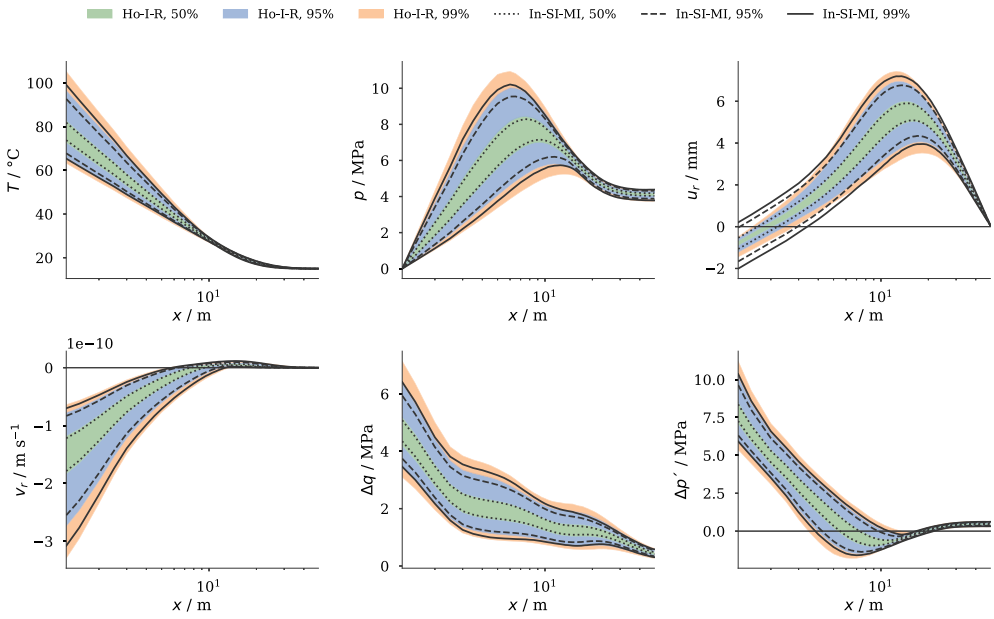

Fig. 12. Plot showing the comparison of various inter-percentile ranges (IPR) of output variables along the radial distance (x-axis) from the heat source for 10 000 realizations of random fields of input parameters. The comparison includes homogeneous-isotropic (Random) and inhomogeneous-isotropic (SI-MI) cases. The x-axis represents the radial distance in meters on a logarithmic scale.

material anisotropy seems to dominate here as we see somewhat similar ellipsoidal contours as observed for the homogeneous and anisotropic case shown in Fig. 6. This may, in addition to the previously discussed causes, be due to the fact that the tensor component along the y-axis is completely dependent on the component along the x-axis by the constant scaling factor s. Thus, it would be interesting to analyze the anisotropic effects by considering a softer or more flexible approach while generating correlated RFs along both axes.

To gain further insights into the overall extent to which the output variables vary, we adapt a common statistical measure known as inter-percentile range (IPR) which we already used partially for the homogeneous Random cases in the previous section. The IPR provides the range of data that lies between any two chosen percentiles. For example, the 50 % IPR gives the central region of data, which is the data between P25 and P75. The IPR can not only help filter out the desired data by eliminating the outliers of one's choice, but is also a useful tool to compare different data sets, which is also what we require. Thus, to compare the homogeneous and inhomogeneous study cases or the subcases within the inhomogeneous cases, we choose 50 % IPR which will give us the central range of output parameter variability, 95 % IPR which corresponds to the input data range in Table 2 as well as 99 % IPR solely for the sake of comparison.

Fig. 12 shows the comparison between homogeneous-isotropic (Random) (Ho-I-R) and inhomogeneous-isotropic (In-SI-MI) case using above-mentioned IPR values for all output variables for 10 000 realizations plotted along the x-axis. Apart from some negligible differences, both Ho-I-R and In-SI-MI give almost the same central range (50 %) for all output variables. In case of 95 % and 99 % IPR, the In-SI-MI case gives a somewhat lower range of variability for all output variables except in drift convergence. In case of T, 95 %, which corresponds to our data range, where the Ho-I-R case already shows peak values below 100 °C, the In-SI-MI case shows even lower values which might be significant when it comes to thermal repository design considerations. In case of u_r, in the regions with higher values and at relatively larger distance from the heat source, we see the similar pattern mentioned earlier but the Ho-I-R case gives a very narrow band of variability next to the heat source in contrast to the inhomogeneous case.

We also took a look at the comparison between the anisotropic cases (Ho-A-R and In-SI-MA) provided overall similar results (data not shown). To make it clearer, the results in homogeneous Random (isotropic and anisotropic) cases show differences mainly in case of u_r near the heat source. Furthermore, in case of T in Ho-A-R case, we observed the peak temperatures of 100 °C (Fig. 8) but in In-SI-MA case, we observe lower peak values of T (not shown here). This overall behavior of inhomogeneous cases giving narrower range of variability than the homogeneous Random cases might be significant in context of the site selection or repository design.

4.2.2. Statistically anisotropic cases

Due to the layered and anisotropic nature of the Opalinus clay and other sedimentary rocks, correlation lengths parallel to the bedding plane are typically much longer than perpendicular to the bedding plane. The ratio values of correlation lengths between perpendicular and parallel bedding planes for E typically vary between 0.1–0.01 and can be even smaller for k [84–86]. We refer to the cases with a non-unit correlation length ratio as statistically anisotropic cases. For this study, a ratio of $\gamma = 0.1$, specifically a correlation length of 50 m along the x-axis and 5 m along the y-axis is chosen for all input parameters. Again, we consider two study cases in this context; a statistically anisotropic and materially isotropic (SA-MI) case with $f = \mathrm{RF}(f_{\mathrm{iso}})$ and a statistically and materially anisotropic (SA-MA) case with $\mathbf{f} = \mathrm{diag}\left[\mathrm{RF}(f_{\parallel}), s\,\mathrm{RF}(f_{\parallel})\right]$. Similar to the other study cases, we used 10 000 realizations of RFs for each study case.

Fig. 13 shows a single realization of a statistically anisotropic (and materially isotropic) random field for all three input parameters. The effectiveness of the method used to generate the RFs with different correlation lengths can be observed here by the layered structure typical for sedimentary formations. Similarly, the lognormal behavior in case of k is also reproduced, showing the lower values of k dominating the majority of the RF values, yet maintaining the layered nature resulting from different correlation lengths. Fig. 14 shows the contour plots up to $r = 40$ m for the output variables for the inhomogeneous SA-MI case for which the RFs shown in Fig. 13 are used as input parameters. The effects of inhomogeneity or statistical anisotropy are again not noticeable, or in other words difficult to see in case of T. What is interesting is that not only the effects of random inhomogeneity are visible for p, u_r as well as the stress measures, but also the effects caused by statistical anisotropy can also be observed. To be specific, for

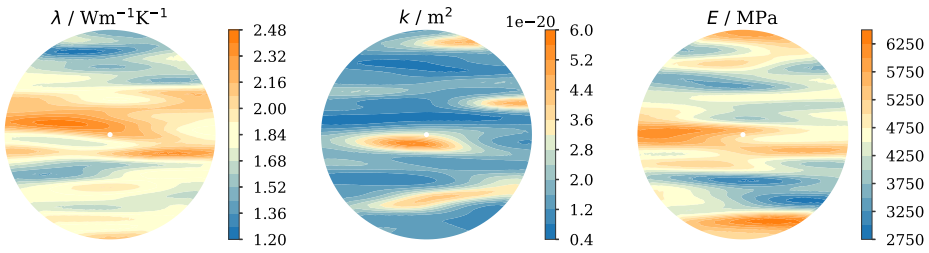

Fig. 13. One realization of statistically anisotropic inhomogeneous random field for input parameters.

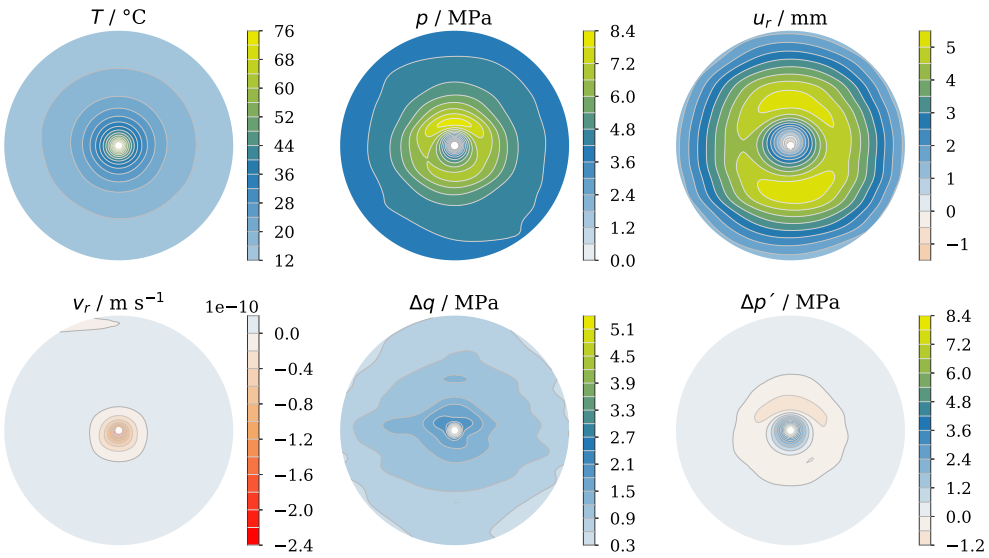

Fig. 14. Contour plot for the statistically anisotropic and materially isotropic case using the random realization of the random field shown in Fig. 13 for input parameters up to a radius of $r = 40$ m.

example, in case of p and u_r, the peak values tend to be along or very near to the vertical axis as was observed in case of material anisotropy for the homogeneous (Fig. 6) and inhomogeneous case (Fig. 11). The statistical anisotropy effects seen in Fig. 14 were also observed in several other realizations analyzed but not shown here. As far as the magnitude of the output variables is concerned, it is not of much benefit to discuss for the single realization shown here, as it varies significantly among the 10 000 simulations performed.

Fig. 15 shows the contour plots for all output variables for statistically and materially anisotropic (SA-MA) case using one random realization of random field (not shown) up to a radius of $r = 40$ m. The contours of output variables are qualitatively very similar to the homogeneous and anisotropic case (Fig. 6) for most of the domain, although the random inhomogeneity effects are still visible in some regions. As far as the magnitudes of the output variables are concerned, although most of the output variables show higher peak values than the single SA-MI case, but, as mentioned earlier, individual random realizations are not suitable for comparing peak values.

Thus, we require some additional statistical analysis similar to that shown in Fig. 12 but such a spatial plot showing IPR's along a single line is not suitable when the anisotropy comes into play. A suitable alternative is to integrate the IPR's as shown in Fig. 12 into the polar plot already introduced for the homogeneous and anisotropic (reference) case (Fig. 7) for any chosen constant distance from the heat source.

Fig. 16 shows the polar plot for all output variables showing 50 %, 95 % and 99 % inter-percentile range (IPR) for the statistically isotropic and anisotropic cases in the absence of material anisotropy at a distance of 3 m from the heat source. As a reminder, the radius of lines in such a polar plot does not represent the distance from the heat source, rather the magnitude of the output variables (IPR's in this case). In case of T, both SI-MI and SA-MI cases give the same central range (50 % IPR) but in case of 95 % and 99 % IPR, SA-MI case gives a slightly lower range of variability. A similar but even less pronounced difference is observed in the case of Δq. The effect of statistical anisotropy (elliptical shape of curves) on the other hand is hardly visible in case of T but can be slightly observed in case of Δq. The most prominent differences are observed in case of u_r where we not only observe the effects of SA, but also the switch in previously discussed major axes of the ellipsoids due to the physical couplings involved. In case of p, v_r and $\Delta p'$, we do not observe any differences neither in magnitude of variability for the SI-MI and SA-MI cases, nor any effects of statistical anisotropy. This observation is in contrast with the one made in a single random case of SA (Fig. 14). In our opinion, there can be several possible reasons for this. Firstly, the inherent nature of the type of plot i.e., percentiles or inter-percentile range across 10 000 simulations results in somewhat smoother contours than the single case, thus losing the variability or randomness if the scale of randomness is small. Secondly, although the plot gives significant information about the anisotropy and magnitude of variability across several thousand simulations it is restricted to a certain distance of the heat source.

Fig. 17 shows the polar plot for all output variables showing 50 %, 95 % and 99 % inter-percentile range (IPR) for the statistically isotropic and anisotropic cases in the absence of material anisotropy at a distance of 10 m from the heat source. It can be observed that as we move away from the heat source, the effects of statistical anisotropy can be observed even for T and the SA-MI case gives higher variability in contrast to the case nearer to the heat source. Furthermore, the switch in major axes of ellipsoids for the upper and lower bounds is also observed. Although these differences may be not play a significant

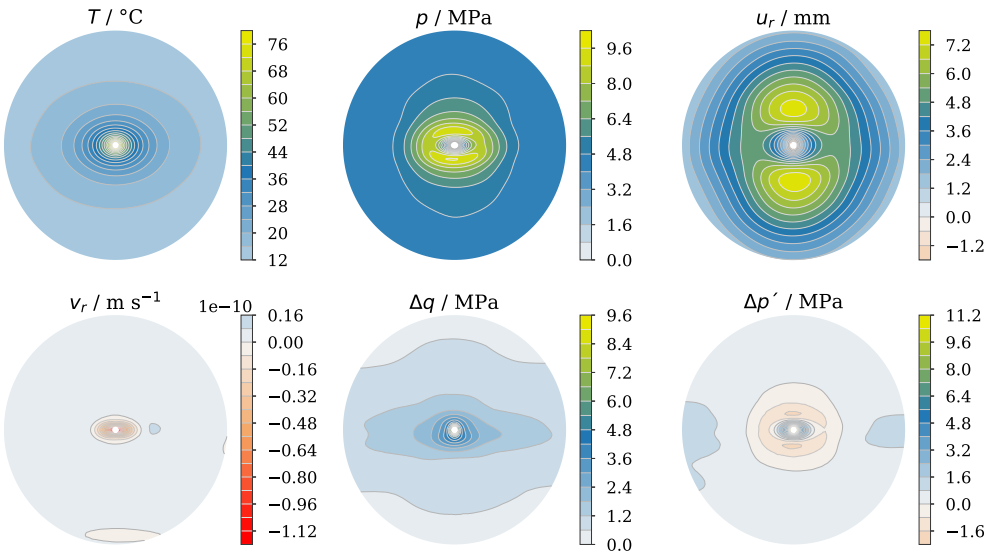

Fig. 15. Contour plot for statistically and materially anisotropic case using one random realization of a random field (not shown) for all output variables up to a radius of $r = 40$ m.

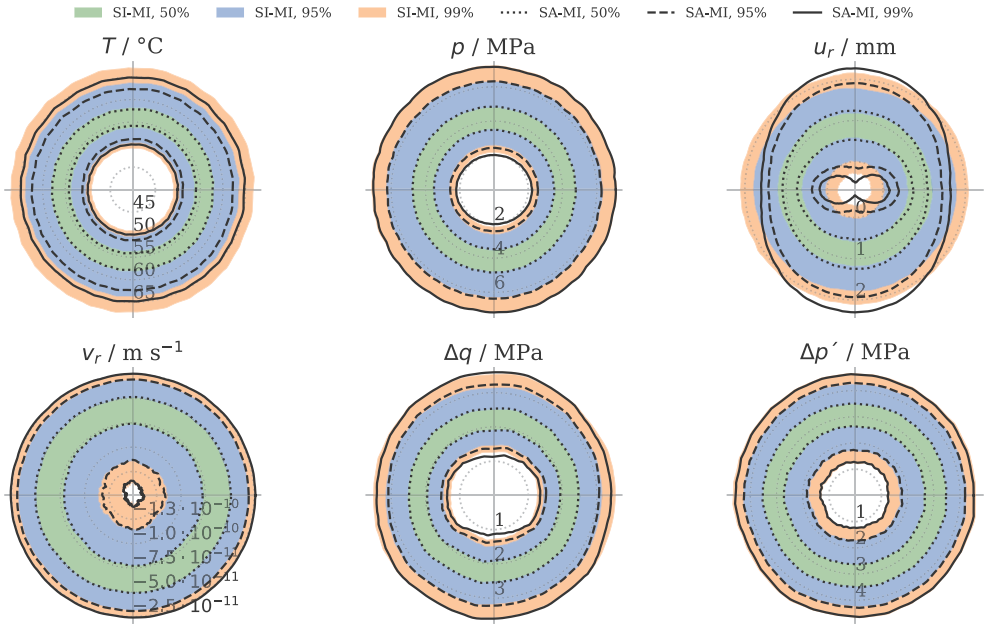

Fig. 16. Polar plot showing 50 %, 95 % and 99 % IPR for the statistically isotropic and anisotropic cases in the absence of material anisotropy for all output variables at a distance of 3 m from the heat source.

role due to the very low magnitudes of T, they illustrate the interaction between physical and statistical length scales. Furthermore, the output variables like p, v_r and $\Delta p'$ also show some influence of statistical anisotropy in contrast to the case nearer to the heat source. It is worth noting here that irrespective of whether the statistical anisotropy is involved or not, not only Δq shows higher upper bounds than $\Delta p'$ but also the values of $\Delta p'$ themselves are negative for most of the realizations at this distance. From the integrity criteria perspective, as mentioned earlier, the latter can be indicative of tensile failure. One can argue that the IPR shows the data from all 10 000 realizations in a single plot and thus does not give a direct measure of ratio of Δq and $\Delta p'$. Again, conclusive interpretation requires information on the in-situ stress field and is not attempted here.

So far, we have only compared the effects of statistical anisotropy in the absence of material anisotropy in this section. As the material anisotropy is relatively strongly embedded in the numerical model by tensor-valued random fields, it would be interesting to see if SA has any influence in the MA cases and to what extent. Fig. 18 thus shows the polar plot for all output variables showing 50 %, 95 % and 99 % inter-percentile range (IPR) for the statistically isotropic and anisotropic cases in the presence of material anisotropy at a distance of 3 m from the heat source. Except for the case of u_r, all other output variables show remarkable consistency in the overall variability in the IPR's. The u_r again shows significant differences in the presence of statistical anisotropy. What can further be observed is that the outer bounds of IPR's appear to be much smoother than in the absence of MA which was also observed and discussed earlier in the single random realization of SA-MA case (Fig. 15). As we move away from the heat source (not shown here), the strong effects of SA in case of u_r start to diminish, similar to the case observed in the absence of MA. Furthermore, other output variables does not show any distinguishable differences arising from SA as we move away from the heat source.

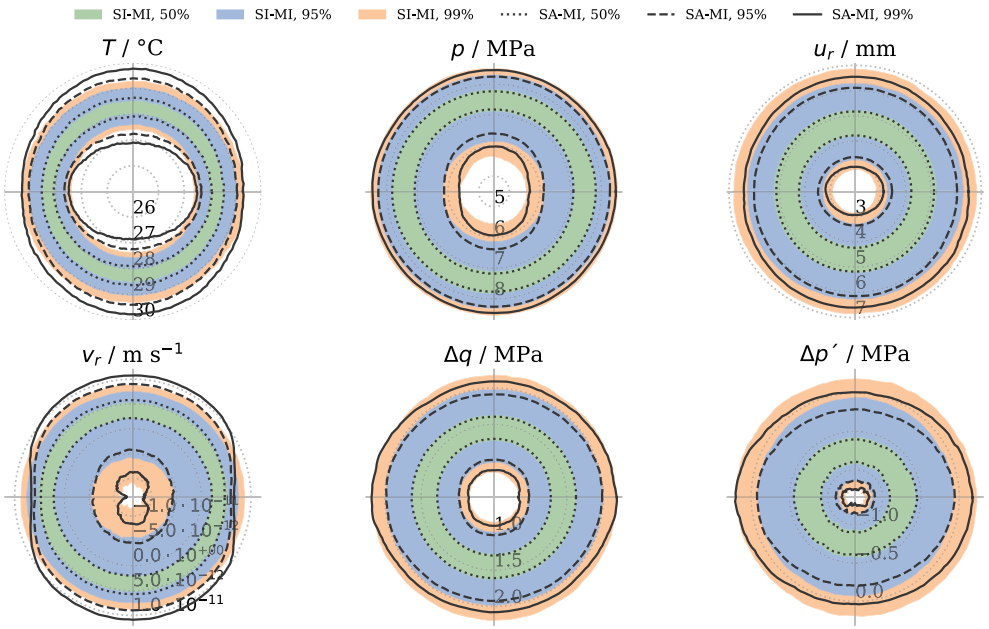

Fig. 17. Polar plot showing 50 %, 95 % and 99 % IPR for the statistically isotropic and anisotropic cases in the absence of material anisotropy for all output variables at a distance of 10 m from the heat source.

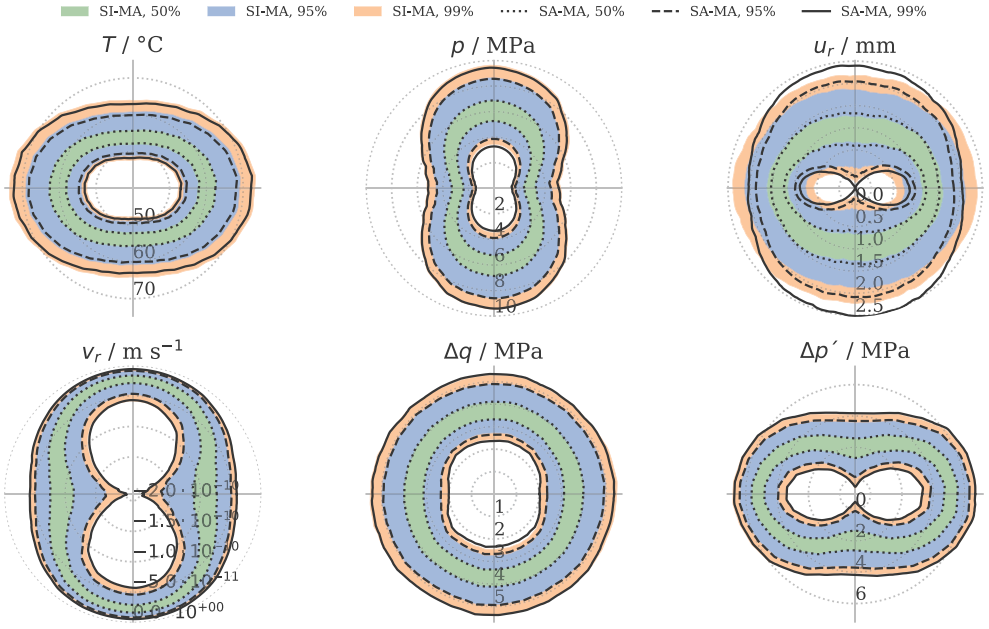

Fig. 18. Polar plot showing 50 %, 95 % and 99 % IPR for the statistically isotropic and anisotropic cases in the presence of material anisotropy for all output variables at a distance of 3 m from the heat source.

4.3. Percentiles over whole domain

To effectively analyze and interpret the results when the data comes from tens of thousands of simulation runs and from several different study cases, is a challenge. The difficulty further increases due to other factors like the presence of anisotropy or a large number of output variables each of which varies at different length and/or time scales and when for some the lowest values are of interest and the highest for others. Here we extend the contour plots used for single cases like Fig. 5 to show the percentiles for the whole domain. It is worth mentioning that computing the percentiles across the whole domain is a memory intensive process as at each node (4096 in our case), a large number of values (10000 in our case) need to be stored at the same time for all output variables. The post-processing for this purpose is done using the tool VTUFileHandler [87] written in the Julia programming language. Fig. 19 shows the 2.5 percentiles of all output variables for the statistically and materially anisotropic inhomogeneous case for the whole domain, whereas Fig. 20 shows 97.5 percentiles of all output variables for the same case for the whole domain. As mentioned earlier as well, these two percentiles are chosen based on the fact that these correspond to input data range used in this study. From the interpretation perspective, Fig. 19 can thus be said to show the lowest possible values of output variables, while Fig. 20 represents the majority of values among 10000 realizations. Such a percentile contour plot can be helpful when it comes to design decisions based on temperature or integrity criteria, for example. As can be observed in Fig. 20, one

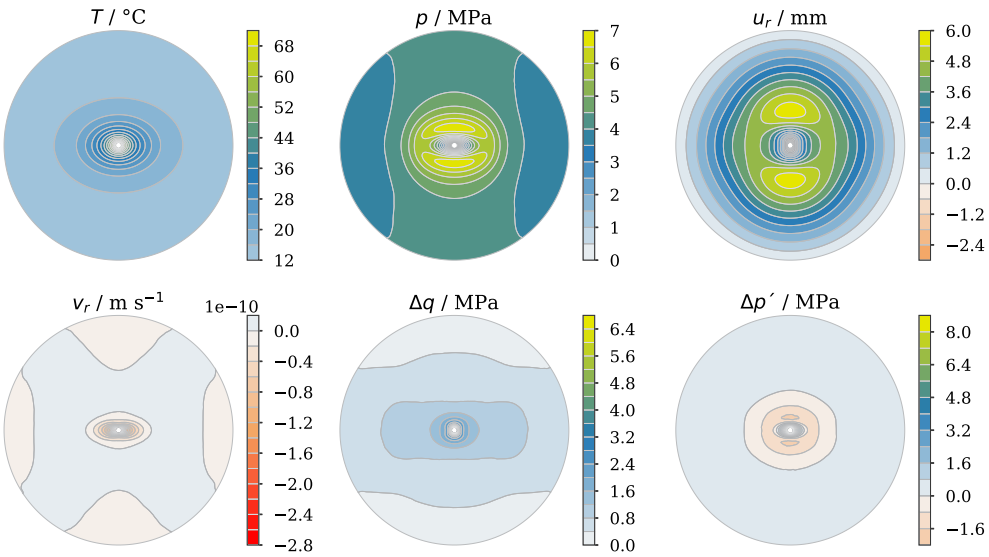

Fig. 19. 2.5 percentile plot for all output variables for the statistically and materially anisotropic inhomogeneous case for the whole domain.

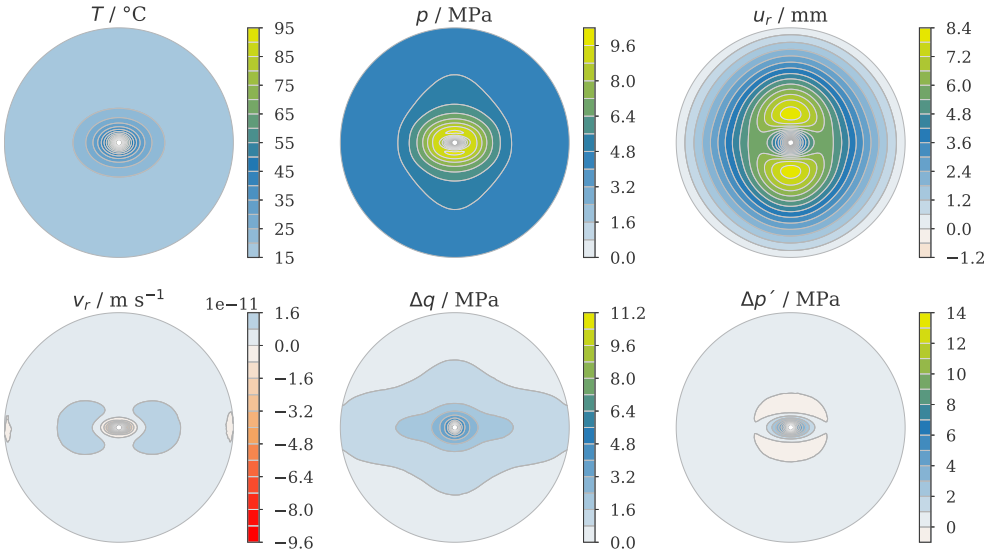

Fig. 20. 97.5 percentile plot for all output variables for the statistically and materially anisotropic inhomogeneous case for the whole domain.

can say that the peak value of T anywhere in the whole domain in any of the 10000 random scenarios analyzed, does not exceed 95 °C. It should be understood though, this is a representative benchmark study in which many structural elements were not considered and many simplifying assumptions were made. Furthermore, in this section we only show the results for the SA-MA case taking into account the largest feature set. The percentile contour plots can indeed be modified and extended to show other significant statistical measures like mean, median, standard deviation and so on. Some the statistical measures although may require prior knowledge or assumptions on the type of distribution of the output variables.

5. Conclusions

This study provides an in-depth analysis of the effects of inhomogeneity and anisotropy on THM simulations, particularly in the context of assessing the integrity of geological barriers for the deep geological disposal of radioactive waste. Through a series of numerical simulations, we analyzed the behavior of host rocks under varying conditions of key material properties, namely thermal conductivity, intrinsic permeability, and Young's modulus. The results from the homogeneous and isotropic case provided a baseline understanding of the system's behavior under uniform conditions, showing expected patterns of heat transfer, fluid flow, and mechanical deformation in such an ideal scenario. By contrast, the introduction of anisotropy revealed more complex interactions, leading to pronounced variations in the output variables. Our analysis further explored the Extreme and Random cases, highlighting the variability in outcomes due to the probabilistic uncertainty of the input parameters. The analysis showed that, although the Extreme cases are computationally inexpensive and can be effective in indicating extreme output bounds as well as better understanding of the effects of couplings, they are not associated with likelihood measures in contrast to the bounds in the Random case. The comparison between homogeneous and inhomogeneous cases revealed that the latter provides a more realistic and comprehensive understanding of the potential variability in system responses and can indicate local concentrations in flow or stress fields. The homogeneous random cases, while avoiding the need to generate random fields, might overestimate variability observed in inhomogeneous simulations. Statistically anisotropic cases, which incorporated realistic correlation

lengths based on geological formations like Opalinus clay, provided a closer approximation of real-world scenarios. These cases demonstrated that the statistical anisotropy of the material properties plays a crucial role in the accuracy of THM simulations. To be specific, the effects of statistical anisotropy were observed for different variables at different distances from the heat source and were less prominent in case of pressure than temperature and displacement. The effects of material anisotropy were understandably observed to be stronger than that of statistical anisotropy. This observation points again to the relevance of length scales, specifically the ratio between those of spatial property fluctuation, physical process and geometry/domain. In fact, as alluded to in the introduction, material anisotropy can often be interpreted as a manifestation of structural inhomogeneity/heterogeneity, upscaled to the representative elementary volume (REV) level, transitioning from the micro to meso-scale of the current model's length scale.

One of the key findings of this study is the impact of anisotropy and inhomogeneity on stress measures, i.e., the von Mises stresses and effective hydrostatic stresses. The results showed that in the presence of anisotropy, there was a notable increase in stress-based indicators, pointing to a higher likelihood of failure, especially near the heat source. This is particularly relevant for site selection and repository design.

Furthermore, this study underscores the importance of extending THM analyzes to capture the long-term impacts of anisotropy and inhomogeneity on geological barrier integrity. While the current work provides critical insights into material variability at a fixed instance in time linked to the chosen scale, future studies should focus on exploring the effects of inhomogeneity and anisotropy over larger scales and extended time frames. Building on previous work [44,88], such advancements would significantly enhance understanding of safety analyses for deep geological repositories.

While generating the random fields, the study adapted a moderately conservative approach. A more conservative approach for the future study can be to use the same random field realizations, i.e., fully correlated random fields, for all input parameters, while a less conservative approach can be to make the random fields along x and y-axis completely independent. Future work could also explore the implications of varying correlation lengths or ratios to reflect different geological settings, as well as the effects of incorporating additional uncertain parameters. Investigating the behavior under unsaturated conditions or including structural components such as bentonite in the model could also provide deeper insights into the coupled processes affecting barrier integrity. Another valuable direction would be the integration of field-scale data into simulations, allowing for the calibration and validation of models against real-world conditions. Given the computationally intensive nature of the experiments in this work, it would also be of interest to consider first building a surrogate to emulate the mapping from random fields to quantities of interest, as demonstrated in [31], and then perform analysis based on the surrogate in place of the costly THM modeling.

CRediT authorship contribution statement

Aqeel Afzal Chaudhry: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Chao Zhang:** Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Oliver G. Ernst:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Thomas Nagel:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Table A.4
Summary of physical variables and mathematical symbols used in Section 2.1. Dimensions are denoted as follows: L (length), t (time), T (temperature), and M (mass). Units are provided for a consistent system of units.

Description	Symbol	Dimension	Unit
Density of the medium	ρ	ML^{-3}	$kg\,m^{-3}$
Specific heat capacity	c_p	$L^2t^{-2}T^{-1}$	$J\,kg^{-1}\,K^{-1}$
Density of gas phase	ρ_{GR}	ML^{-3}	$kg\,m^{-3}$
Density of liquid phase	ρ_{LR}	ML^{-3}	$kg\,m^{-3}$
Specific heat capacity of liquid	c_{pL}	$L^2t^{-2}T^{-1}$	$J\,kg^{-1}\,K^{-1}$
Temperature	T	T	K
Porosity	ϕ	–	–
Liquid saturation	S_L	–	–
Capillary pressure	p_{cap}	$ML^{-1}t^{-2}$	Pa
Gas-phase pressure	p_{GR}	$ML^{-1}t^{-2}$	Pa
Liquid-phase pressure	p_{LR}	$ML^{-1}t^{-2}$	Pa
Volumetric latent heat of water vaporization	L_0	ML^2t^{-2}	$J\,m^{-3}$
Volumetric vapor content	θ_{vap}	–	–
Effective thermal conductivity	λ_{eff}	$MLt^{-3}T^{-1}$	$W\,m^{-1}\,K^{-1}$
Advective mass flux of liquid	A_L	$ML^{-2}t^{-1}$	$kg\,s^{-1}\,m^{-2}$
Diffusive mass flux of water in gas phase	J_G^W	$ML^{-2}t^{-1}$	$kg\,s^{-1}\,m^{-2}$
Density of water in gas phase	ρ_{GR}^W	ML^{-3}	$kg\,m^{-3}$
Biot coefficient	α_B	–	–
Solid compressibility	$\beta_{p,SR}$	$L^1M^{-1}t^2$	Pa^{-1}
Solid displacement vector	$\mathbf{u}_S$	L	m
Solid thermal expansion coefficient	$\alpha_{T,SR}$	T^{-1}	K^{-1}
Heat source term	Q_T	$ML^{-1}t^{-3}$	$W\,m^{-3}$
Source/sink term for liquid	Q_H	$ML^{-3}t^{-1}$	$kg\,m^{-3}\,s^{-1}$
Dynamic viscosity of liquid	μ_{LR}	$ML^{-1}t^{-1}$	Pa s
Specific heat capacity of water vapor	$c_{p,vap}$	$L^2t^{-2}T^{-1}$	$J\,kg^{-1}\,K^{-1}$
Material time derivative	$\frac{d}{dt}$	t^{-1}	s^{-1}
Divergence operator	div	–	–

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research was funded by the Federal Company for Radioactive Waste Disposal (BGE), Germany under Grant Number STAFuE-21-4-Klei, and we gratefully acknowledge their support. This study is conducted as part of the URS research cluster established by BGE, which is dedicated to identifying a suitable site for a high-level radioactive waste repository in Germany [89]. Chao Zhang was additionally supported by the Villum Foundation, Denmark under Grant Number 25893. We would also like to extend our gratitude to Dr.-Ing. Julia Sorgatz and Dr.-Ing. Maximilian Bittens for their valuable comments and technical assistance throughout this work.

Appendix A. Nomenclature

The variables and symbols used in Section 2.1 are summarized in Table A.4. This table serves as a quick reference to facilitate the understanding of the equations and analyses presented. Each entry provides the symbol, its corresponding physical dimension, and the unit of measurement in a consistent system of units.

Appendix B. Rescaling of random fields

A Gaussian or lognormal random field $\{a(x) : x \in D\}$ on a bounded domain D will take on all values in $\mathbb{R}$ or $\mathbb{R}^+$, respectively, with positive, if small, probability. If the uncertain quantity modeled by the random field a is known to lie in an interval $[a_{\min}, a_{\max}]$, strictly enforcing this range requires applying a mapping to a with range in this interval. An approximate alternative would be to at least rescale the random field

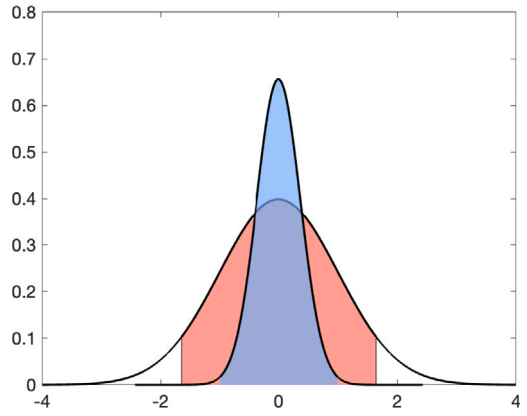

Fig. 21. Density of standard normal distribution with centered (red) portion containing $\alpha = 0.9$ probability mass, density of scaled RV $Y = cZ$ with c chosen so that same shaded (blue) probability mass lies in interval $[-1, 1]$.

in such a way that a given, sufficiently large fraction $\alpha \in [0, 1]$ of its probability mass lies in the desired interval.

B.1. Affine rescaling of scalar random variables

We begin with some scaling considerations for scalar random variables.

B.1.1. Gaussian random variables

For a scalar standard normal random variable (RV) $Z \sim \mathsf{N}(0, 1)$, the probability mass contained in an interval symmetric about the origin $[-b, b]$ for some $b > 0$, i.e., the probability that Z lies in this interval, is given by

$$\begin{aligned}\mathbf{P}(-b \le Z \le b) &= \int_{-b}^{b} \varphi(z)\,\mathrm{d}z \\ &= \int_{-\infty}^{b} \varphi(z)\,\mathrm{d}z - \int_{-\infty}^{-b} \varphi(z)\,\mathrm{d}z = \Phi(b) - \Phi(-b) = 2\Phi(b) - 1,\end{aligned} \tag{19}$$

where φ denotes the standard normal pdf, Φ the standard normal cdf and we have used the symmetry property $\Phi(-b) = 1 - \Phi(b)$. Thus, requiring $\mathbf{P}(Z \in [-b, b]) = \alpha$ for some $\alpha \in [0, 1]$ is equivalent with $\Phi(b) = (\alpha + 1)/2$. The red shading in Fig. 21 indicates the area under the pdf of the standard normal distribution centered at the origin containing the probability mass $\alpha = 0.9$.

Suppose now that b is given and we wish to rescale the RV $Z \sim \mathsf{N}(0, 1)$ with a positive scaling factor c in such a way that a given probability mass α of $Y = cZ$ is contained in $[-b, b]$. Since $\mathbf{Var}(cX) = c^2\mathbf{Var}\,X$ for any RV X with finite variance, we have $Y = cZ \sim \mathsf{N}(0, c^2)$. Moreover, the cdf F_Y of Y may be expressed in terms of Φ as

$$F_Y(y) = \mathbf{P}(Y \le y) = \mathbf{P}(cZ \le y) = \mathbf{P}\left(Z \le \frac{y}{c}\right) = \Phi\left(\frac{y}{c}\right), \tag{20}$$

hence $F_Y(y) = \Phi(y/c)$. The requirement $\mathbf{P}(-b \le Y \le b) = \alpha$ is thus equivalent with

$$\alpha = \mathbf{P}(-b \le Y \le b) = F_Y(b) - F_Y(-b) = \Phi\left(\frac{b}{c}\right) - \Phi\left(\frac{-b}{c}\right) = 2\Phi\left(\frac{b}{c}\right) - 1, \tag{21}$$

from which we obtain the desired scaling factor as

$$c = \frac{b}{\Phi^{-1}\left(\frac{\alpha+1}{2}\right)}, \tag{22}$$

in which Φ^{-1} denotes the inverse of the (strictly monotone) The blue shaded region in Fig. 21 shows the area under the pdf of a rescaled Gaussian RV for which the interval $[-1, 1]$ centered at the origin contains the probability mass $\alpha = 0.9$. By adding an offset $\mu \in \mathbb{R}$ a given probability mass may be assigned to any bounded interval in this way. Specifically, for a not necessarily centered interval $[a_{\min}, a_{\max}] \subset \mathbb{R}$, we shift the standard normal Z by $\mu := (a_{\min} + a_{\max})/2$, rescale by $c > 0$ and require that for $Y := \mu + cZ$ we have

$$\begin{aligned}\alpha = \mathbf{P}(a_{\min} \le Y \le a_{\max}) &= \mathbf{P}(a_{\min} \le \mu + cZ \le a_{\max}) = \mathbf{P}\left(\frac{a_{\min} - \mu}{c} \le Z \le \frac{a_{\max} - \mu}{c}\right) \\ &= \Phi\left(\frac{a_{\max} - \mu}{c}\right) - \Phi\left(\frac{a_{\min} - \mu}{c}\right) = \Phi\left(\frac{a_{\max} - a_{\min}}{2c}\right) - \Phi\left(-\frac{a_{\max} - a_{\min}}{2c}\right) \\ &= \Phi\left(\frac{a_{\max} - a_{\min}}{2c}\right) - \left[1 - \Phi\left(\frac{a_{\max} - a_{\min}}{2c}\right)\right] \\ &= 2\Phi\left(\frac{a_{\max} - a_{\min}}{2c}\right) - 1.\end{aligned} \tag{23}$$

Solving for the scaling factor c then yields, analogous to the centered interval,

$$c = \frac{a_{\max} - a_{\min}}{2\Phi^{-1}\left(\frac{\alpha+1}{2}\right)}. \tag{24}$$

B.1.2. Lognormal random variables

A lognormal RV has the form $X = \exp(\mu + \sigma Z)$ with $Z \sim \mathsf{N}(0, 1)$, hence $\log X \sim \mathsf{N}(\mu, \sigma^2)$. Since the range of X is the positive real line, we can rescale X such that for suitable $c \ge 0$ there holds $\mathbf{P}(cX \in [0, b]) = \alpha$ for given $\alpha \in [0, 1]$ (see Fig. 22).

Since $\log(cX) = \log c + \mu + \sigma Z$, we have

$$cX \in [0, b] \Leftrightarrow Z \le \frac{-\mu + \log b - \log c}{\sigma} \tag{25}$$

and thus requiring $\mathbf{P}(cX \in [0, b]) = \alpha$ leads to

$$\Phi\left(\frac{-\mu + \log b - \log c}{\sigma}\right) \overset{!}{=} \alpha, \quad \text{or} \quad c = b\exp\left(-\mu - \sigma\Phi^{-1}(\alpha)\right). \tag{26}$$

To determine the pdf of the rescaled density $Y = cX = c\exp(\mu + \sigma Z)$, we first note that the cdf is given by

$$F_Y(y) = \mathbf{P}(Y \le y) = \mathbf{P}\left(c\exp(\mu + \sigma Z) \le y\right) = \Phi\left(\frac{-\mu + \log(y/c)}{\sigma}\right) \tag{27}$$

and obtain the associated pdf as

$$\begin{aligned}f_Y(y) &= \frac{\mathrm{d}}{\mathrm{d}y}\Phi\left(\frac{-\mu + \log(y/c)}{\sigma}\right) = \varphi\left(\frac{-\mu + \log(y/c)}{\sigma}\right)\frac{1}{\sigma y} \\ &= \frac{1}{2\pi\sigma y}\exp\left(-\frac{(\mu + \log(y/c))^2}{2\sigma^2}\right).\end{aligned} \tag{28}$$

If a more symmetric scaling is desired, i.e., given an interval $[a_{\min}, a_{\max}]$ with $0 \le a_{\min} < a_{\max}$ which we would like to contain a probability mass $\alpha \in [0, 1]$ with equal tail probabilities

$$\mathbf{P}(0 \le X \le a_{\min}) = \frac{1 - \alpha}{2} \quad \text{and} \quad \mathbf{P}(X \ge a_{\max}) = \frac{1 - \alpha}{2} \tag{29}$$

then the left inequality is equivalent with

$$\begin{aligned}\frac{1 - \alpha}{2} &= \mathbf{P}(\exp(\mu + \sigma Z) \le a_{\min}) = \mathbf{P}\left(\mu + \sigma Z \le \log a_{\min}\right) \\ &= \mathbf{P}\left(Z \le \frac{-\mu + \log a_{\min}}{\sigma}\right) = \Phi\left(\frac{-\mu + \log a_{\min}}{\sigma}\right),\end{aligned} \tag{30}$$

which yields the relation

$$\frac{-\mu + \log a_{\min}}{\sigma} = \Phi^{-1}\left(\frac{1 - \alpha}{2}\right). \tag{31}$$

In the same way, the right inequality gives

$$\mathbf{P}(X > a_{\max}) = 1 - \mathbf{P}(X \le a_{max}) = \frac{1 - \alpha}{2} \Leftrightarrow \mathbf{P}(X \le a_{\max}) = \frac{1 + \alpha}{2} \tag{32}$$

and thus

$$\frac{1 + \alpha}{2} = \mathbf{P}\left(\exp(\mu + \sigma Z) \le a_{\max}\right) = \mathbf{P}\left(Z \le \frac{-\mu + \log a_{\max}}{\sigma}\right)$$

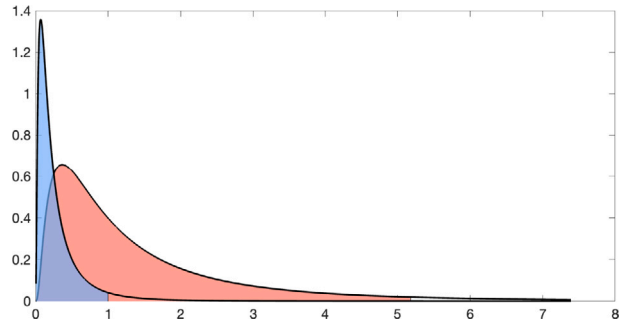

Fig. 22. Density of lognormal distribution $X = \exp(Z)$ for $Z \sim \mathsf{N}(0,1)$ (red) with probability mass $\alpha = 0.9$ shaded and density of rescaled RV $Y = cX$ such that same probability mass is contained in $[0,1]$.

$$= \Phi\left(\frac{-\mu + \log a_{\max}}{\sigma}\right) \tag{33}$$

and hence a second relation

$$\frac{-\mu + \log a_{\max}}{\sigma} = \Phi^{-1}\left(\frac{1+\alpha}{2}\right). \tag{34}$$

Solving these for the shift and rescaling parameters μ and σ gives

$$\sigma = \frac{\log a_{\max} - \log a_{\min}}{q_+ - q_-}, \qquad \mu = \frac{1}{q_+ - q_-}\left(q_+ \log a_{\min} - q_- \log a_{\max}\right) \tag{35}$$

with

$$q_+ := \Phi^{-1}\left(\frac{1+\alpha}{2}\right), \qquad q_- := \Phi^{-1}\left(\frac{1-\alpha}{2}\right). \tag{36}$$

B.2. Random fields

Now assume we are given a random field $a = a(x) = a(x,\omega)$ on a bounded domain D and underlying probability space $(\Omega, \mathfrak{A}, \mathbf{P})$ by its KL expansion

$$a(x) = a_0(x) + \sum_{m=1}^{\infty} \sqrt{\lambda_m} a_m(x)\, \xi_m, \qquad x \in D, \tag{37}$$

with uncorrelated RV $\xi_m = \xi_m(\omega)$ having zero mean and unit variance as well as eigenfunctions a_m of the associated covariance operator normalized so that $\|a_m\|_{L^2(D)} = 1$, and we would like to rescale a in such a way that its fluctuation $\tilde{a} = a - a_0$ has a given probability mass in an interval centered about the origin, i.e., $\mathbf{P}(|\tilde{a}| \le b) = \alpha$.

Denoting by $c = c(x,y)$ the two-point covariance function of a, we have by definition

$$c(x,y) = \mathbf{Cov}(a(x), a(y)) = \mathbf{E}\left[\left(a(x) - a_0(x)\right)\left(a(y) - a_0(y)\right)\right], \quad x, y \in D. \tag{38}$$

It is a consequence of Mercer's theorem (or follows directly once the KL expansion (37) is given) that

$$c(x,y) = \sum_{m=1}^{\infty} \lambda_m a_m(x) a_m(y). \tag{39}$$

In particular, at any point $x \in D$, the variance of the random field a is given by

$$\mathbf{Var}\, a(x) = c(x,x) = \sum_{m=1}^{\infty} \lambda_m a_m(x)^2. \tag{40}$$

We will restrict our considerations to covariance functions from the Matérn family of stationary covariance kernels normalized so that $c(x,x) = 1$.

B.2.1. Truncation

The covariance eigenvalues form a square summable sequence of nonnegative numbers converging to zero at a rate determined by the smoothness of the covariance kernel. For computational purposes it is therefore natural to truncate the KL expansion after a finite number $M \in \mathbb{N}$ of terms to

$$a(x) \approx a_M(x) := a_0(x) + \sum_{m=1}^{M} \sqrt{\lambda_m} a_m(x)\, \xi_m, \qquad x \in D. \tag{41}$$

At a single point $x \in D$, (40) becomes

$$\mathbf{Var}\, a_M(x) = \sum_{m=1}^{M} \lambda_m a_m(x)^2 \tag{42}$$

and so by

$$\mathbf{Var}\, a(x) - \mathbf{Var}\, a_M(x) = \sum_{m=M+1}^{\infty} \lambda_m a_m(x)^2 = \mathbf{E}\left[(a(x) - a_M(x))^2\right] \ge 0 \tag{43}$$

the variance of a_M always underestimates that of a.

Viewing the random fields as square integrable random variables with values in $L^2(D)$, the approximation error is naturally measured in the norm of $L^2(\Omega; L^2(D))$, which is given by

$$\begin{aligned} \|a - a_M\|^2_{L^2(\Omega;L^2(D))} &= \mathbf{E}\left[\|a - a_M\|^2_{L^2(D)}\right] = \mathbf{E}\left[\left\|\sum_{m>M} \sqrt{\lambda_m} a_m(x)\, \xi_m\right\|^2_{L^2(D)}\right] \\ &= \mathbf{E}\left[\int_D \left(\sum_{m>M} \sqrt{\lambda_m} a_m(x)\, \xi_m\right)^2 \mathrm{d}x\right] \\ &= \mathbf{E}\left[\int_D \sum_{m>M} \lambda_m a_m(x)^2\, \xi_m^2\, \mathrm{d}x\right] = \sum_{m>M} \lambda_m, \end{aligned} \tag{44}$$

where we have used the fact that the eigenfunctions are orthonormal in $L^2(D)$ and the RVs ξ_m are pairwise uncorrelated with unit variance.

Exchanging the order of integration results in

$$\begin{aligned} \|a - a_M\|^2_{L^2(\Omega;L^2(D))} &= \mathbf{E}\left[\|a - a_0\|^2_{L^2(D)} - \|a_M - a_0\|^2_{L^2(D)}\right] \\ &= \int_D \mathbf{Var}\, a(x)\, \mathrm{d}x - \sum_{m=1}^{M} \lambda_m = |D| - \sum_{m=1}^{M} \lambda_m, \end{aligned} \tag{45}$$

where $|D|$ denotes the Lebesgue measure of the domain D and we have assumed the field to be stationary with normalized covariance. Thus an approximation of the M dominant covariance eigenvalues yields an estimate of the error due to truncation of the KL expansion.

B.2.2. Scaling of random fields

If the random field a is Gaussian, then the random coefficients ξ_m in the KL expansion are each standard normal and pairwise independent and the same holds for the random coefficients in the KL expansion of a_M. Thus, the variation of the random field at a given location $x \in D$ is a random superposition of all included eigenmodes a_m, and the scaling approach outlined above for scalar random variables does not apply, and a rough global scaling based on the local mean and global total variance can be applied.

Data availability

A significant portion of the code utilized in this work is open-source and has been appropriately referenced. The remaining code can be made available upon request, subject to feasibility.

References

[1] Bear J. Dynamics of fluids in porous media. Courier Corporation; 2013.

[2] Zhang C. Thermo-hydro-mechanical behavior of clay rock for deep geological disposal of high-level radioactive waste. J Rock Mech Geotech Eng 2018;10(5):992–1008.

[3] Plúa C, Vu MN, Armand G, Rutqvist J, Birkholzer J, Xu H, et al. A reliable numerical analysis for large-scale modelling of a high-level radioactive waste repository in the Callovo-Oxfordian claystone. Int J Rock Mech Min Sci 2021;140:104574.

[4] Hudson JA, Stephansson O, Andersson J. Guidance on numerical modelling of thermo-hydro-mechanical coupled processes for performance assessment of radioactive waste repositories. Int J Rock Mech Min Sci 2005;42(5–6):850–70.

[5] Narkuniene A, Poskas P, Kilda R, Bartkus G. Uncertainty and sensitivity analysis of radionuclide migration through the engineered barriers of deep geological repository: Case of RBMK-1500 SNF. Reliab Eng Syst Saf 2015;136:8–16.

[6] Zheng L, Rutqvist J, Xu H, Birkholzer JT. Coupled THMC models for bentonite in an argillite repository for nuclear waste: Illitization and its effect on swelling stress under high temperature. Eng Geol 2017;230:118–29.

[7] Tosoni E, Salo A, Govaerts J, Zio E. Comprehensiveness of scenarios in the safety assessment of nuclear waste repositories. Reliab Eng Syst Saf 2019;188:561–73.

[8] Tsang C, Barnichon J, Birkholzer J, Li XL, Liu H, Sillen X. Coupled thermo-hydro-mechanical processes in the near field of a high-level radioactive waste repository in clay formations. Int J Rock Mech Min Sci 2012;49:31–44.

[9] Cadini F, Gioletta A, Zio E. Improved metamodel-based importance sampling for the performance assessment of radioactive waste repositories. Reliab Eng Syst Saf 2015;134:188–97.

[10] Saveleva E, Svitelman V, Blinov P, Valetov D. Sensitivity analysis and model calibration as a part of the model development process in radioactive waste disposal safety assessment. Reliab Eng Syst Saf 2021;210:107521.

[11] Buchwald J, Kolditz O, Nagel T. Design-of-experiment (DoE) based history matching for probabilistic integrity analysis—A case study of the FE-experiment at Mont Terri. Reliab Eng Syst Saf 2024;244:109903.

[12] Coduto DP, Yeung MR, Kitch WA. Geotechnical engineering: Principles and practices. Pearson; 2010.

[13] Rechard RP, Stockman CT. Waste degradation and mobilization in performance assessments for the Yucca Mountain disposal system for spent nuclear fuel and high-level radioactive waste. Reliab Eng Syst Saf 2014;122:165–88.

[14] Chen X, Bose N, Brito M, Khan F, Thanyamanta B, Zou T. A review of risk analysis research for the operations of autonomous underwater vehicles. Reliab Eng Syst Saf 2021;216:108011.

[15] Mazzoccola D, Hudson J. A comprehensive method of rock mass characterization for indicating natural slope instability. Q J Eng Geol Hydrogeol 1996;29(1):37–56.

[16] Schön JH. Physical properties of rocks: fundamentals and principles of petrophysics. Elsevier; 2015.

[17] Jiwen T, Shaobai X, Zhongjie Z, Hongbing L, Zhouxun Y, Kezhong S, et al. Deep structure pattern, anisotropy and continental geodynamics revealed by geophysical profiles and transects in China. In: Geophysics. CRC Press; 2020, p. 21–40.

[18] Liu W, Li A, Fang W, Love PE, Hartmann T, Luo H. A hybrid data-driven model for geotechnical reliability analysis. Reliab Eng Syst Saf 2023;231:108985.

[19] Cao B, Yin Q, Guo Y, Yang J, Zhang L, Wang Z, et al. Field data analysis and risk assessment of shallow gas hazards based on neural networks during industrial deep-water drilling. Reliab Eng Syst Saf 2023;232:109079.

[20] Zhang Y. Stochastic methods for flow in porous media: coping with uncertainties. Soil Science Society of America; 2006.

[21] Wellmann JF, Horowitz FG, Schill E, Regenauer-Lieb K. Towards incorporating uncertainty of structural data in 3D geological inversion. Tectonophysics 2010;490(3–4):141–51. http://dx.doi.org/10.1016/j.tecto.2010.04.022.

[22] Thiele ST, Jessell MW, Lindsay M, Wellmann JF, Pakyuz-Charrier E. The topology of geology 2: Topological uncertainty. J Struct Geol 2016;91:74–87. http://dx.doi.org/10.1016/j.jsg.2016.08.010.

[23] Thiele ST, Jessell MW, Lindsay M, Ogarko V, Wellmann JF, Pakyuz-Charrier E. The topology of geology 1: Topological analysis. J Struct Geol 2016;91:27–38. http://dx.doi.org/10.1016/j.jsg.2016.08.009.

[24] Alibeikloo M, Khabbaz H, Fatahi B. Random field reliability analysis for time-dependent behaviour of soft soils considering spatial variability of elastic visco-plastic parameters. Reliab Eng Syst Saf 2022;219:108254.

[25] Bossart P, Bernier F, Birkholzer J, Bruggeman C, Connolly P, Dewonck S, et al. Mont Terri rock laboratory, 20 years of research: introduction, site characteristics and overview of experiments. 2018, p. 3–22, Mont Terri Rock Laboratory, 20 Years: Two Decades of Research and Experimentation on Claystones for Geological Disposal of Radioactive Waste.

[26] Pyrcz M, Deutsch C. Geostatistical reservoir modeling. Oxford; 2014.

[27] Wackernagel H. Multivariate geostatistics: an introduction with applications. Springer Science & Business Media; 2003.

[28] Hariri-Ardebili MA. Safety and reliability assessment of heterogeneous concrete components in nuclear structures. Reliab Eng Syst Saf 2020;203:107104.

[29] Follin S, Hartley L, Rhén I, Jackson P, Joyce S, Roberts D, et al. A methodology to constrain the parameters of a hydrogeological discrete fracture network model for sparsely fractured crystalline rock, exemplified by data from the proposed high-level nuclear waste repository site at Forsmark, Sweden. Hydrogeol J 2014;22(2):313–31.

[30] Joyce S, Hartley L, Applegate D, Hoek J, Jackson P. Multi-scale groundwater flow modeling during temperate climate conditions for the safety assessment of the proposed high-level nuclear waste repository site at Forsmark, Sweden. Hydrogeol J 2014;22(6):1233.

[31] Ernst OG, Sprungk B, Zhang C. Uncertainty modeling and propagation for groundwater flow: a comparative study of surrogates. GEM Int J Geomath 2024;15(1):11.

[32] Cherpeau N, Caumon G, Lévy B. Stochastic simulations of fault networks in 3D structural modeling. C R Géosci 2010;342(9):687–94.

[33] Wang H, Wellmann JF, Li Z, Wang X, Liang RY. A segmentation approach for stochastic geological modeling using hidden Markov random fields. Math Geosci 2017;49:145–77.

[34] Cvetkovic V, Frampton A. Transport and retention from single to multiple fractures in crystalline rock at Äspö (Sweden): 2. Fracture network simulations and generic retention model. Water Resour Res 2010;46(5):1–17. http://dx.doi.org/10.1029/2009WR008030, URL http://doi.wiley.com/10.1029/2009WR008030.

[35] Norberg T, Rosén L, Baran A, Baran S. On modelling discrete geological structures as Markov random fields. Math Geol 2002;34:63–77.

[36] Hadgu T, Karra S, Kalinina E, Makedonska N, Hyman JD, Klise K, et al. A comparative study of discrete fracture network and equivalent continuum models for simulating flow and transport in the far field of a hypothetical nuclear waste repository in crystalline host rock. J Hydrol 2017;553:59–70.

[37] Kim H, Inoue J. Analytical approach for anisotropic permeability through a single rough rock joint under shear deformation. J Geophys Res: Solid Earth 2003;108(B8).

[38] Tsang C, Neretnieks I, Tsang Y. Hydrologic issues associated with nuclear waste repositories. Water Resour Res 2015;51(9):6923–72.

[39] Shao H, Sönnke J, Morel J, Krug S. In situ determination of anisotropic permeability of clay. Phys Chem Earth Parts A/ B/ C 2011;36(17–18):1688–92.

[40] Gao F, Xue Y, Gao Y, Zhang Z, Teng T, Liang X. Fully coupled thermo-hydro-mechanical model for extraction of coal seam gas with slotted boreholes. J Nat Gas Sci Eng 2016;31:226–35.

[41] Xue Y, Liu J, Ranjith P, Liang X, Wang S. Investigation of the influence of gas fracturing on fracturing characteristics of coal mass and gas extraction efficiency based on a multi-physical field model. J Pet Sci Eng 2021;206:109018.

[42] Salager S, François B, Nuth M, Laloui L. Constitutive analysis of the mechanical anisotropy of Opalinus Clay. Acta Geotech 2013;8:137–54.

[43] Barton N, Quadros E. Anisotropy is everywhere, to see, to measure, and to model. Rock Mech Rock Eng 2015;48(4):1323–39.

[44] Chaudhry AA, Buchwald J, Nagel T. Local and global spatio-temporal sensitivity analysis of thermal consolidation around a point heat source. Int J Rock Mech Min Sci 2021;139:104662.

[45] Buchwald J, Chaudhry AA, Yoshioka K, Kolditz O, Attinger S, Nagel T. DoE-based history matching for probabilistic uncertainty quantification of thermo-hydro-mechanical processes around heat sources in clay rocks. Int J Rock Mech Min Sci 2020;134:104481.

[46] Wang W, Rutqvist J, Görke U, Birkholzer JT, Kolditz O. Non-isothermal flow in low permeable porous media: A comparison of Richards' and two-phase flow approaches. Environ Earth Sci 2010;62:1197–207. http://dx.doi.org/10.1007/s12665-010-0608-1.

[47] Pitz M, Kaiser S, Grunwald N, Kumar V, Buchwald J, Wang W, et al. Non-isothermal consolidation: A systematic evaluation of two implementations based on multiphase and Richards equations. Int J Rock Mech Min Sci 2023;170:105534.

[48] Bilke L, Flemisch B, Kalbacher T, Kolditz O, Helmig R, Nagel T. Development of open-source porous media simulators: principles and experiences. Transp Porous Media 2019;130:337–61.

[49] Richards LA. Capillary conduction of liquids through porous mediums. Physics 1931;1(5):318–33.

[50] Grunwald N, Lehmann C, Maßmann J, Naumov D, Kolditz O, Nagel T. Non-isothermal two-phase flow in deformable porous media: Systematic open-source implementation and verification procedure. Geomech Geophys Geo-Energ Geo-Resour 2022;8(107).

[51] Wang W, Shao H, Nagel T, Kolditz O. Analysis of coupled thermal-hydro-mechanical processes during small scale in situ heater experiment in Callovo-Oxfordian clay rock introducing a failure-index permeability model. Int J Rock Mech Min Sci 2021;142:104683. http://dx.doi.org/10.1016/j.ijrmms.2021.104683, URL https://linkinghub.elsevier.com/retrieve/pii/S1365160921000691.

[52] Graupner BJ, Thatcher K, Friedenberg L, Guo R, Hadgu T, Hou W, et al. An international study on THM modelling of the Full-Scale Heater experiment at Mont Terri Laboratory. Geomech Energy Environ 2025;100631. http://dx.doi.org/10.1016/j.gete.2024.100631, URL https://linkinghub.elsevier.com/retrieve/pii/S2352380824000984.

[53] Sharqawy MH, Lienhard JH, Zubair SM. Thermophysical properties of seawater: a review of existing correlations and data. Desalination Water Treat 2010;16(1–3):354–80.

[54] Kitanidis PK. Introduction to geostatistics: Applications to hydrogeology. Cambridge University Press; 1997.

[55] Stein ML. Interpolation of spatial data: Some theory for kriging. Springer series in statistics, Springer-Verlag; 1999.

[56] Lord GJ, Powell CE, Shardlow T. An introduction to computational stochastic PDEs, vol. 50. Cambridge University Press; 2014.
[57] Montoya-Noguera S, Zhao T, Hu Y, Wang Y, Phoon K. Simulation of non-stationary non-Gaussian random fields from sparse measurements using Bayesian compressive sampling and Karhunen-Loève expansion. Struct Saf 2019;79:66–79.
[58] Zheng Z, Beer M, Nackenhorst U. An iterative multi-fidelity scheme for simulating multi-dimensional non-Gaussian random fields. Mech Syst Signal Process 2023;200:110643.
[59] Dietrich CR, Newsam GN. Fast and exact simulation of stationary Gaussian processes through circulant embedding of the covariance matrix. SIAM J Sci Comput 1997;18(4):1088–107. http://dx.doi.org/10.1137/S1064827592240555.
[60] Harbrecht H, Peters M, Schneider R. On the low-rank approximation by the pivoted Cholesky decomposition. Appl Numer Math 2012;62(4):428–40. http://dx.doi.org/10.1016/j.apnum.2011.10.001.
[61] Bachmayr M, Graham IG, Nguyen VK, Scheichl R. Unified analysis of periodization-based sampling methods for matérn covariances. SIAM J Numer Anal 2020;58(5):2953–80. http://dx.doi.org/10.1137/19M1269877.
[62] Karhunen K. Über lineare methoden in der Wahrscheinlichkeitsrechnung. Ann Acad Sci Fenn Ser A. I. Math-Phys 1947;37:1–79.
[63] Loève M. Probability theory II. 4th ed.. Graduate texts in mathematics, vol. 46, New York: Springer-Verlag; 1978.
[64] Anstett-Collin F, Goffart J, Mara T, Denis-Vidal L. Sensitivity analysis of complex models: coping with dynamic and static inputs. Reliab Eng Syst Saf 2015;134:268–75.
[65] Huang S, Liu L. New Kriging methods for efficient system slope reliability analysis considering soil spatial variability. Reliab Eng Syst Saf 2024;245:109989.
[66] Pronzato L. Sensitivity analysis via Karhunen–Loève expansion of a random field model: Estimation of Sobol'indices and experimental design. Reliab Eng Syst Saf 2019;187:93–109.
[67] Thapa M, Missoum S. Uncertainty quantification and global sensitivity analysis of composite wind turbine blades. Reliab Eng Syst Saf 2022;222:108354.
[68] Ghanem R, Spanos PD. Stochastic finite elements: A spectral approach. New York: Springer-Verlag; 1991.
[69] Van Trees HL. Detection, estimation and modulation theory: Part 1. Wiley; 1968.
[70] Sauter SA, Schwab C. Boundary element methods. Springer series in computational mathematics, vol. 39, Springer; 2011, http://dx.doi.org/10.1007/978-3-540-68093-2.
[71] Wu K, Simon H. Thick-restart lanczos method for large symmetric eigenvalue problems. SIAM J Matrix Anal Appl 2000;22(2):602–16.
[72] Wu L, Xue F, Stathopoulos A. TRPL+K: Thick-restart preconditioned Lanczos+K method for large symmetric eigenvalue problems. SIAM J Sci Comput 2019;41(2):A1013–40. http://dx.doi.org/10.1137/17M1157568.
[73] Eiermann M, Ernst OG, Ullmann E. Computational aspects of the stochastic finite element method. Comput Vis Sci 2007;10(1):3–15.
[74] Hackbusch W. Hierarchical matrices: Algorithms and analysis, vol. 49. Springer; 2015.
[75] Kriemann R. HLIBpro. 2024, URL https://www.hlibpro.com.
[76] Hernandez V, Roman JE, Vidal V. SLEPc: A scalable and flexible toolkit for the solution of eigenvalue problems. ACM Trans Math Software 2005;31(3):351–62.
[77] Chilès J, Delfiner P. Geostatistics: Modeling spatial uncertainty. 2nd ed.. New York: John Wiley and Sons, Inc.; 2012.
[78] Gneiting T, Kleiber W, Schlather M. Matérn cross-covariance functions for multivariate random fields. J Amer Statist Assoc 2010;105(491):1167–77. http://dx.doi.org/10.1198/jasa.2010.tm09420.
[79] Cho H, Venturi D, Karniadakis GE. Karhunen–Loève expansion for multi-correlated stochastic processes. Probab Eng Mech 2013;34:157–67. http://dx.doi.org/10.1016/j.probengmech.2013.09.004.
[80] Lisjak A, Garitte B, Grasselli G, Müller H, Vietor T. The excavation of a circular tunnel in a bedded argillaceous rock (Opalinus Clay): short-term rock mass response and FDEM numerical analysis. Tunn Undergr Space Technol 2015;45:227–48.
[81] Bossart P, Thury M. Characteristics of the Opalinus clay at Mont Terri. Rep Swiss Geol Surv 2011;3.
[82] Bock H. RA experiment. Updated review of the rock mechanics properties of the Opalinus clay of the Mont Terri URL based on laboratory and field testing. Unpubl Mont Terri Tech Rep 2009;68.
[83] Geuzaine C, Remacle J. Gmsh: A 3-D finite element mesh generator with built-in pre-and post-processing facilities. Internat J Numer Methods Engrg 2009;79(11):1309–31.
[84] Houben ME, Urai J. In situ characterization of the microstructure and porosity of Opalinus Clay (Mont Terri rock laboratory, Switzerland) [Ph.D. thesis], Hochschulbibliothek der Rheinisch-Westfälischen Technischen Hochschule Aachen; 2013.
[85] Bertrand F, Collin F. Anisotropic modelling of Opalinus Clay behaviour: From triaxial tests to gallery excavation application. J Rock Mech Geotech Eng 2017;9(2):235–46, URL https://www.sciencedirect.com/science/article/pii/S167477551630052X.
[86] Zhang C, Laurich B. Investigation on anisotropy of mechanical properties of Callovo-Oxfordian claystone. Eng Geol 2019;254:1–10, URL https://www.sciencedirect.com/science/article/pii/S0013795218313917.
[87] Bittens M. VTUFileHandler: A VTU library in the Julia language that implements an algebra for basic mathematical operations on VTU data. J Open Source Softw 2022;7(73):4300.
[88] Chaudhry AA, Buchwald J, Kolditz O, Nagel T. Consolidation around a point heat source (correction and verification). Int J Numer Anal Methods Geomech 2019;43(18):2743–51.
[89] Kurgyis K, Achtziger-Zupančič P, Bjorge M, Boxberg MS, Broggi M, Buchwald J, et al. Uncertainties and robustness with regard to the safety of a repository for high-level radioactive waste: introduction of a research initiative. Environ Earth Sci 2024;83(2):82.