

This is the preprint of the contribution published as:

Buchwald, J., Kolditz, O., Nagel, T. (2024):

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. **244** , art. 109903

The publisher's version is available at:

<https://doi.org/10.1016/j.res.2023.109903>

Design-of-Experiment (DoE) based history matching for probabilistic integrity analysis – a case study of the FE-experiment at Mont Terri

J. Buchwald^{a,b,*}, O. Kolditz^{a,c}, T. Nagel^{b,d}

^a*Helmholtz Centre for Environmental Research – UFZ, Permoserstraße 15, 04318 Leipzig, Germany*

^b*Chair of Soil Mechanics and Foundation Engineering, Geotechnical Institute, Technische Universität Bergakademie Freiberg, Freiberg, Germany*

^c*Applied Environmental Systems Analysis, Technische Universität Dresden, Germany*

^d*Freiberg Center for Water Research – ZeWäF, Technische Universität Bergakademie Freiberg, Freiberg, Germany*

Abstract

We present an application of design-of-experiment (DoE) based history matching as an approach to reduce and investigate parameter uncertainties in finite-element models for repositories of high-level radioactive waste. We combine experimental data from the FE-experiment at the Mont Terri underground research laboratory in Switzerland with thermo-hydro-mechanical modeling using the open-source package OpenGeoSys. Uncertainties were reduced by an initial parameter screening to find heavy hitters and an experiment-matching procedure using Monte-Carlo sampling on a Gaussian proxy model to fit the error between modeling response and the experiment. Furthermore, we performed a global sensitivity analysis based on the proxy model, demonstrating the spatial impact of parameter sensitivities. Very good agreement between the experimental data and the model was found for the temperature response, whereas the pressure match hints at a significant remaining gap in the physical models and/or structure. This gap could not be filled within the scope of our contribution and needs further investigation.

Keywords: Design of Experiments, history matching, thermo-hydro-mechanical, radioactive waste, geological repository, uncertainty quantification, sensitivity analysis, OpenGeoSys

2000 MSC: 62K20

2000 MSC: 49Q12

2000 MSC: 60-08

2000 MSC: 60G15

1. Introduction

To date, the computational uncertainty assessment as part of simulations of high-level radioactive waste repositories remains a rather challenging task. This is mainly due to

*Corresponding author

Email addresses: joerg.buchwald@ufz.de (J. Buchwald), thomas.nagel@ufz.de (T. Nagel)

Preprint submitted to Reliability Engineering & System Safety

September 11, 2024

model complexity and large spatial and temporal scale ranges relevant to the repository system evolution. These challenges explain, at least in part, why literature data are only available on single, very specific aspects of the entire repository system. Although a lot of research has been done based solely on transport phenomena, e.g. [1, 2, 3, 4], only a few studies exist focusing on uncertainties involving thermo-hydro-mechanical processes [5, 6, 7, 8, 9, 10]. However, the hydraulic and mechanical integrity of geotechnical and geological barriers play a central role in retarding radionuclide migration in the post-closure phase of a repository. Therefore, the uncertainty of THM processes must be taken into account in safety assessment procedures. A common approach to speed-up coupled simulations is by the replacement of specific physical processes with surrogate-models, as done by Hu and Pfingsten for the saturation profile of granular bentonite in the FE-experiment [10]. However, these models require large experimental data sets or extensive modelling for building the proxy and have usually only very restricted validity.

The origins of uncertainties in computational models mimicking those processes are manifold: We distinguish between (i) parameter uncertainties stemming from their natural variability over space and time or a lack of detailed knowledge, including measurement uncertainties. Furthermore, there are (ii) model uncertainties stemming from reducible simplifications like geometric simplifications, mathematical simplifications, or the numerical discretization. This includes the more challenging gaps in a model due to unknown or deliberately omitted effects, and also rare events that present challenges both for the elaboration of quantitative physical theories and their statistical description. Specific to repository models, significant disturbances to the natural state are caused by the excavation of the mine itself, which dominates the early stages of repository evolution [11]. The excavation-induced changes of hydro-mechanical boundary conditions and chemical equilibria in the near-field also represent initial conditions for the post-closure phase. In this study, we include the process of excavation to generate a reasonable initial mechanical state for the heating phase. Nevertheless, we perform the excavation run deterministically, neglecting the effects coming from its variability because of a lack of data and to keep the problem still tractable. The most significant post-closure effects are due to emitted decay heat, triggering major changes in the physical properties and states of the host rock in the near-field. Chemical and biological processes also play an important role in the later stages when canister corrosion begins to affect it. While there exist several studies dealing with heterogeneities and transport phenomena [12, 13, 14, 15], to date, it remains unclear what role heterogeneities at different scales play in the safety assessment when studying the thermo-hydro-mechanical evolution of a repository. The types of processes and also their impact depends very much on the host rock in which the canister is emplaced [16]. An overview can be found elsewhere [17, 18, 19].

Based on the Full-Scale Emplacement (FE) experiment in the underground research lab at Mont Terri [20], we provide a first attempt to model uncertainties based on a coupled three-dimensional thermo-hydro-mechanical (THM) model with focus on uncertainties relating purely to the host rock. To speed up simulations, we incorporate an effective description of mechanical effects into a non-isothermal flow model, as proposed by [21] for saturated conditions and generalize the approach to unsaturated media under the Richards assumption in the present study.

For uncertainty quantification, we use a history matching workflow based on design-of-experiments (DoE), a surrogate model approach widely used in different disciplines (e.g., for reservoir modeling [22]) and proposed by the authors in [8] to treat uncertainties in

the evolution of radioactive waste repositories. The use of metamodeling tools is not new in the field, e.g. [15] used a surrogate model to accelerate uncertainty quantification of computationally expensive multiphase flow simulations involving heterogeneous porous media with high-dimensional input and function-valued output also in the context of radioactive waste repositories. Also, other studies exist that use surrogate modeling as part of the performance assessment of radioactive waste repositories [23, 24, 25]. While we show that the metamodeling tools in this study are adequate for the given problem, we are aware at the same time that future studies potentially require the inclusion of a number of additional parameters that might show the limits of the classical tools used herein. Therefore, we would also like to stress that future workflow advancements will profit from an immense amount of work done over recent years to improve these models namely by implementing better learning strategies using adaptive methods [26, 27, 28, 29] and to better deal with high-dimensional problems by building the proxy in active subspaces [30, 31, 32]. However, the main feature of the used workflow is the combination of a surrogate model with prior parameter uncertainties that are further reduced by comparing the model outcome to the experimentally measured responses which is a new approach for dealing with the reliability of a disposal site for high-level radioactive waste. Such a history matching approach is very similar to a Bayesian inference approach (see e.g., [33]), however, some differences exist as pointed out below. Combining both, metamodeling and Bayesian inference for studying the reliability of complex technological systems is still a field of active research in the reliability engineering community [34, 35, 36] as pointed out by [37] recently:

Addressing modeling complexity remains a significant challenge for Bayesian inference applications because integrating metamodeling techniques is not trivial. The challenge here is to establish a fully automated integration process that addresses different degrees of competency for the end user and a wide range of application problems with a certain degree of robustness.

Our history matching workflow employs a 'brute-force' direct Monte Carlo sampling method to overcome the main flaws of widely used Markov chain (MCMC) and minimization techniques for parameter fitting on response surfaces with multiple minima. This reduces the issue of workflow-robustness mainly to the issue of meaningful parameter identification and a proxy error given the appropriateness of the models being used. The proxy error as well as poor chain convergence in MCMC techniques are typical reasons why true posteriors in a Bayesian sense won't be generated.

In general, uncertainty quantification and sensitivity analysis is not just a final step of the safety assessment, but can be understood as part of an iterative process of model development and calibration [38]. We see our work as a contribution subscribing to this idea. So, despite neglecting certain aspects and effects (uncertainties stemming from the geometry, bentonite, excavation/construction, extended sets of physical effects, etc.), we aim to reduce the entire uncertainty space to some extent in the present work, leaving room for further studies to build upon this work.

The paper is structured as follows: We have divided the entire article into two parts, deterministic modeling and statistical modeling. The deterministic section consists of theoretical derivation of the governing equations, computational details, and model preparation of the initial state. The statistical modeling section comprises the descrip-

tion of the methodology, subsections for software and data, and finally the conclusions' subsection.

2. Deterministic modeling

2.1. Theoretical formulation

The evolution of temperature, pressure and displacement of the host rock around a heat-emitting canister emplaced in an underground rock laboratory can be best described in terms of a coupled thermo-hydro-mechanical (two-phase flow) and component model, like the TH²M model (see Tab. 1 for a model overview) introduced by Grunwald et al. (2022) [39]. Pitz et al. (2023) [40] showed that temperature and pore pressure can be very well-matched by applying Richards' assumption, as long as excess gas pressures can dissipate quickly and dissolution of gas in the liquid phase can be neglected. Under this assumption, the governing equations simplify to the following set consisting of an energy balance (1), a mass balance (3) and the momentum balance equation (4) [41].

$$0 = (\rho c_p)_{\text{eff}} \frac{\partial T}{\partial t} + L_0 \frac{\partial \theta_{\text{vap}}}{\partial t} - \text{div}(\mathbf{K}_{\text{eff}} \text{grad} T) + \text{div} \left(\frac{L_0 \mathbf{q}_v}{\rho_{\text{LR}}} \right) \quad (1)$$

$$+ \text{grad} T \cdot (c_{pL} \rho_{\text{LR}} \tilde{\mathbf{w}}_{\text{LS}} + c_{p, \text{vap}} \rho_{\text{vap}} \tilde{\mathbf{w}}_{\text{vapS}})$$

where L_0 represents the volumetric latent heat of water vaporization:

$$L_0 = \rho_{\text{LR}} \left[2.501 \times 10^6 \text{ J kg}^{-1} - 2369.2 \text{ J kg}^{-1} \left(\frac{T}{\text{K}} - 273.15 \right) \right] \quad (2)$$

$$0 = (\rho_{\text{LR}} - \rho_{\text{vap}}^{\text{W}}) [\phi + p_{\text{LR}} S_{\text{L}} (\alpha_{\text{B}} - \phi) \beta_{p, \text{SR}}] \frac{dS_{\text{L}}}{dt} + \rho_{\text{LR}} S_{\text{L}} [\phi \beta_{p, \text{LR}} + S_{\text{L}} (\alpha_{\text{B}} - \phi) \beta_{p, \text{SR}}] \frac{\partial p_{\text{LR}}}{\partial t}$$

$$+ \phi \left[(1 - S_{\text{L}}) \frac{\partial \rho_{\text{vap}}^{\text{W}}}{\partial p_{\text{LR}}} \right] \frac{\partial p_{\text{LR}}}{\partial t} + \text{div}(\rho_{\text{LR}} \tilde{\mathbf{w}}_{\text{LS}} + \rho_{\text{vap}}^{\text{W}} \tilde{\mathbf{w}}_{\text{vapS}}) + S_{\text{L}} \rho_{\text{LR}} \alpha_{\text{B}} \text{div} \frac{\partial u}{\partial t}$$

$$+ \left[\phi (1 - S_{\text{L}}) \frac{\partial \rho_{\text{vap}}^{\text{W}}}{\partial T} - \rho_{\text{LR}} S_{\text{L}} [\phi \beta_{T, \text{LR}} + (\alpha_{\text{B}} - \phi) \text{tr}(\boldsymbol{\alpha}_{T, \text{SR}})] \right] \frac{\partial T}{\partial t} \quad (3)$$

$$\mathbf{0} = \text{div}(\boldsymbol{\sigma}' - \alpha_{\text{B}} \chi p_{\text{LR}} \mathbf{I}) + [(1 - \phi) \rho_{\text{SR}} + S_{\text{L}} \phi \rho_{\text{LR}}] \mathbf{b} \quad (4)$$

The effective stress rate $\boldsymbol{\sigma}'$ is given by a generalized form of Hooke's law to account for thermal effects:

$$\boldsymbol{\sigma}' = \mathcal{C}: \left(\dot{\boldsymbol{\epsilon}} - \boldsymbol{\alpha}_{T, \text{SR}} \dot{T} \right) \quad (5)$$

Where the double dot product ':' denotes the contraction of the last two indices of the elasticity tensor \mathcal{C} with the two indices of the generalized strain rate tensor. Throughout the paper, we will focus on temperature and pressure predictions in comparison to experimental results. To speed up the scoping simulations, we simplify the model further, following the approach published by the authors in an earlier contribution [42] for saturated media. Using this approach, the anisotropic elastic behavior of the solid is

reduced to a scalar (isotropic) expression that can effectively represent the impact of the volumetric deformation on the pressure changes. This simplification reduces the number of needed primary variables from five (T , p , u_x , u_y and u_z) to just two (T and p), such that the momentum balance is dropped from the set of PDEs and the last term of Eq. (3) can be written in terms of temperature and pressure derivatives:

$$\phi \left((\rho_{\text{LR}} - \rho_{\text{vap}}^{\text{W}}) \frac{\partial S_{\text{L}}}{\partial p_{\text{LR}}} + (1 - S_{\text{L}}) \frac{\partial \rho_{\text{vap}}^{\text{W}}}{\partial p_{\text{LR}}} + S_{\text{L}} (\rho_{\text{LR}} \beta_p + \chi_{\text{S}} \alpha_{\text{B}}^2 \beta^\sigma) \right) \frac{\partial p_{\text{LR}}}{\partial t} \quad (6)$$

$$+ \left(\phi(1 - S_{\text{L}}) \frac{\partial \rho_{\text{vap}}^{\text{W}}}{\partial T} - S_{\text{L}} \rho_{\text{LR}} (\phi \beta_{T,\text{LR}} + (\alpha_{\text{B}} - \phi) \text{tr}(\boldsymbol{\alpha}_{T,\text{SR}}) - \alpha_T^\sigma) \right) \frac{\partial T}{\partial t} \quad (7)$$

$$+ \text{div}(\mathbf{q}_w + \mathbf{q}_v) = Q_H . \quad (8)$$

Here, α^σ and β^σ are terms describing the thermal expansivity and storage contributions coming from the elastic behavior of the solid. In [42], two models were derived corresponding to two limiting cases in terms of stress and kinematics. One is based on the assumption of hydrostatic stress conditions, and the other only allows uniaxial deformation. Although both conditions are not met in general cases, the latter model has been shown to give relatively good pressure prediction for the given test case under saturated conditions. Even for anisotropic materials, good agreement could be found if the material axis is rotated along the main diagonal of the material coordinate system such that all elastic constants enter into the correction factor in an average sense. The coefficients for the uniaxial model were calculated according to Eq. (27) in [42]:

$$\beta^{\sigma=\text{uniaxial}} = \frac{1}{C_{3333}} \quad (9)$$

$$\alpha^{\sigma=\text{uniaxial}} = \alpha_{\text{B}} \left(\alpha_{TS;z} + \alpha_{TS;x} \frac{C_{1133}}{C_{3333}} + \alpha_{TS;y} \frac{C_{2233}}{C_{3333}} \right) \quad (10)$$

C_{ijkl} corresponds to the entries of the elasticity tensor. The given equations are exact if only uniaxial deformation along the third direction (regarding the material coordinate system) occurs. Generally, there is no preferential direction. In the given problem, the transverse-isotropic elasticity tensor of clay was rotated along the main diagonal of the material coordinate system (i.e. normal 111) for an effective averaging of all three principal directions. Eqn. (6) provides a straightforward generalization of the equations presented in [42] to unsaturated conditions based on Richard's approximation. The energy balance equation is unaffected by these changes.

These simplifications allow us to evaluate the THM model with the efficiency of a hydro-thermal model under a moderate loss in accuracy only. A brief overview of models and their abbreviations used throughout this publication is given in Tab. 1.

2.2. Computational details

All simulations were performed using the OpenGeoSys numerical simulator (OGS-6)¹ [43, 44], an open-source code for modeling thermo-hydro-mechanical-chemical processes

¹<https://www.opengeosys.org/>

Table 1: Explanation of numerical models and their abbreviations used in this work.

Model abbreviation	Description
THM	This abbreviation refers to the umbrella term of thermo-hydro-mechanical coupled models in this paper. It includes saturated uniphase models as well multiphase and multi-component models.
TH ² M	The model featuring non-isothermal two-phase (liquid and gaseous phase) two-component (water and non-aqueous component) flow with linear poro-elasticity. This model is the most comprehensive model described herein.
TRM	The non-isothermal Richards mechanics implementation based on the non-isothermal Richards flow equation coupled to linear poro-elasticity. It includes a vapor diffusion model and can be used to model both saturated and unsaturated problems.
TRuni	The same as TRM without the mechanical coupling. It contains correction terms in the mass balance equation accounting for the thermo-mechanical coupling.

in porous and fractured media. The three-dimensional hexahedral model domain with dimensions 100 m x 80 m] x 100 m is intended to capture the near-field thermo-hydro-mechanical behavior around a tunnel filled with high-level radioactive waste. The tunnel is aligned along the y -axis and starts at $y = 0$ and has a length of 50 m. A detailed description of the implementation of the experiment can be found elsewhere [20]. A two-dimensional vertical cut through the mesh is depicted in Fig. 1. The model domain was subdivided into 164 471 linear hexahedral elements. The sampling runs were conducted with an assembly based on the non-isothermal flow equations with simplified elasticity corrections (later referred to as TRuni) as given by Eqs. (1) and (6). For verification, additional runs were conducted based on the mechanically fully coupled model (later referred to as TRM) as given in Eqs. (1) to (4). Both are assembled and solved using a monolithic scheme with an iteration-based adaptive time stepper over a simulation time of $t = 3.83$ a. Nonlinearities are resolved with an iterative-incremental Newton-Raphson solution strategy, while time integration is fully implicit based on a backward Euler scheme. The relative tolerances for the nonlinear solver were set to 10^{-6} for all primary variables. The linear system was solved using the Eigen library’s iterative solvers IDRSTABL [45, 46] for TRM and BiCGSTAB [47] for TRuni, both with an error tolerance of 10^{-12} and using the Jacobi preconditioner.

2.3. Model preparation

2.3.1. Equilibration

Prior to excavation, we prepared an equilibrated initial state by setting the displacement everywhere to zero and applying a linear temperature and pressure profile ($T_{\max} = T(z = -50 \text{ m}) = 293.15 \text{ K}$, $T_{\min} = T(z = 50 \text{ m}) = 287.15 \text{ K}$, $p_{\max} = p(z = -50 \text{ m}) = 2.5 \text{ MPa}$, $p_{\min} = p(z = 50 \text{ m}) = 1.5 \text{ MPa}$). The initial effective stress was computed from the given total stress and pressure profile using the relation $\sigma'_{ij} = \sigma_{ij} + \alpha_B p \delta_{ij}$.

Table 2: Van Genuchten parameters used for bentonite

Material	S_0	ϕ	S_{res}	S_{max}	m	p_b
bentonite pellets (GBM)	0.56	0.331	0.01	1	0.45	18 MPa
bentonite blocks	0.19	0.331	0.01	1	0.45	30 MPa

The numerical values were taken from ranges given in the literature [48, 49]:

$$\sigma'_{xx} = -4.5 \text{ MPa} + 0.85 \times (2.0 - 0.01 \text{ m}^{-1}z) \text{ MPa} \quad (11)$$

$$\sigma'_{yy} = -2.5 \text{ MPa} + 0.85 \times (2.0 - 0.01 \text{ m}^{-1}z) \text{ MPa} \quad (12)$$

$$\sigma'_{zz} = -6 \text{ MPa} + 0.85 \times (2.0 - 0.01 \text{ m}^{-1}z) \text{ MPa} \quad (13)$$

$$\sigma'_{ij}|_{i \neq j} = 0 \text{ MPa} \quad (14)$$

An equilibration step was skipped, as the given conditions satisfy the momentum balance equation $\partial_j(\sigma'_{ij} - \alpha_B p \delta_{ij}) + \rho g_i = 0$. All remaining out-of-balance forces stemming from numerical inaccuracies are discarded when assembling the discrete force balance. As boundary conditions, we set all surface-normal displacements to zero along all outer boundaries. Additionally, Dirichlet conditions were also applied at $z = -50 \text{ m}$ as well as at $z = 50 \text{ m}$ setting temperature and pressure to their corresponding maximal and minimal values compatible with geothermal and hydrostatic gradients. All remaining boundaries were taken as free boundaries (homogeneous Neumann type).

2.3.2. Excavation and application of shotcrete

The excavation was performed over a time span of 181 days for the cavern and 91 days for the tunnel. For this purpose, all material domains inside the tunnel were gradually deactivated, beginning with the cavern entry and ending at $y = 50 \text{ m}$. The pressure along the newly generated boundary was set to a value that corresponds to 90% relative humidity. The excavation phase was followed by a shotcreting and ventilation phase in which the shotcrete was applied with initial and boundary conditions corresponding to 90% relative humidity followed by linear drying to 60% relative humidity over a time span of 336 (FE-cavern) and 778 days (FE-tunnel).

2.4. Back-filling

The entire tunnel was then closed with a porous concrete wall, a bentonite block wall, emplaced with heaters, concrete and bentonite pedestals and backfilled granular bentonite and finally sealed with a plug. Initial pressures based on water saturation were calculated from the van Genuchten curves of the respective materials. The numerical values of the water content and the van Genuchten parameters used in our model are given in Tab. 2.

The reactivation steps of the material groups are summarized in Tab. 3. The temperature of all back-filled materials were set to 291.15 K.

2.5. Heating phase

After back-filling, the cross-sectional 2D slice of the entire model domain through the tunnel looks as depicted in Fig. 1. The heaters, modeled by applying thermal properties

Table 3: Durations used for backfilling.

step	section on y-axis	duration
porous concrete wall + bentonite block wall	59.8-54.4	30 d
heater 1 + GBM	59.8-42.2	32 d
heater 2 + GBM	59.8-33.6	61 d
heater 3 + GBM	59.8-25	31 d
plug + completion	59.8-0	14 d

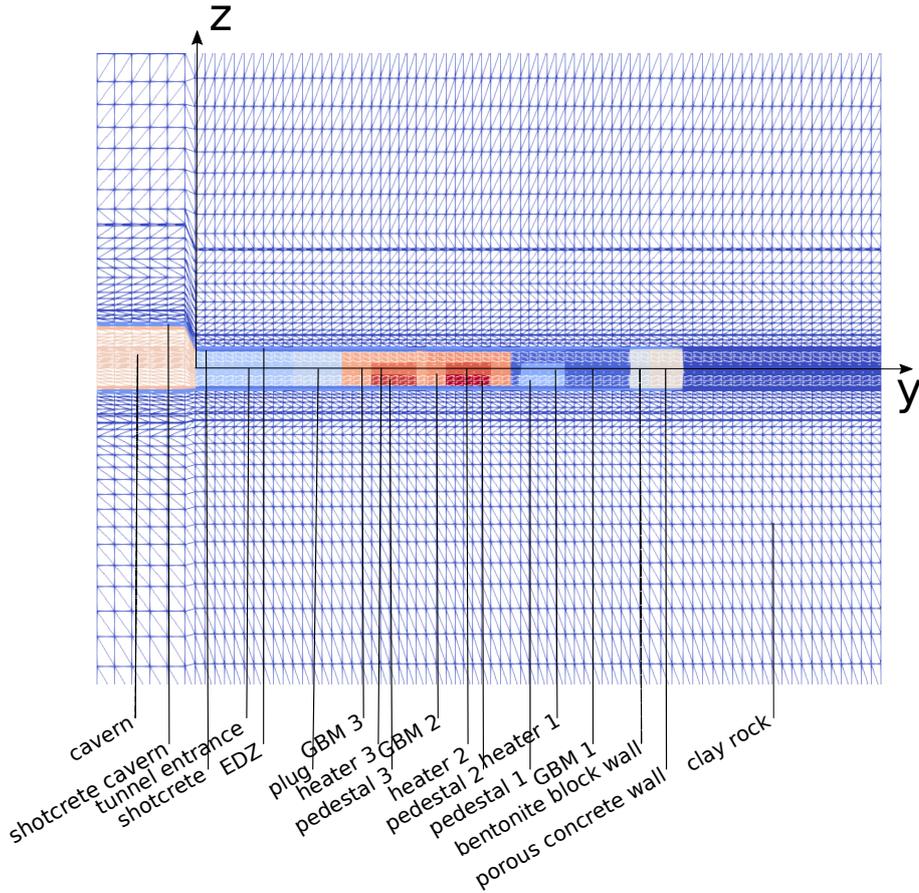


Figure 1: Vertical-cross section of the 3D model at $x = 0$ of the mesh sketching the color-coded material domains.

of steel lying on compacted bentonite blocks (pedestals) are surrounded by GBM pellets as filling material. The tunnel consists of a 20 cm excavation damage zone with reduced clay parameter values. The surface was backed by a 20 cm layer of shotcrete. The heating profile with their exact dates can be found elsewhere [20]. First, heater 1 (the one that was back-filled first) was turned on at a power of 500 W. One month later, the heating power was increased to 1000 W. Another month later, all three heaters were turned on

at a power of 1350 W. Results for specific observation points for the full TRM model and the TRuni model together with their corresponding experimental results can be found in subsection 3.4.5.

3. Statistical modeling

3.1. Methodology

This paper follows the approach introduced out in [8]. A sketch of the workflow for uncertainty quantification is depicted in Fig. 2. The first step requires the aggregation of experimental material parameters ①. Detailed knowledge of spatial variations is advantageous, as well as data on parameter distributions, as they reduce the uncertainty and the efficiency of the workflow significantly. However, the data provided by technical reports provides in most cases only min and max values. In the second step, the objective function needs to be defined ②. In the present work, we will use a history match error defined as an integral measure over all time steps for all response variables and observation points. For the history match filtering procedure, we will use an accumulated normalized error over all six observation points for both response variables p and T . In the third step, parameter screening is conducted to identify heavy-hitters, i.e. parameters that contribute most to the results, that will be used in further analysis ③. For this purpose, different screening designs can be used. We restrict our screening procedure to one-variable-at-a-time (OVAT) and Plackett-Burman designs. Using the identified heavy-hitters, a space-filling design like Latin-hypercube sampling can be used to build a proxy for the objective functions ④. Typical proxies include multidimensional polynomials, neural networks, or Gaussian process regression. In our study, we use the latter. Having built the proxy, it can be deployed for cheap evaluation either as part of a global sensitivity analysis ⑩, see also [50], or to perform a direct Monte Carlo sampling ⑤ to explore the entire parameter uncertainty space. Although considered a 'brute-force' method, it is a very effective way to perform a model calibration, regardless of multiple local minima that may be present in the model prediction. The results of the sampling can be used to apply a filter ⑥ to match results to the experimental response data allowing for parameter estimation or runs of the full deterministic model ⑦ for confirmation, further evaluation ⑨ e.g. to analyze safety measures like integrity criteria in terms of cumulative distribution functions or to perform a forecast to get projections for experimentally available data.

3.2. Software

The entire workflow was implemented in Python and uses ogs6py [51] and VTUinterface [52] for OGS pre- and post-processing [21]. The experimental designs were created using the pyDOE2² library, while GPy [53] was used for proxy modeling. The global sensitivity analysis was done using the package SALib³. The t-test was conducted with the help of the statsmodels package [54].

²<https://pypi.org/project/pyDOE2/>

³<https://salib.readthedocs.io/>

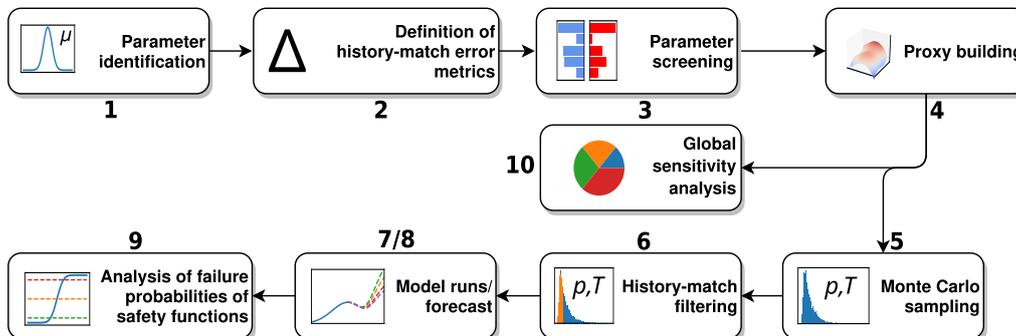


Figure 2: Schematic sketch of the workflow.

3.3. Data

The material properties for the Opalinus clay were mainly compiled from swisstopo’s database and reports [55] and [56]. In Tab. 4, we provide the used low/best/high values. If data was present in both resources, we took the min/max values resulting in the greater interval. In this study, we assume a uniform/loguniform distribution for all parameters, as in many cases no information on probability density functions could be obtained. In most cases, the low/mid/max values correspond to the natural variability or uncertainty in measurements of the properties, while for other parameters like c_s the resulting variability also arises from different temperature regimes ($T = 20^\circ\text{C}$: $c_s = (1040 \pm 170) \text{ J kg}^{-1}\text{K}^{-1}$ and $T = 80^\circ\text{C}$: $c_s^p = (1380 \pm 10) \text{ J kg}^{-1}\text{K}^{-1}$). Parameters used for other material groups can be found in Section Appendix B. It has been discussed elsewhere (e.g., [57, 58, 59]) that variations of the thermodynamic fluid properties of water might play a significant role. Therefore, we used the IAPWS water model for the viscosity and fluid density, resulting in non-constant water thermal expansivity and compressibility [60].

3.4. Results and discussion

The analysis was conducted at six observations points listed in Tab. 5:

Keeping in mind that the heaters were placed along y -axis, this implies that a cross-section between the first and the second heater is chosen. The points 1-3 are aligned in the bedding plane, the points 4-6 in a direction normal to the bedding plane. All points lie in the Opalinus clay and span a distance between 3.5 and 13 m from the tunnel wall and correspond to actual sensors that were installed in boreholes drilled from within the FE-Tunnel during the instrumentation phase. An illustration can be found in Fig. 8 of the implementation paper by Müller et al. [20]. The history match error metrics are then defined for each response quantity by

$$\text{Ehm} = \sqrt{\frac{1}{n} \sum_{i=1}^n (d_i^{\text{obs}} - d_i^{\text{sim}})^2}. \quad (15)$$

Here, n is the number of measured instances (data points) in time. Different derived and combined metrics for the various quantities and observation points are discussed later in the manuscript.

Table 4: Used parameters for clay

Parameter	symbol	low	best	high	unit	distribution
Young's modulus	E_1 [56]	$2.2 \cdot 10^9$	$7.2 \cdot 10^9$	$8.1 \cdot 10^9$	Pa	uniform
Young's modulus	E_3 [56]	$6.1 \cdot 10^8$	$2.8 \cdot 10^9$	$3.5 \cdot 10^9$	Pa	uniform
Young's modulus EDZ	E_1 [56]	$0.2 \cdot 10^9$	$2.16 \cdot 10^9$	$4.86 \cdot 10^9$	Pa	uniform
Young's modulus EDZ	E_3 [56]	$0.61 \cdot 10^8$	$0.84 \cdot 10^9$	$2.1 \cdot 10^9$	Pa	uniform
Shear modulus	G_{13} [56]	$1.2 \cdot 10^9$	$2.4 \cdot 10^9$	$3.7 \cdot 10^9$	Pa	uniform
Poisson's ratio	ν_{12} [56]	0.28	0.33	0.38	-	-
Poisson's ratio	ν_{13} [56]	0.16	0.24	0.32	-	-
lin. thermal expansion coefficient of the solid	α_s^\perp [55]	$0.8 \cdot 10^{-5}$	$1.9 \cdot 10^{-5}$	$3 \cdot 10^{-5}$	K^{-1}	uniform
lin. thermal expansion coefficient of the solid	α_s^\parallel [55]	$1.0 \cdot 10^{-5}$	$1.4 \cdot 10^{-5}$	$1.8 \cdot 10^{-5}$	K^{-1}	uniform
Porosity	ϕ [56]	0.13	0.16	0.21	-	-
Porosity EDZ	ϕ [56]	0.13	0.16	0.21	-	-
Solid grain density	ρ_s [55]	2680.0	2700.0	2720.0	$kg\ m^{-3}$	uniform
Specific isobaric heat capacity of the solid	c_s [55]	870.0	1040.0	1390.0	$J\ kg^{-1}K^{-1}$	uniform
Heat conductivity of the solid	K_s^\perp [56]	1.0	1.2	3.1	$W\ m^{-1}K^{-1}$	uniform
Heat conductivity of the solid	K_s^\parallel [56]	1.0	2.1	3.1	$W\ m^{-1}K^{-1}$	uniform
Intrinsic permeability	k_s^\perp [55]	$0.5 \cdot 10^{-20}$	$0.6 \cdot 10^{-20}$	$0.7 \cdot 10^{-20}$	m^2	loguniform
Intrinsic permeability	k_s^\parallel [55]	$1.3 \cdot 10^{-20}$	$1.5 \cdot 10^{-20}$	$2 \cdot 10^{-20}$	m^2	loguniform
Intrinsic permeability EDZ	k_s^\perp [55]	$1.2 \cdot 10^{-19}$	$1.0 \cdot 10^{-18}$	$1.2 \cdot 10^{-17}$	m^2	loguniform
Intrinsic permeability EDZ	k_s^\parallel [55]	$1.2 \cdot 10^{-19}$	$1 \cdot 10^{-18}$	$2 \cdot 10^{-17}$	m^2	loguniform
EDZ thickness	l_{EDZ}	0	0.5	1	m	uniform

Table 5: Observation points

Observation point	Coordinates (x,y,z) in meter
pt1	(4.593, 32.966, 2.517)
pt2	(7.142, 32.588, 4.287)
pt3	(11.66, 32.765, 8.508)
pt4	(-3.602, 32.598, 3.714)
pt5	(-4.998, 32.612, 6.282)
pt6	(-8.956, 32.638, 11.702)

3.4.1. Parameter Screening (Step 3)

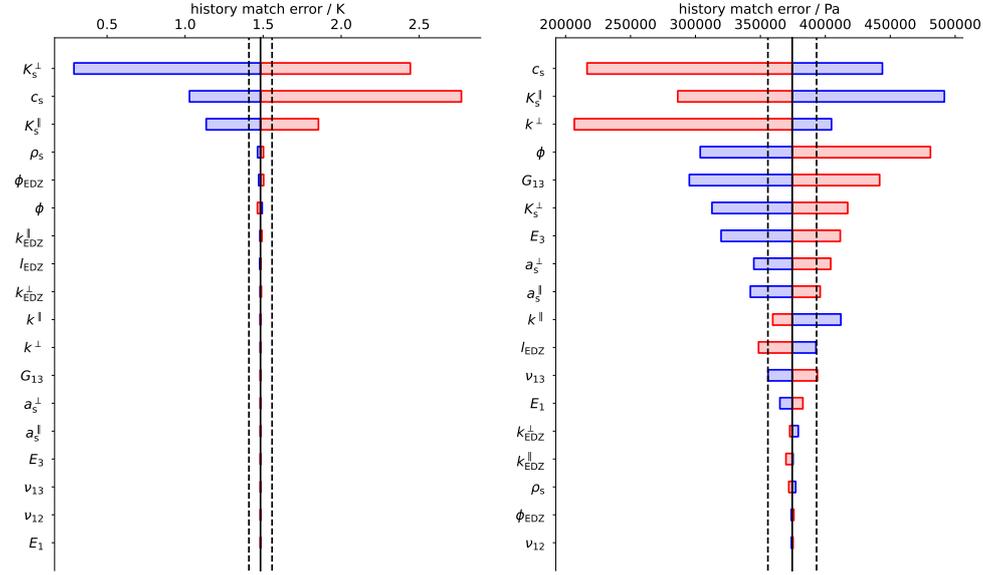
The purpose of the initial screening step (workflow step ③ in Fig. 2) is to identify relevant parameter uncertainties and parameters whose variability has no significant impact on the output response. This behavior is not just dependent on the considered value range and the considered quantity itself, it is also a function of space and time (see [50]). As the purpose is to reduce the computational complexity at later workflow steps, no thorough investigation should be done at the screening stage. Therefore, we used relatively simple (i.e. small) designs for the parameter screening: a one variable at a time (OVAT) screening for non-interfered main effects and a folded Plackett-Burman design for creating Pareto charts to obtain a complementary picture based on main effects. Both designs were evaluated for every response variable and at all six observation points. For the OVAT design, we took the 'best' values (see table 4) as reference and changed all parameters separately to their min/max values, whereas the Plackett-Burman design uses the min/max values only. The results of both designs were compared for pt1 and

temperature (Figs. 3a and 3c) as well as for pt2 and pressure (Figs. 3b and 3d). Both combinations were selected because in both cases the t-test revealed significant results. For other combinations, the null hypothesis from the F-test couldn't be rejected, so their results lack significance (see [61] for more details on t-test and hypothesis testing in the context of regression analysis). This confirms the experiences we had with a simpler model of a heat source in a saturated, homogeneous medium [8]. We conclude that even the folded design is not suitable for parameter screening with the history match error and these kinds of thermo-hydro-mechanical coupled models. Nevertheless, the results clearly show the impact of different screening methods and also give rise to the conclusion that the sensitivity is affected by the choice of observation points (compare the spatio-temporal sensitivity from [50]). Although not surprising, it is an effect one should keep in mind already during the initial screening. Although there are well-known downsides of OVAT-screening designs such as the possibility that potential effects could be overlooked, we conclude that the combination of the modelers experience and his/her insight into the physical model together with testing of multiple response variables at different locations should be sufficient to select all relevant parameters for proxy building. Other (OVAT) screening designs, like the Morris method ([62]) require multiple parameter sweeps to cover the sampling space as only small parameter changes are considered.

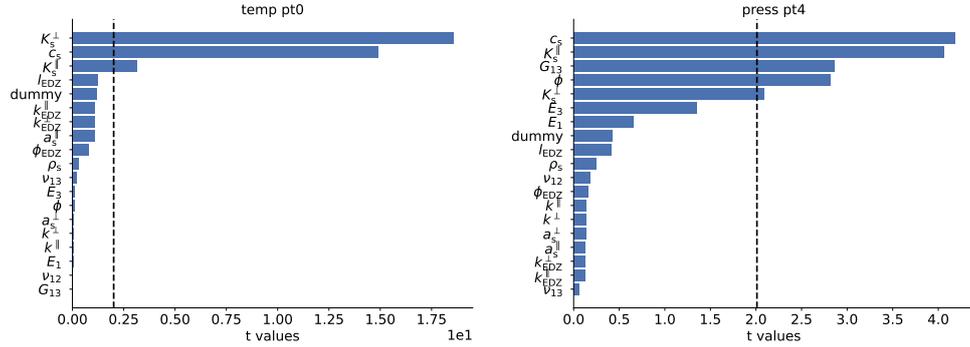
Combining both screening methods, we were able to reduce the uncertainty space from originally 20 parameters to the following 14 parameters which we consider to be significant for further investigation: $K_s^\perp, K_s^\parallel, c_s, G_{13}, E_3, a_s^\parallel, a_s^\perp, k^\parallel, k^\perp, \phi, l_{\text{EDZ}}, \phi_{\text{EDZ}}, k_{\text{EDZ}}^\parallel, k_{\text{EDZ}}^\perp$.

3.4.2. Proxy building (Step 4)

Based on the 14 heavy hitters found in the previous step, we used a Latin-hypercube sampling scheme with 500 sampling points on which the proxy is defined workflow step ④. The size of the sampling scheme was chosen from the tradeoff between the required computational effort and proxy error. For this purpose, we calculated the coefficient of determination R^2 as well as the root-mean-square error (RMSE) for each response variable and at each observation point. The results are summarized in Tab. 6. For this purpose, the LH sample was split up into a training set with 450 points and a testing set with 50 points. Both measures were calculated using ten proxy rebuilds by selecting every time new training and testing points. The final values were taken from the minimum (R^2) and maximum (RMSE) value of one of the rebuilds for each point and variable. For all runs, the deterministic TRuni forward model was used to obtain temperature and pressure responses. In Figs. 4a-4f we provide the time series output of all simulated curves at all observation points for both response variables to display the full prior parameter uncertainty. Additionally, the experimental data are displayed. What can be seen already at this stage is that the qualitative behavior of the experimental curves is mostly met by the modeling curves, except for the pressure dip of the most distant observation point. However, it is not clear whether the optimization regarding one variable at one particular point contradicts the optimization with regard to another point or variable, especially because the experimental curves are, for some cases on opposing ends of the sampled response spectrum. However, this will be discussed in a thorough analysis later. For all observation points and both response variables, we normalized the input data using the mean and standard deviation, as the numerical values of different parameters are in very



(a) OVAT analysis of the temperature at pt1. Units are in K (b) OVAT analysis of the pressure at pt5. Units are in Pa.



(c) T-test for linear effects of the temperature at pt1 using a folded Plackett-Burman design (d) T-test for linear effects of the pressure at pt5 using a folded Plackett-Burman design

Figure 3: Sensitivity screening with the OVAT method and a t-test of all included THM clay parameters.

different ranges. The normalized input was used to build a proxy of the history match error based on Eq. (15) using Gaussian process regression. For each combination of observation point and variable the prediction of the history match error was taken from the conditional mean of the Gaussian process. This was done with the Python package GPy using a Matern52 kernel. The Gaussian process uses a constant mean function and a separate length scale for each parameter. The hyperparameters were optimized using a scaled conjugated gradient algorithm as implemented in GPy.

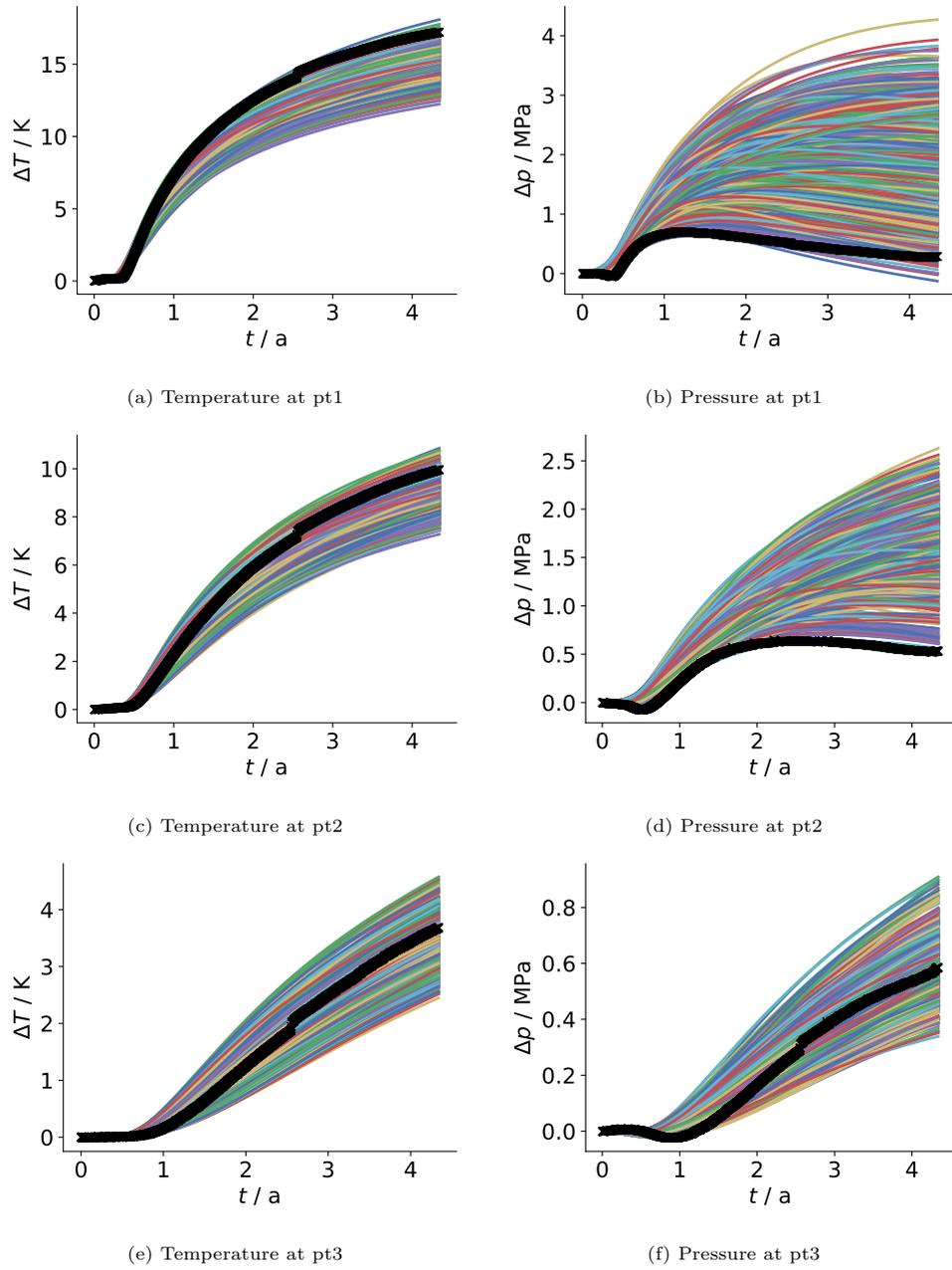


Figure 4: Ensemble run temperature and pressure curves generated from LHS design based on prior distributions for the first three observation points. The measured response is painted black. Colored curves present modeled responses of TRuni model for different sets of parameters.

Table 6: Proxy Quality measures.

	Temperature		Pressure	
	R ²	RMSE	R ²	RMSE
pt1	0.999	0.01 K	0.991	51513 Pa
pt2	0.999	0.01 K	0.996	7398 Pa
pt3	0.998	0.00 K	0.896	8753 Pa
pt4	0.997	0.01 K	0.943	47973 Pa
pt5	0.993	0.01 K	0.988	10277 Pa
pt6	0.995	0.01 K	0.993	1060 Pa

3.4.3. Monte Carlo (MC) sampling (Step 5)

Given the (uniform) input parameter distributions (Tab. 4), the entire parameter space was sampled using a direct Monte Carlo sampling strategy with a size of $N = 10^6$ (workflow step ⑤ in Fig. 2). The proxy models optimized using the Latin-hypercube-training data set were evaluated on the entire sampling space, providing an estimation of the history match error for each set of parameters at each observation point for temperature and pressure. The result gives us a distribution of the history match error on the one hand and allows for solving the inverse problem (i.e. parameter estimation) on the other.

3.4.4. History Matching (Step 6)

The selection of a subsample of MC analysis results that is compatible with the experimental data set allows us to reduce the uncertainty of the problem further. This selection is done by filtering of the MC sample with history match criteria (workflow step ⑥ in Fig. 2). In general, the history match criteria are based on three sources of error. The proxy error E_{proxy} , the experimental error E_{exp} representing the uncertainty coming from measurements of temperature and pressure and the model error E_{model} . As no data are available for the experimental error and all curves are only analyzed in terms of their changes with respect to the first time step, we neglect the experimental error and try to get an estimation of the model error in the following analysis. The proxy error can be evaluated in terms of the RMSE of the proxy, $E_{\text{proxy}}^{\text{VAR,PT}} = 2\text{RMSE}^{\text{VAR,PT}}$ assuming a 95 % confidence interval. As it is assumed to be stochastic, we assume that the remaining systematic deviations constitute the model error, which we thereby estimate by the minimal history match error. In the following analysis, the unit afflicted history match errors are normalized by their maxima and analyzed by taking the sum over all points and response variables. Let $\{x_k\} \subset \mathcal{A}$ be a sample of the sampling space \mathcal{A} . Then, we have a prediction of the history match error for each variable and each observation point given by the proxy:

$$x_k \mapsto \text{Ehm}_k^{\text{VAR,PT}}, \text{VAR} \in \{T, p\}, \text{PT} \in \{\text{pt1}, \text{pt2}, \text{pt3}, \text{pt4}, \text{pt5}, \text{pt6}\} \quad (16)$$

To be able to combine errors for different observation points and for different variables, we define a normalized error by

$$\widehat{\text{Ehm}}_k^{\text{VAR,PT}} = \frac{\text{Ehm}_k^{\text{VAR,PT}}}{\max_k \left(\left\{ \text{Ehm}_k^{\text{VAR,PT}} \right\} \right)} \quad (17)$$

We also define condensed error metrics in the following way for the variables

$$\widehat{\text{Ehm}}_k^{\text{VAR}} = \sum_{PT} \widehat{\text{Ehm}}_k^{\text{VAR,PT}}, \quad (18)$$

and analogous for the points

$$\widehat{\text{Ehm}}_k^{\text{PT}} = \sum_{VAR} \widehat{\text{Ehm}}_k^{\text{VAR,PT}} \quad (19)$$

to allow for variable wise and point-wise analyses. Finally, we have a overall metric summed over all points and variables:

$$\widehat{\text{Ehm}}_k = \sum_{VAR} \widehat{\text{Ehm}}_k^{\text{VAR}}. \quad (20)$$

An analogous notation is also used for the proxy error. The model error estimations based on temperature matches and pressure matches separately can be found via the following subsamples that uses for each variable or for each point the corresponding minimum and considers all sampling points that are within the proxy error. So we obtain for each variable

$$\{x_k\}_{|\text{VAR}} := \left\{ x_k \mid \widehat{\text{Ehm}}^{\text{VAR}}(x_k) < \min_k \left(\widehat{\text{Ehm}}_k^{\text{VAR}} \right) + \widehat{\text{Eproxy}}^{\text{VAR}} \right\} \quad (21)$$

and similarly, for each point

$$\{x_k\}_{|\text{PT}} := \left\{ x_k \mid \widehat{\text{Ehm}}^{\text{PT}}(x_k) < \min_k \left(\widehat{\text{Ehm}}_k^{\text{PT}} \right) + \widehat{\text{Eproxy}}^{\text{PT}} \right\}. \quad (22)$$

The estimates for the model error can then be found by taking the minimum of the overall metric for each subsample:

$$\widehat{\text{Emodel}}_{\text{VAR}} = \min_k \left(\left\{ \widehat{\text{Ehm}}(x_k) \mid x_k \in \{x_k\}_{|\text{VAR}} \right\} \right) \quad (23)$$

As corresponding numerical values we find $\widehat{\text{Emodel}}_{|T} = 1.88$ and $\widehat{\text{Emodel}}_{|p} = 2.344$ keeping in mind that both values are normalized, and their theoretical maximum is $2 \times 6 = 12$ as the upper bound for each individual term is one.

With these relations, we can now define two thresholds for the history matching procedure:

$$\hat{\Theta}_{T/p} = \widehat{\text{Emodel}}_{|T/p} + \widehat{\text{Eproxy}} \quad (24)$$

In the upcoming analysis, we will apply both thresholds to investigate their influence on the distribution of modeling results, referred to as the 'posterior.' It is essential

to emphasize that the term 'posterior,' consistently used throughout this publication, is not a posterior in a strict Bayesian sense. This departure primarily arises from the introduction of extra uncertainties attributed to the proxy model, which is why we enclose the term in single quotation marks.

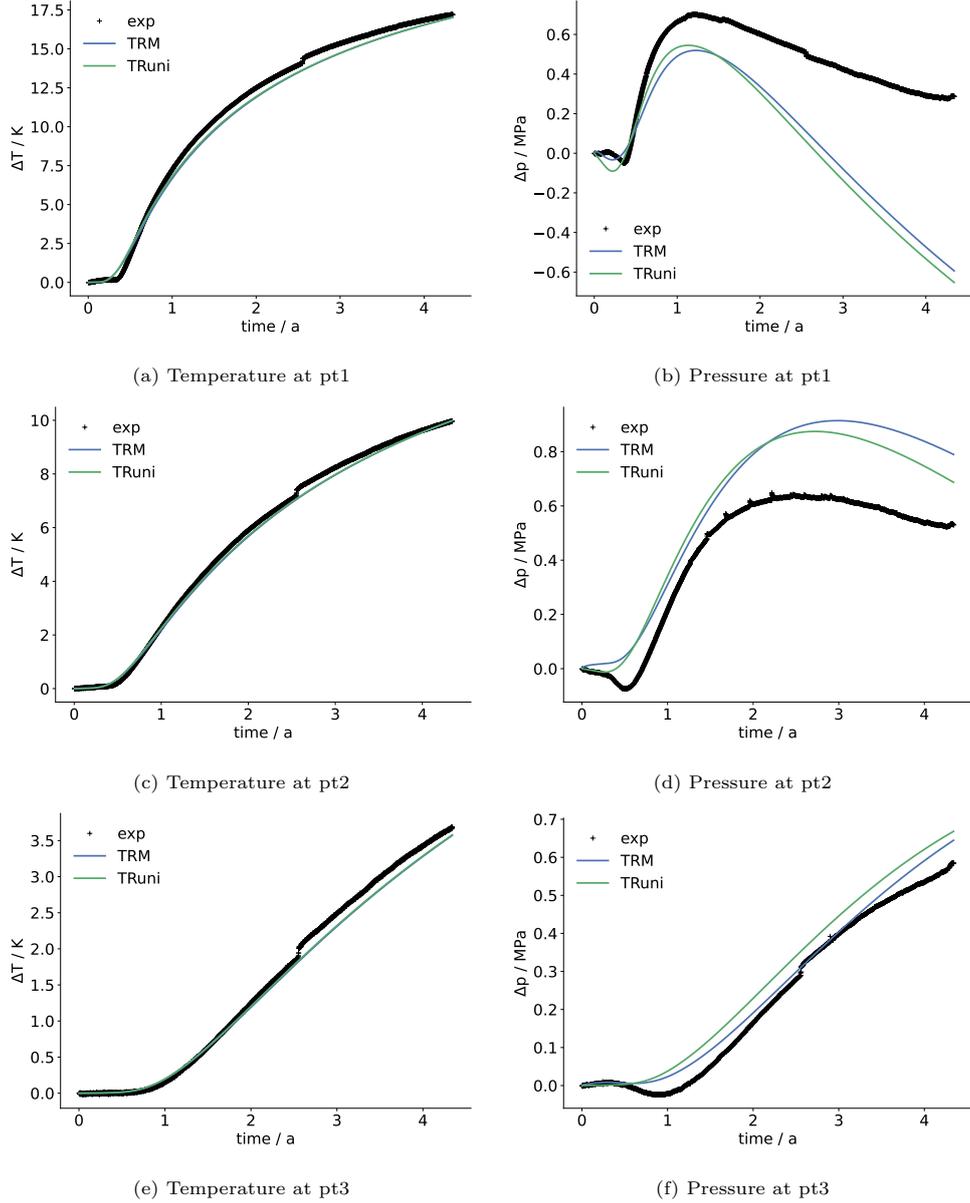


Figure 5: Temperature and pressure responses at the first three observation points: Best fit results. The measured curve is plotted alongside the TRUni and the fully coupled TRM response.

In Figs. 5a-5f, the experimental results for the first three observation points are depicted together with simulation results of the best overall match for a complete TRM model and the TR model with simplified HM-coupling based on the uniaxial strain approximation. The results for the remaining points are presented in the appendix (Fig. A.12). As already shown for saturated media, nearly perfect agreement between TRM and TRuni is found for the temperature, and also relatively good agreement for the pressure response. We deduce from the theoretical derivation of the TRuni model that similar agreement can be found also for other parameter combinations. Having shown the validity of the simplified forward model, we can now analyze the history matching capabilities. The smaller numerical value of the model error $\widehat{\text{Emodel}}|_T$ is reflected by the decent temperature match for all observation points. However, for pressure, we find somehow conflicting behavior: an underestimation of the pressure peak with an increased drainage at the closest point whereas the point in the middle exhibits an overestimation. The reason can be manifold: stemming from oversimplified models or inaccurate data of the bentonite and shotcrete, a greater extent of the excavation damaged zone and fracturing therein [63], the tight coupling of material interfaces to neglected effects of thermo-filtration and thermo-osmosis which are assumed to play a significant role in clay-rich formations under thermal gradients [64]. Additionally, the pressure dip that happens in the first year observed at pt1 and pt2 is weaker pronounced in the TRuni model possibly hinting to a reach in the HM coupling. For pt3 the pressure dip is not reflected at all in the models at least for the given parameter set.

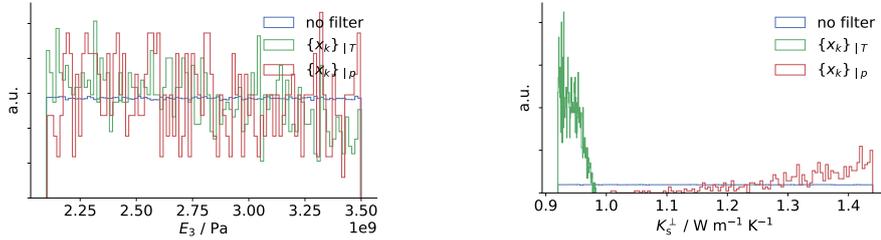


Figure 6: Parameter estimation for two selected parameters based on subsamples $\{x_k\} | T$ and $\{x_k\} | p$ compared to the prior parameter distribution. The counts are given in arbitrary units (a.u.).

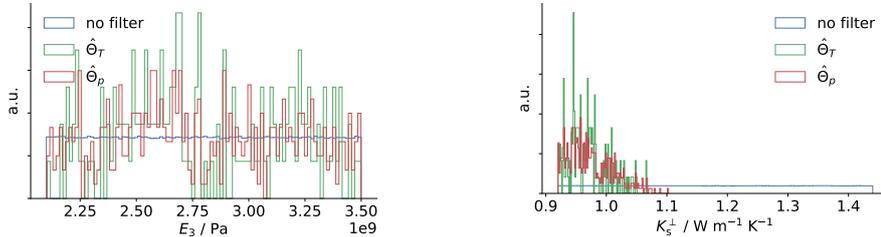


Figure 7: Parameter estimation for two selected parameters based on criteria thresholds $\hat{\Theta}_T$ and $\hat{\Theta}_p$ compared to the prior parameter distribution. The counts are given in arbitrary units (a.u.).

3.4.5. Parameter estimation

The subsamples, Eq. (21) obtained from applying the matching criteria, were analyzed in terms of their input parameters and compared to the full MC sample, Fig. 6. The results for the remaining parameters can be found in the appendix A.13. While some parameters, like the Young’s modulus, don’t show a selection effect, other parameters like K_s^\perp lead to an inconclusive selection when T and p are selected separately. This hints to greater model errors suspected for the pressurization in clay. To account for these model errors, we analyzed the parameter input based on the overall history match error and the thresholds given in Eq. 24. Two selected representative parameter distributions are plotted in Fig. 7. The remaining parameters can be found in A.15. As expected, we observed a very similar behavior in the Young’s modulus E_3 , a parameter that is less sensitive to the model outcome for most points according to the initial screening. However, for K_s^\perp both samples exhibit a clustering at lower values, very similar to the temperature-based subsample in Fig. 6 which is due to the fact that the best matches are dominated by the temperature match which is overall better, which can be observed in Fig. 8. This is due to the fact that temperature propagation is essentially reduced to a simple heat-conduction problem, whereas the physics of pressure propagation is significantly more complex.

In the next step (7), both parameter subsets were used to run the complete TRuni forward model. To limit the computational effort, we selected randomly 100 sampling points from both subsamples. We also calculated the RMSE values for all runs. The exact values are provided in the appendix (see A.7 and A.7). They agree mostly with the values already estimated from the latin-hypercube sampling plan, however some values tend to be a bit larger, which we attribute to the fact that the MC sampling covers a greater uncertainty space, i.e., the interpolation problem becomes an extrapolation problem for some sampling points. In Fig. 8a-8f we plotted the response by filling the space between the lowest and highest projection. Here, we see clearly the difference between the temperature-based and pressure-based threshold. While the pressure uncertainty width/range seems to be doubled for the first point, it is about the same range for the last point. This is also reflected in the fact that the qualitative behavior of the model for the last observation point agrees quite well, while for the point closest to the heater the model predicts a rapid pressure decrease, which is not found in the experiment.

3.5. Point-based analysis

Adding up the normalized history matches for temperature and pressure, but for each point separately, we find the best point-based matches by their minimum plus twice the RMSE of the proxy. This allows us to perform a parameter estimation for each individual point (Fig. 9). This analysis gives us the ability to identify contradictions in the model prediction. To focus on the effects, this analysis was restricted to the first three observation points only. The parameter that stands out most is the porosity, where lower values are preferred for pt5, whereas in contrast higher values are preferred for pt6 whereas pt4 does not exhibit a clear tendency. A similar behavior can also be found for the shear modulus, but contradicting tendencies can also be observed for other parameters as well.

The results of the corresponding full forward model runs are presented in Figs. 10a-10f. While all colors lie more or less above each other for the temperature, we see clear

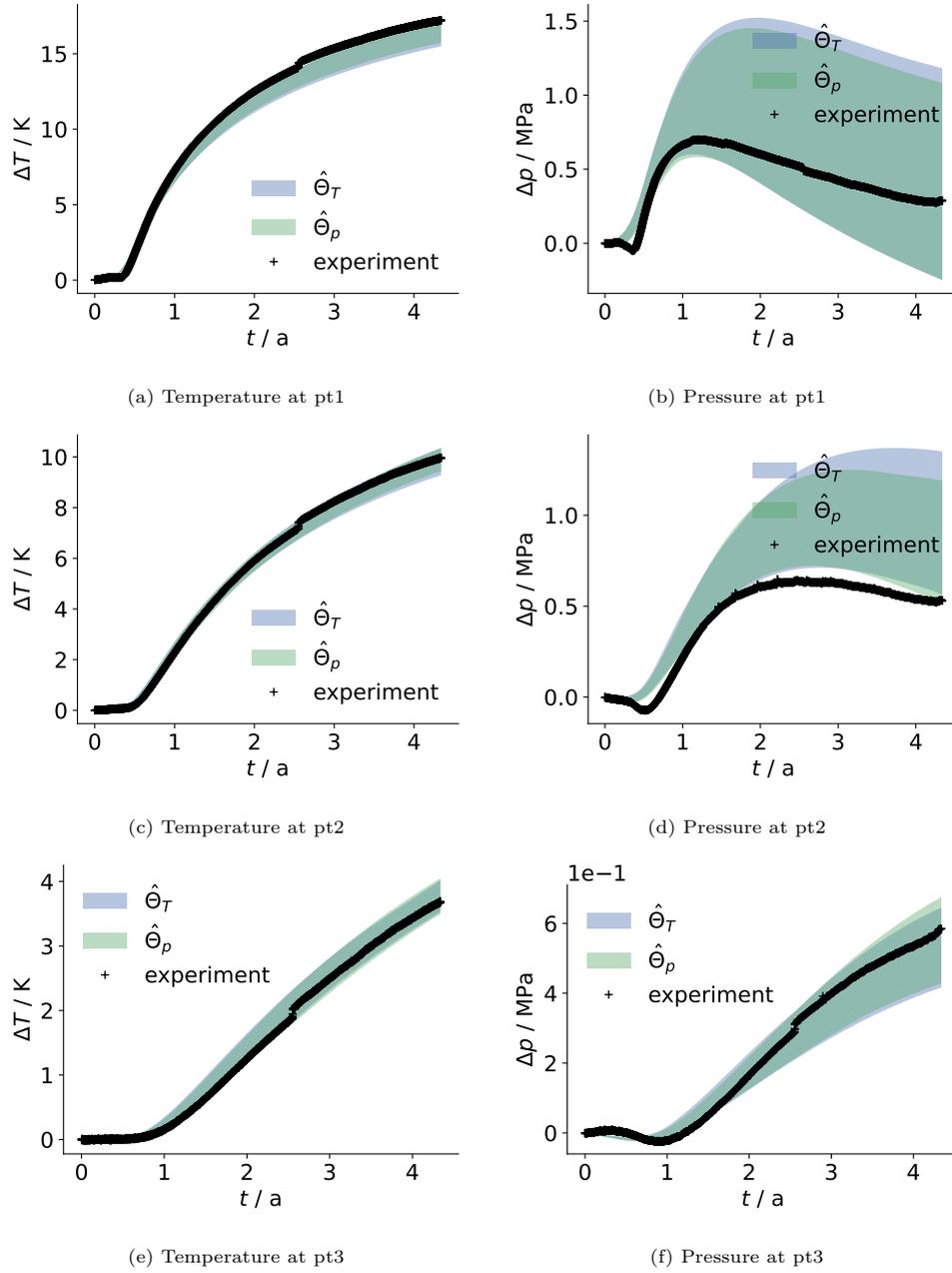


Figure 8: History matched prediction ranges for both thresholds together with the experimental results. The results of the remaining three observation points are plotted in A.17

differences for the pressure. Curves that minimize the history match error for pt1 (blue) are not minimizing the history match error for pt2 (red) or pt3 (green). These results

point to an incompleteness in the model description. This incompleteness could be linked to oversimplified representation of the bentonite, for example. This hypothesis will be tested in later studies. Nevertheless, this result is a clear demonstration of the added value of UQ analyses beyond the aspects of uncertainty quantification and reduction themselves: they provide insights into model insufficiencies that drive hypotheses for model improvement.

3.6. Global Sensitivity Analysis (Step 10)

To attribute the uncertainty to individual parameters and their combinations, we performed a global sensitivity analysis using the proxy for both variables and for all three observation points (workflow step ⑩ in Fig. 2).

To sample the input space, the filtered parameter distributions from thresholds Θ_T and Θ_p were fitted using triangular, uniform, loguniform and truncated normal distributions depending on their shapes. The input space was then sampled using their corresponding fitted distributions employing the sampling scheme of Saltelli and Tarantola [4]. Alike with the parameter estimation, Θ_p yields very similar results, which is why they are not depicted here. The influence of prior and 'posterior' parameter ranges has been discussed elsewhere [8]. The Sobol indices of first and second-order were calculated using their Monte-Carlo estimates based on a sampling size of 32,768. The sampling size was chosen such that for all indices, the width of the 0.95 confidence interval was below 0.02. The results are depicted in Fig. 11. In this analysis, we focus again on three observation points, pt1, pt2, and pt3 only. The results for the remaining points are presented in Fig. A.19. Due to the symmetric representation, we scaled the second-order indices by a factor of 0.5. For temperature, the main impact comes from the thermal conductivity, especially K^\perp and the specific heat capacity c_p . While for pt1 the effects are dominated by the main effects (diagonal), for pt2 and pt3 the impact comes also from the interaction between c_s and K^\perp . The behavior is reversed for the pressure: big changes can be seen between pt1, pt2, and pt3, while pt2 and pt3 diagonal effects seem to have a big impact on pt1 nearly all effect indices have a somewhat comparable contribution. We attribute this behavior to the functional behavior of the pressure over time. As the objective function is an integral measure of the difference between the model and the experimental curve, but the model curves show more or less a similar decline crossing the experiment measurement, no parameter could be identified that, clearly, changes/reduces the error. This is basically an expression of temporal variations of sensitivity at specific measurement points in transient problems [50]. As seen in the distribution of all simulated curves around the experimental one at pt1, most parameters that have an impact on pressure only change the position and the height of the maximum but not the slope of the decline (c.f. Fig. 4b). This behavior favors second-order indices, as combined changes improve the fit. However, it is also a sign of a remaining model error, as also the best fit remains unsatisfactory. According to our the parameter estimation, E_3 , a_s^\perp , a_s^\parallel and k^\perp have the same parameter range as the prior. This behavior is also reflected in the sensitivity analysis for all points. The biggest impact can be found for pt3. Nevertheless, for pt3, the uncertainty reduction due to the filter is the smallest.

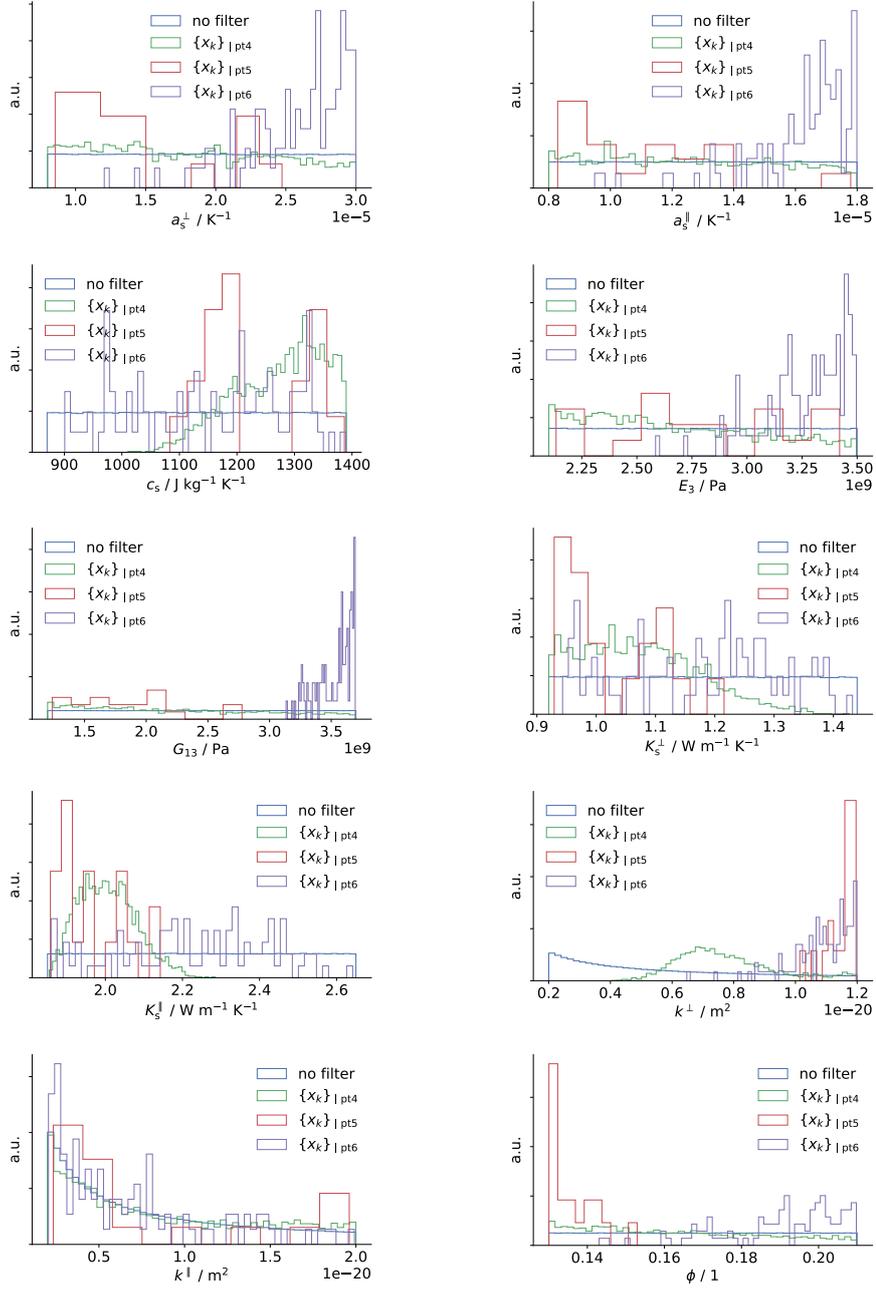


Figure 9: Parameter estimation based on subsamples $\{x_k\}_{\text{pt}T}$ optimized for each point. The counts are given in arbitrary units (a.u.).

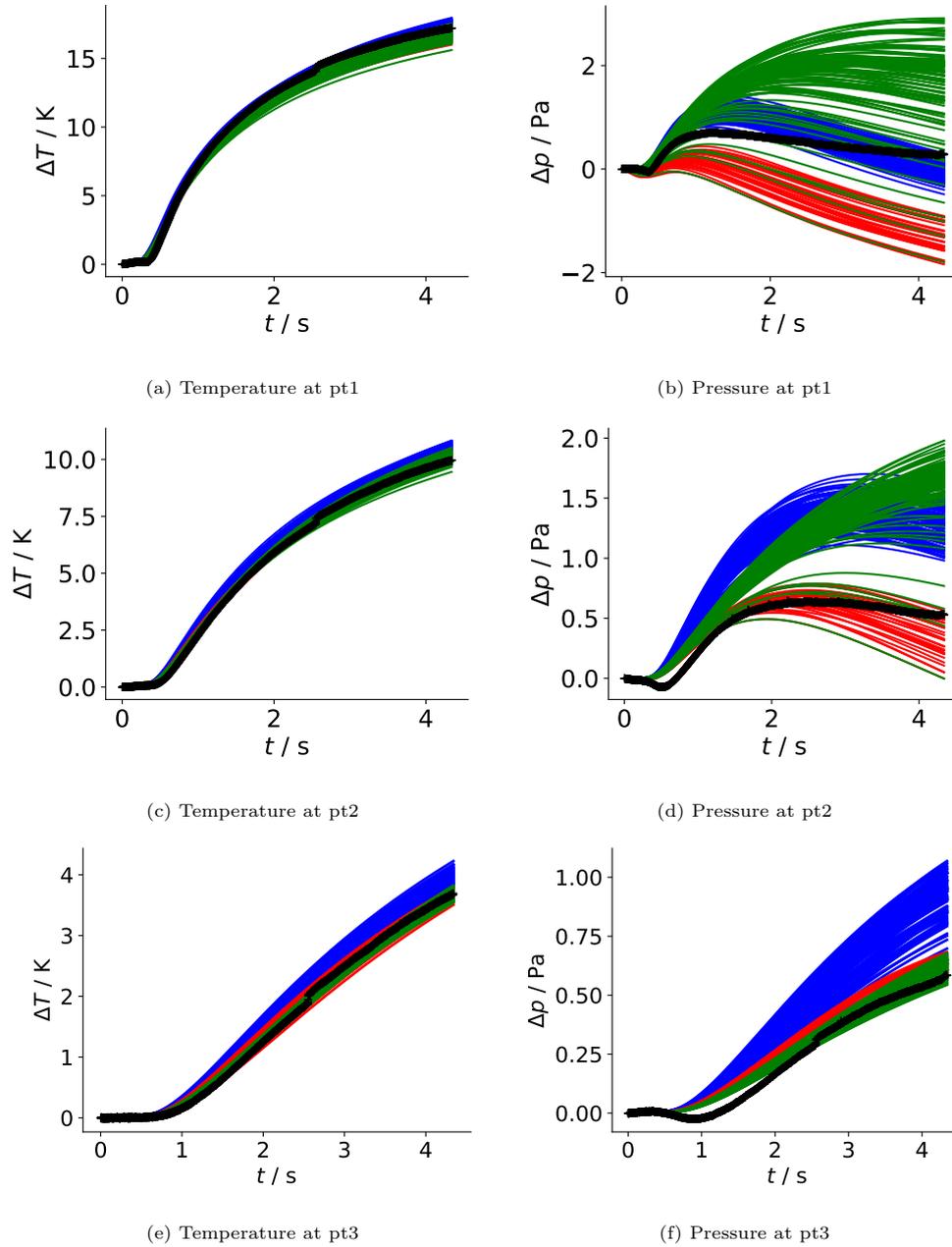


Figure 10: Subsampled observation point-based history matched forward runs. Black: experimental measurement, blue: pt1-match, red: pt2-match, green: pt3-match

4. Conclusions

In this paper, we presented the application of an experimental design-based uncertainty workflow to a full-scale emplacement experiment of a repository for high-level

radioactive waste. We used experimental data from the underground research laboratory at Mont Terri to compare with our simulations. Based on a coupled non-isothermal Richards flow model with different representations of mechanical coupling, we focused on parameter uncertainties of the surrounding clay rock. We combined experimental and modelling results to build a proxy model of the history match error. An initial screening step allowed us to reduce the number of significant uncertain parameters from 20 to 14 needed for building the proxy. This procedure allowed us to perform a parameter estimation, which we used to infer final forward model runs which are believed to be representative. Furthermore, we used the proxy to perform a global sensitivity analysis.

The applied methodology turned out to be a very valuable tool for studying parameter uncertainties and checking model validity in the context of radioactive waste disposal. The Gaussian proxy model allowed us to efficiently investigate the match between experimental and model responses, and to perform parameter estimation that further narrowed our prior expert ("measurements") guesses. For many parameters, the uncertainty range could thus be reduced. The corresponding response curves agreed very well for the temperature, but greater differences were found for the pressure response, especially for points closer to the tunnel where the heaters were placed. These results point us to one or more remaining gaps in the model description, as competing hypotheses remain as to the origin of these differences. Specifically, we suspect them to originate from geometrical aspects of uncertainties and also uncertainties of parameters / constitutive models in other material groups like the bentonite or shotcrete that we neglected in this study. The forward model used consisted of a simplified description of mechanical effects that allowed us to evaluate the THM model with the computational efficacy of a TH model. This model was a generalization of a simplification approach proposed earlier by the authors[42]. Evaluations with the full-complexity THM model showed only relative differences of a few percent in the pressure response within the give time frame, such that we conclude its validity under partially saturated conditions. This is an important result in itself, as this approximation is of great value in situations where primarily temperature and pressure predictions are of interest.

Future work would benefit from more detailed knowledge of the statistical behavior of material parameter properties of all material domains, as specific parameter distributions would have allowed us to reduce the uncertainty in the early stages further.

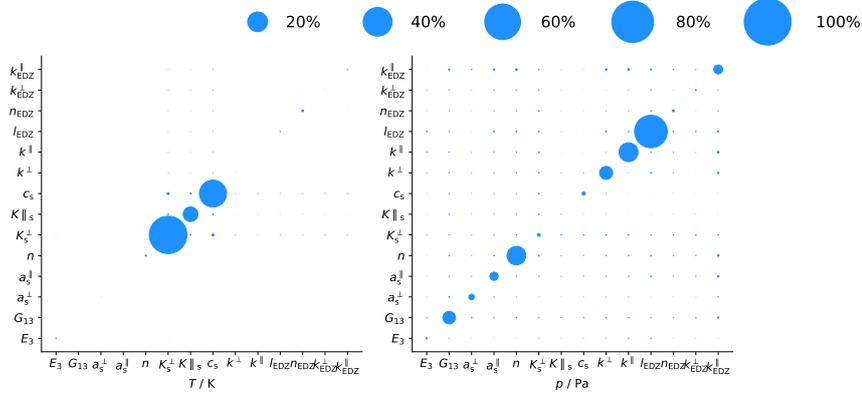
Although it is difficult to rank other potential effects in their suspected impact on uncertainty without a detailed investigation, we hypothesize that remaining uncertainties stem from the representation and parameterization of the bentonite buffer, shotcrete and EDZ as well as geometrical effects which should be the matter of future work.

Although some challenges in tackling complexities remain, we conclude that the workflow presented herein is well suited for examining sensitivities and uncertainties as part of a proper safety assessment and to validate the models employed therein.

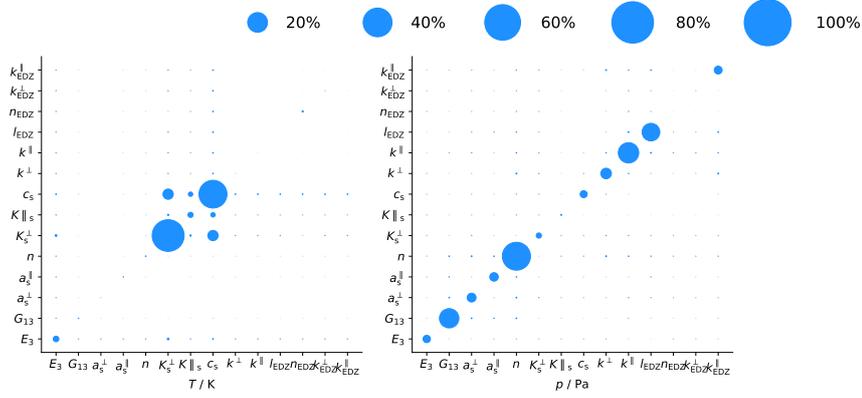
Acknowledgments

The basis of this work is a generic workflow that develops complex uncertainty analyses. The first foundations were laid in the iCROSS project funded by the German Federal Ministry of Education and Research (BMBF, grant 02NUK053E) and the Helmholtz Association (grant SO-093). This work has been co-financed within the framework of EURAD, the European Joint Programme on Radioactive Waste Management (Grant Agree-

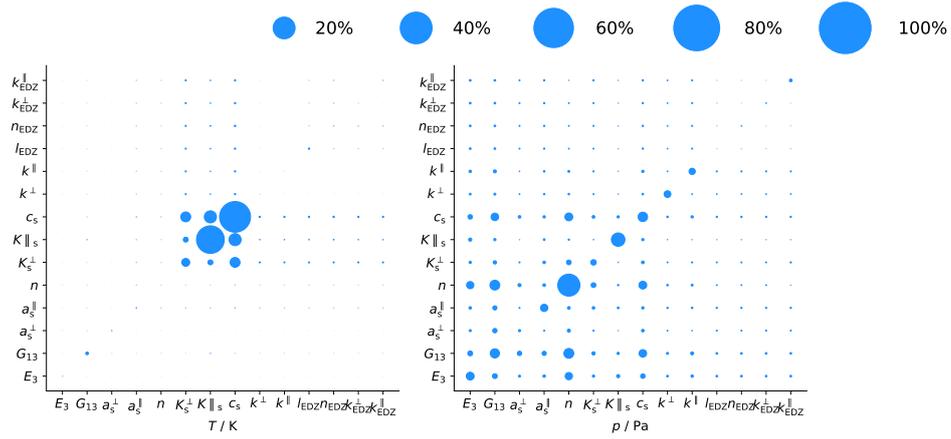
ment No 847593), this support is gratefully acknowledged. We would also like to thank the DECOVALEX team (Task C), the FE Task Force led by Nagra as well as swisstopo for providing data and fruitful discussions for the modelling of the FE experiment in the underground research laboratory at Mont Terri. DECOVALEX (<https://decovallex.org>) is an international research project comprising participants from industry, government and academia, focusing on development of understanding, models and codes in complex coupled problems in subsurface geological and engineering applications; DECOVALEX-2023 is the current phase of the project. The authors appreciate and thank the DECOVALEX-2023 Funding Organisations Andra, BASE, BGE, BGR, CAS, CNSC, COVRA, US DOE, ENRESA, ENSI, JAEA, KAERI, NWMO, RWM, SÚRAO, SSM and Taipower for their financial and technical support of the work described in this paper. The uncertainty analysis work is currently being carried out by the URS project (Uncertainties and Robustness with regard to the Safety of a repository for high-level radioactive waste) funded by the Bundesgesellschaft für Endlagerung mbH (BGE). We would like to express our sincere thanks for past and present funding. We are also very grateful to the Open-GeoSys development team for their enthusiastic and continuous work to further develop and improve the OGS platform for the scientific community.



(a) pt1



(b) pt2



(c) pt3

Figure 11: Global sensitivity analysis based on threshold Θ_T for the first three observation points.

Appendix A. supplemental

Appendix A.1. Comparison between the simplified TRuni and the fully coupled TRM model

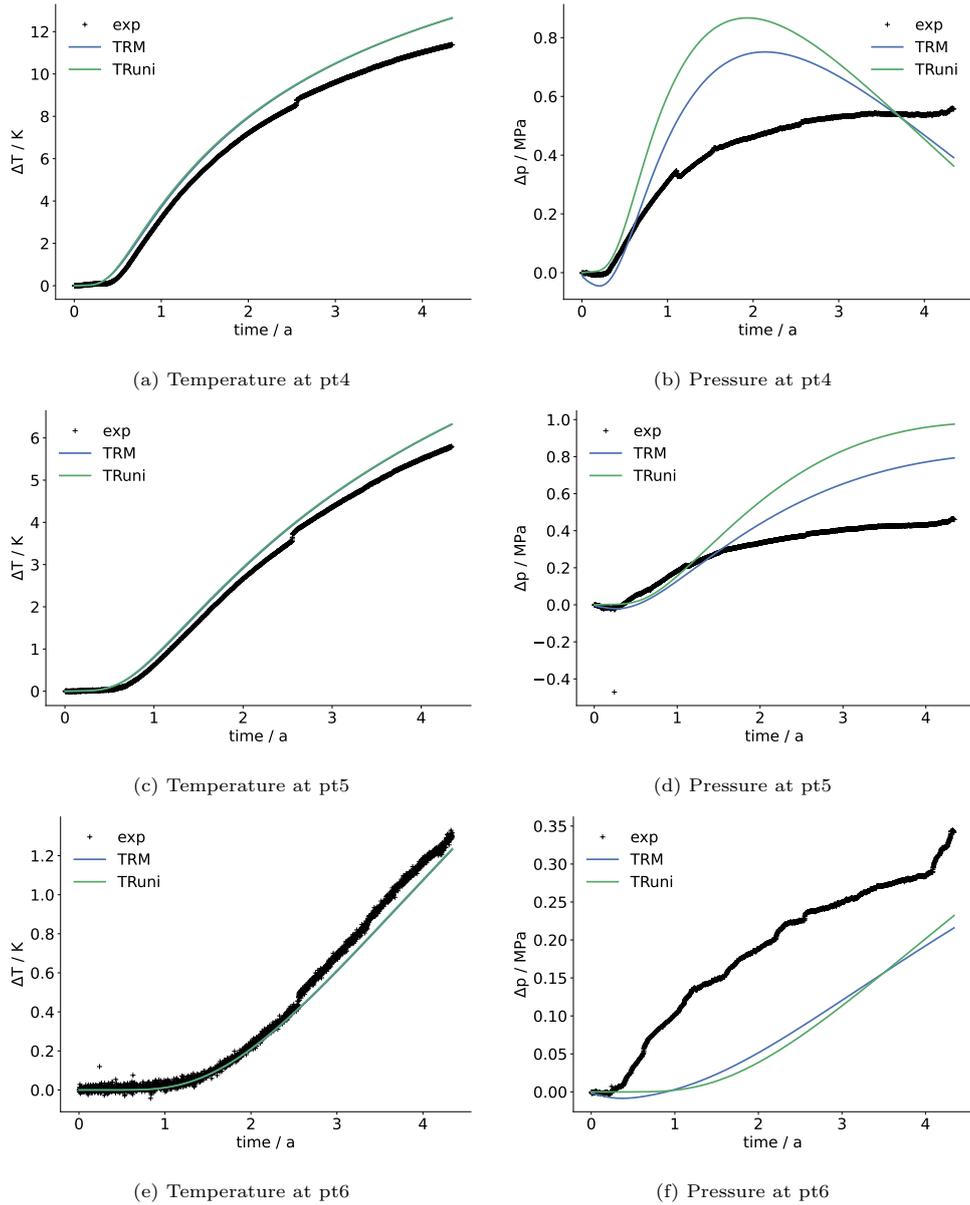


Figure A.12: Temperature and pressure responses at the last three observation points: Best fit results. The measured TR curve is plotted alongside the TRuni and the fully coupled TRM response.

Appendix A.2. Parameter estimation based on each variable selected subsample

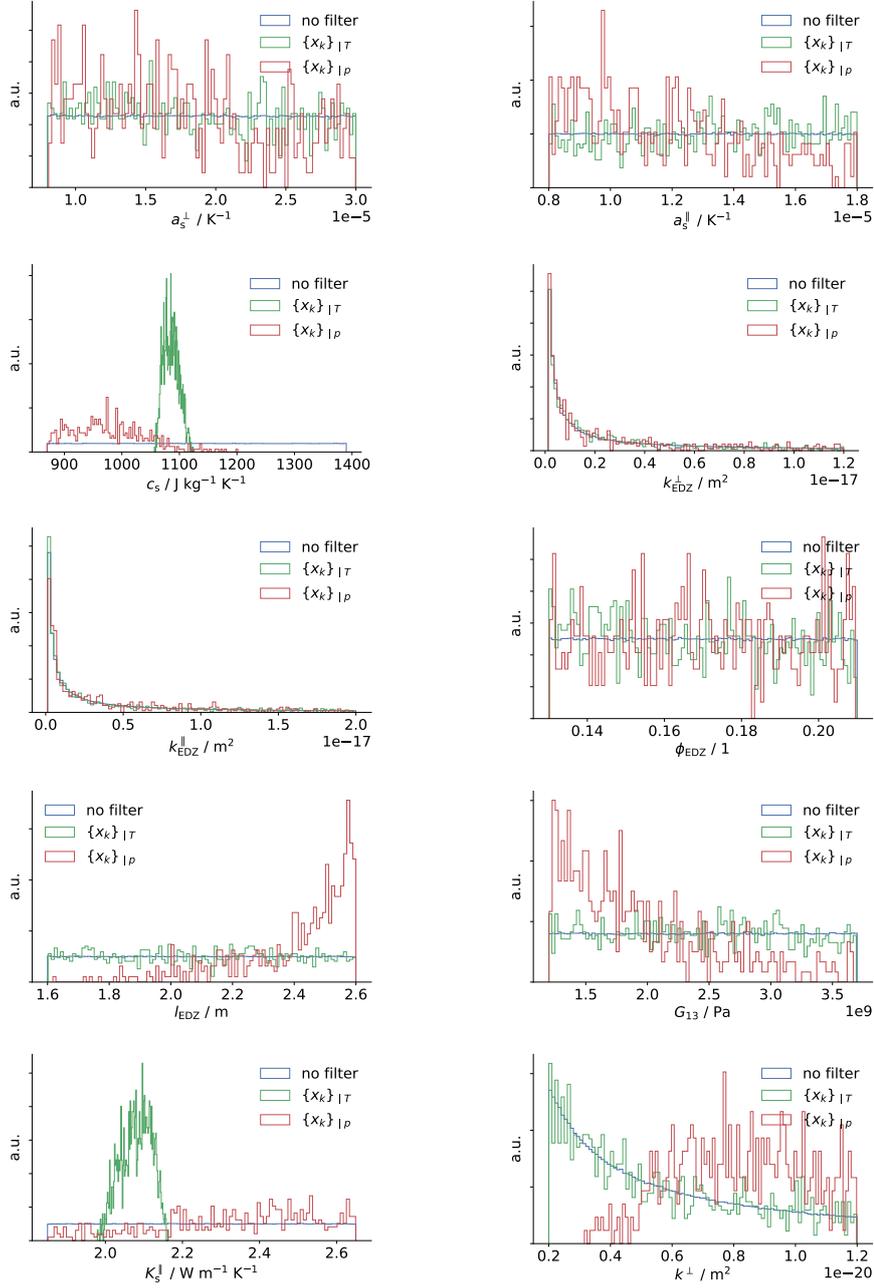


Figure A.13: Parameter estimation based on each variable selected subsample. The counts are given in arbitrary units (a.u.).

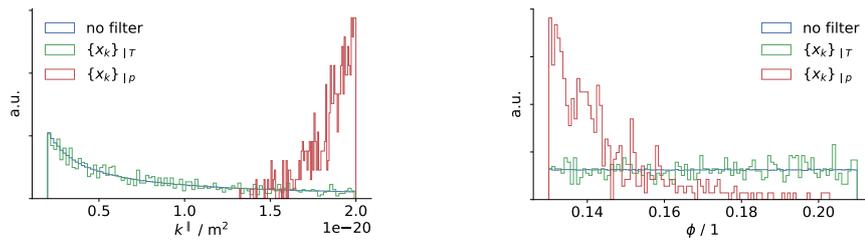


Figure A.14: Parameter estimation based on each variable selected subsample. The counts are given in arbitrary units (a.u.).

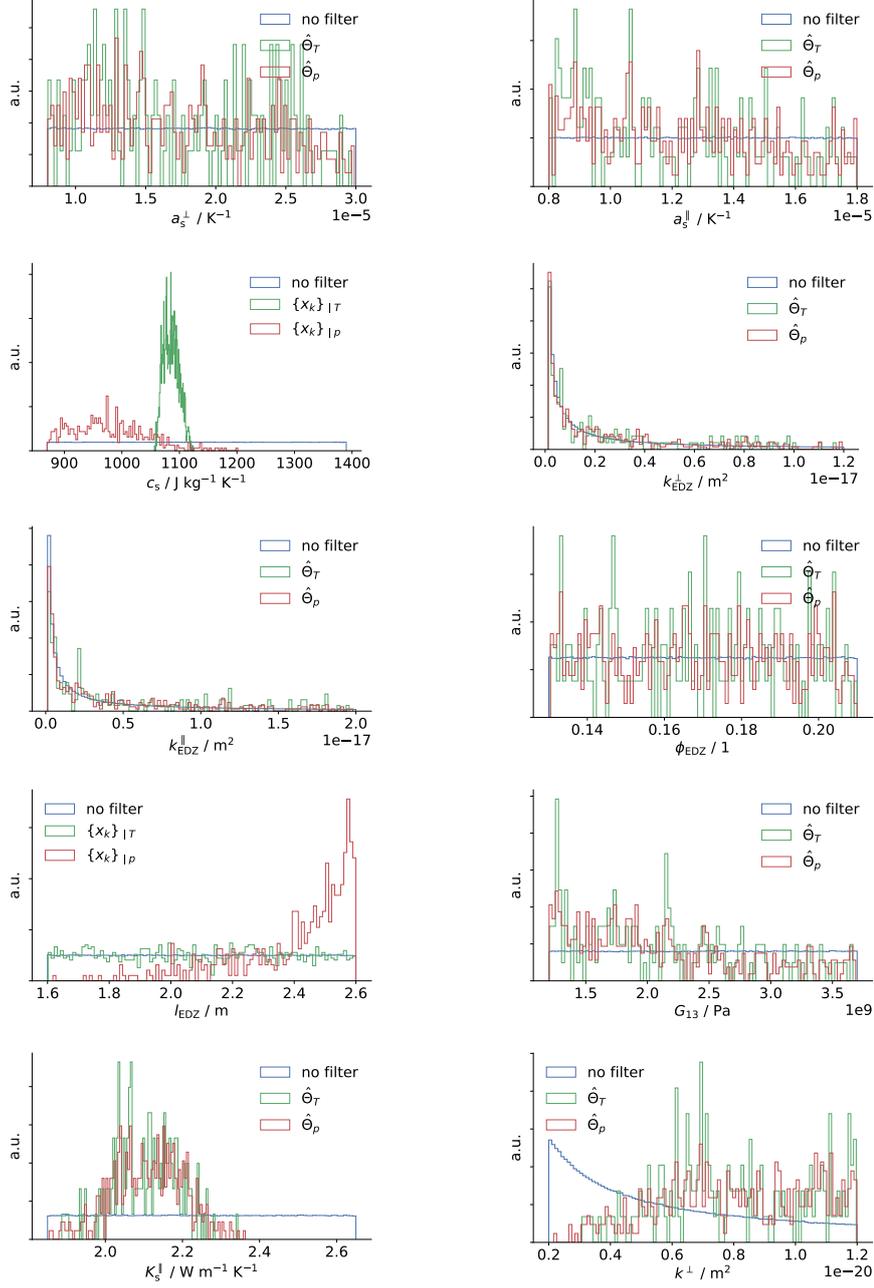


Figure A.15: Parameter estimation for two selected parameters based on criteria thresholds $\hat{\Theta}_T$ and $\hat{\Theta}_p$ compared to the prior parameter distribution. The counts are given in arbitrary units (a.u.).

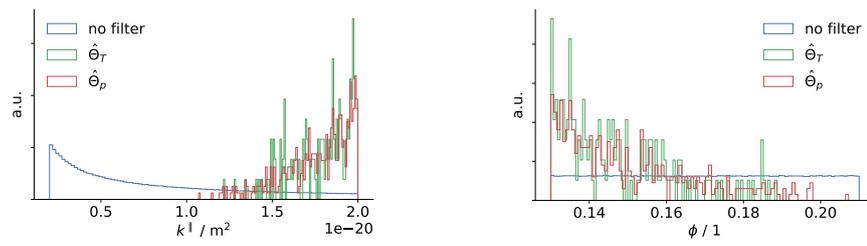
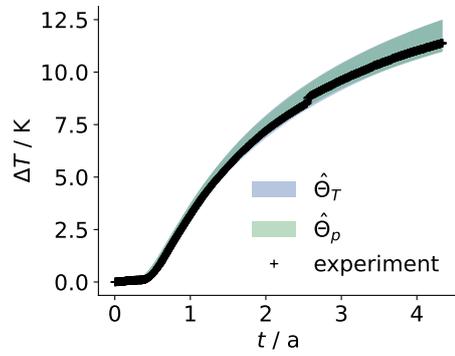
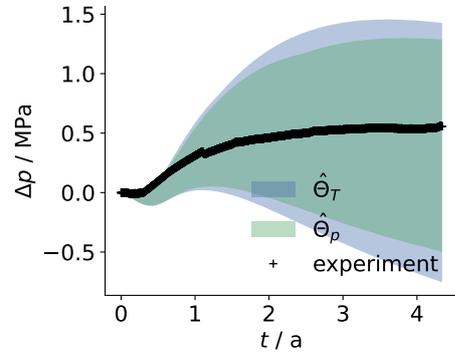


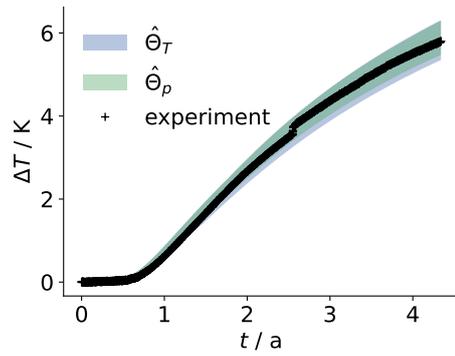
Figure A.16: Parameter estimation for two selected parameters based on criteria thresholds $\hat{\theta}_T$ and $\hat{\theta}_p$ compared to the prior parameter distribution. The counts are given in arbitrary units (a.u.).



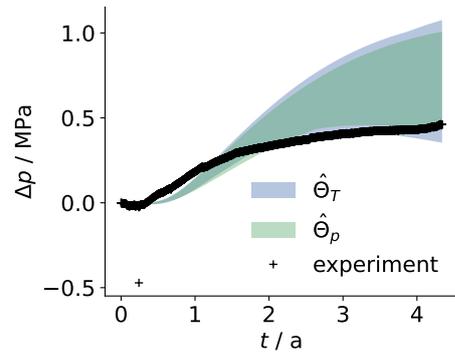
(a) Temperature at pt4



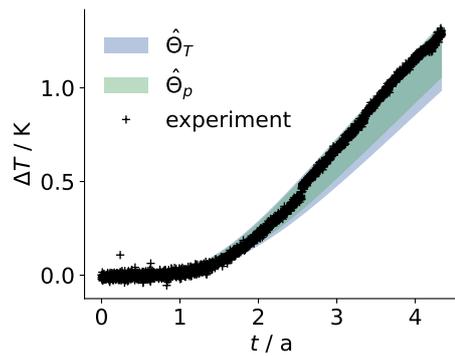
(b) Pressure at pt4



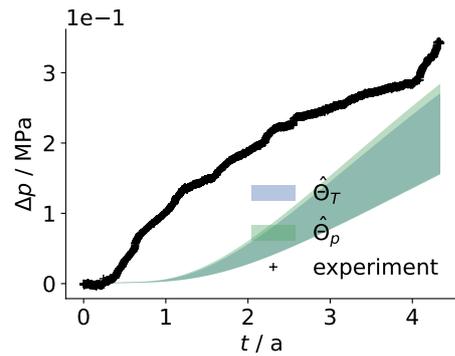
(c) Temperature at pt5



(d) Pressure at pt5



(e) Temperature at pt6



(f) Pressure at pt6

Figure A.17: History matched prediction ranges for both thresholds together with the experimental results.

Table A.7: Proxy errors $\widehat{E}_{\text{proxy}}^{\text{temperature}}$ from the difference between the history-matched forward runs to the proxy predictions.

	RMSE (LHS)	RMSE (MC/ $\hat{\theta}_T$)	RMSE (MC/ $\hat{\theta}_p$)
pt1	0.01 K	0.01 K	0.01 K
pt2	0.01 K	0.02 K	0.02 K
pt3	0.00 K	0.01 K	0.01 K
pt4	0.01 K	0.02 K	0.01 K
pt5	0.01 K	0.03 K	0.02 K
pt6	0.01 K	0.01 K	0.01 K

Table A.8: Proxy errors $\widehat{E}_{\text{proxy}}^{\text{pressure}}$ response from the difference between the history-matched forward runs to the proxy predictions.

	RMSE (LHS)	RMSE (MC/ $\hat{\theta}_T$)	RMSE (MC/ $\hat{\theta}_p$)
pt1	51513 Pa	97232 Pa	101039 Pa
pt2	7398 Pa	23681 Pa	23493 Pa
pt3	8753 Pa	13505 Pa	14305 Pa
pt4	47973 Pa	74623 Pa	60725 Pa
pt5	10277 Pa	9960 Pa	10992 Pa
pt6	1060 Pa	606 Pa	711 Pa

Appendix A.3. Proxy quality from MC match.

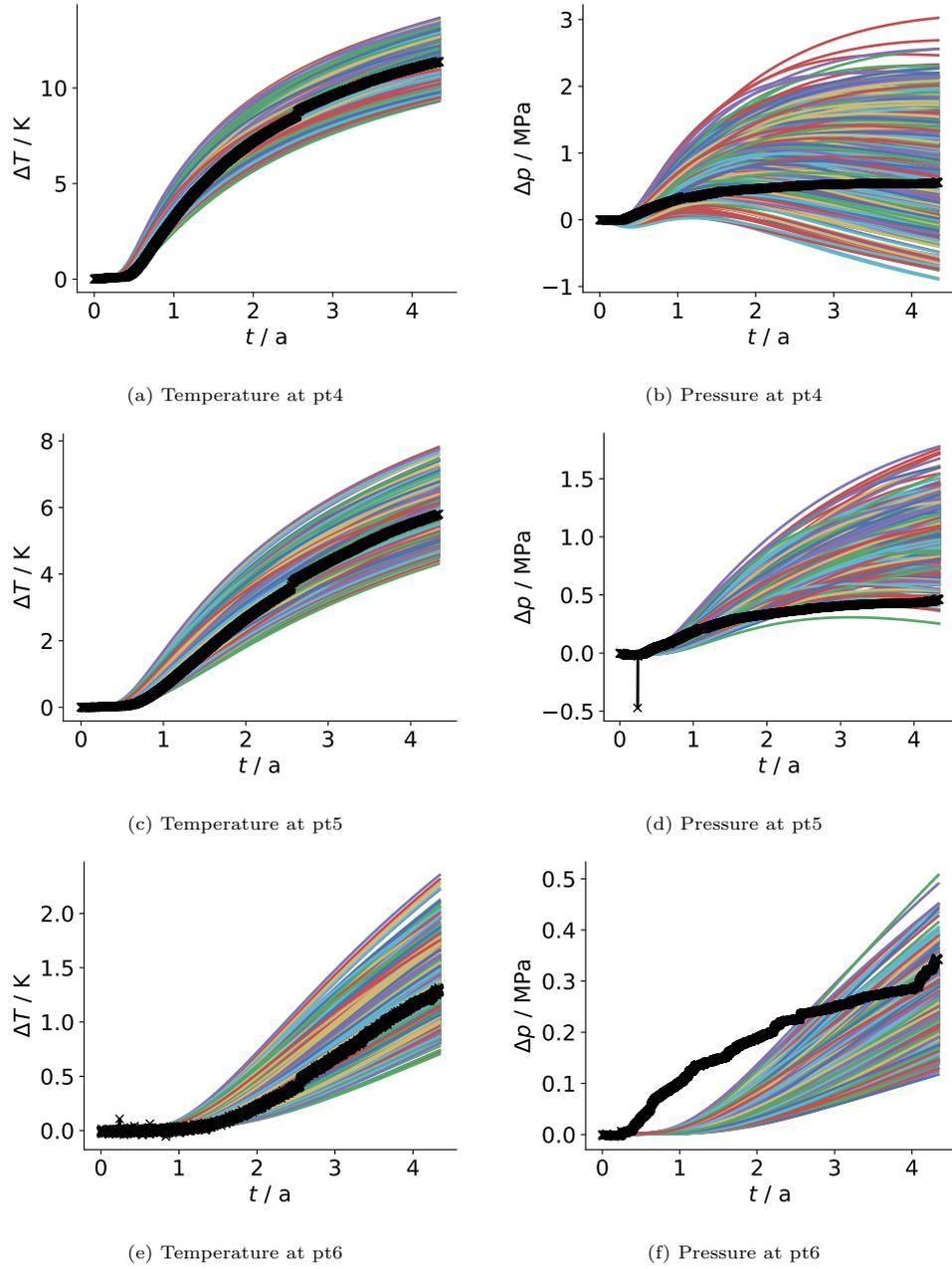
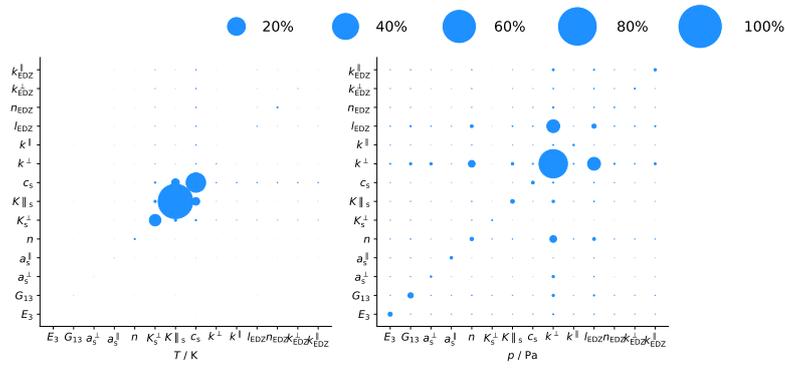
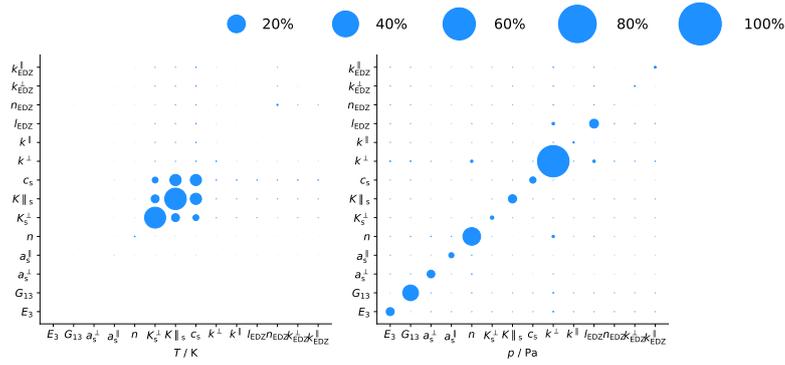


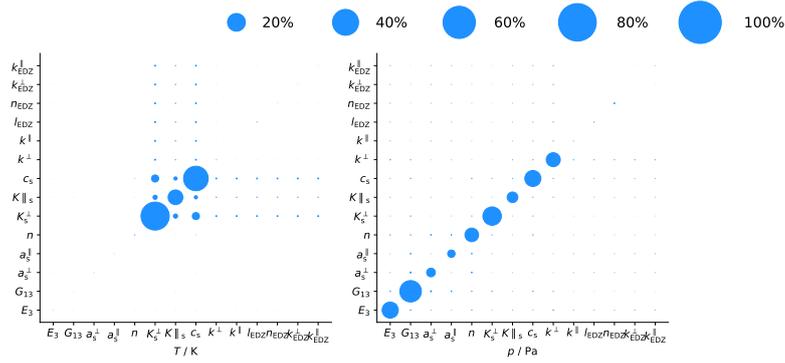
Figure A.18: Ensemble run temperature and pressure curves generated from LHS design based on prior distributions for the first three observation points. The measured response is painted black. Colored curves present modeled responses of TRuni model for different sets of parameters.



(a) pt4



(b) pt6



(c) pt6

Figure A.19: Global sensitivity analysis based on threshold Θ_T .

Appendix B. Other Material Parameters

Table B.9: Material parameters for bentonite and concrete material parts. Thermal conductivity models used in the study: a) $K_{\text{eff}} = (1 - S_L)K_{\text{dry}} + S_L K_{\text{wet}}$, b) $K_{\text{eff}} = \sqrt{1 - S_L}K_{\text{dry}} + \sqrt{S_L}K_{\text{wet}}$, c) $K_{\text{eff}} = K_{\text{dry}}^{1-S_L} K_{\text{wet}}^{S_L}$.

title	symbol	unit	granular bentonite	bentonite blocks	concrete/shotcrete
Solid density	ρ_s	kg m ⁻³	2500.0	2500.0	2706
Specific heat capacity	c_s	J kg ⁻¹ K ⁻¹	800	800	750
Thermal expansivity	a_s	K ⁻¹	3.0e-6	3.0e-6	0.8e-5
Young's modulus	E	Pa	18e6	24e6	2e10
Poisson's ratio	ν	1	0.35	0.2	0.15
saturated thermal conductivity	K_{wet}	W m ⁻¹ K ⁻¹	1.2	0.96	1.83
dry thermal conductivity	K_{dry}	W m ⁻¹ K ⁻¹	0.35	0.26	0.98
shape variation $K(S_L)$	-	-	c	b	a
Intrinsic permeability	k	m ²	3.5e-20	2.5e-21	1e-19
Porosity	ϕ	1	0.331	0.331	0.25
Biot-Willis coefficient	α_B	1	0.6	0.6	0.6
Saturation: Van Genuchten, residual liquid saturation	S_L^r	1	0.01	0.01	0.01
Saturation: Van Genuchten, residual gas saturation	S_g^r	1	1e-15	1e-15	1e-15
Saturation: Van Genuchten, exponent	m	1	0.45	0.45	0.33
Saturation: Van Genuchten, entry pressure	p_b	Pa	18e6	30e6	1.0e6
Relative permeability: Van Genuchten, residual liquid saturation	S_L^r	1	0.01	0.01	0.01
Relative permeability: Van Genuchten, residual gas saturation	S_g^r	1	1e-15	1e-15	1e-15
Relative permeability: Van Genuchten, exponent	m	1	0.45	0.45	0.33
Relative permeability: Van Genuchten, minimum relative permeability	k_r^{min}	1	1e-9	1e-9	1e-9

References

- [1] D. Draper, A. Pereira, P. Prado, A. Saltelli, R. Cheal, S. Eguilior, B. Mendes, S. Tarantola, Scenario and parametric uncertainty in GESAMAC: A methodological study in nuclear waste disposal risk assessment, *Comput Phys Commun* 117 (1) (1999) 142–155.
- [2] E. Zio, G. E. Apostolakis, Two methods for the structured assessment of model uncertainty by experts in performance assessments of radioactive waste repositories, *Reliab Eng Syst Saf* 54 (2-3) (1996) 225–241.
- [3] J. C. Helton, Uncertainty and sensitivity analysis techniques for use in performance assessment for radioactive waste disposal, *Reliab Eng Syst Saf* 42 (2-3) (1993) 327–367.
- [4] A. Saltelli, S. Tarantola, On the relative importance of input factors in mathematical models: Safety assessment for nuclear waste disposal, *J Am Stat Assoc* 97 (459) (2002) 702–709.
- [5] S. J. Wang, K. C. Hsu, Stochastic Analysis of a Thermal Uncoupled Thermal-Hydraulic-Mechanical Model, in: *Poromechanics 2017 - Proceedings of the 6th Biot Conference on Poromechanics*, American Society of Civil Engineers (ASCE), 2017, pp. 787–794.
- [6] L. Nguyen-Tuan, T. Lahmer, M. Datcheva, E. Stoimenova, T. Schanz, A novel parameter identification approach for buffer elements involving complex coupled thermo-hydro-mechanical analyses, *Comput Geotech* (2016).
- [7] N. Watanabe, W. Wang, C. I. McDermott, T. Taniguchi, O. Kolditz, , *Comput Mech* 45 (4) (2010) 263 – 280. doi:10.1007/s00466-009-0445-9.
URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-77955419170&doi=10.1007/2fs00466-009-0445-9&partnerID=40&md5=92adadea7505159789eeb96a48fa138a>
- [8] J. Buchwald, A. Chaudhry, K. Yoshioka, O. Kolditz, S. Attinger, T. Nagel, , *Int J Rock Mech Min Sci* 134 (2020). doi:10.1016/j.ijrmms.2020.104481.
URL <https://www.scopus.com/inward/record.uri?eid=2-s2.0-85090353571&doi=10.1016%2fj.ijrmms.2020.104481&partnerID=40&md5=10fbce5a044c19fc975ca8c6ea21c57b>
- [9] Uncertainty propagation through Thermo-Hydro-Mechanical modelling of concrete cracking and leakage – Application to containment buildings, *International Association for Fracture Mechanics of Concrete and Concrete Structures*, 2019. doi:10.21012/FC10.235507.
- [10] G. Hu, W. Pfungsten, Machine learning-assisted heat transport modelling for full-scale emplacement experiment at Mont Terri underground laboratory, *Int J Heat Mass Transf* 213 (2023) 124290. doi:<https://doi.org/10.1016/j.ijheatmasstransfer.2023.124290>.
URL <https://www.sciencedirect.com/science/article/pii/S0017931023004428>
- [11] M. Williams, M. Ziegler, S. Schennen, S. Loew, Evolution of excavation damaged zones in opalinus clay shale inferred from seismic investigations, *Eng Geol* 299 (2022) 106528. doi:10.1016/J.ENGGEOL.2022.106528.
- [12] R. Ludwig, K. Schelkes, P. Vogel, J. Wollrath, Implications of large-scale heterogeneities for hydraulic model studies at the potential site of a radioactive waste repository at gorleben, germany, *Eng Geol* 61 (2-3) (2001) 119–130.
- [13] T. Yuan, Y. Yang, N. Ait-Mouheb, G. Deissmann, C. Fischer, T. Stumpf, D. Bosbach, A comparative study on heterogeneity of clay rocks using pore-scale diffusion simulations and experiments, *J Geophys Res Solid Earth* 127 (12) (2022) e2022JB025428.
- [14] N. Ait-Mouheb, Y. Yang, L. R. Van Loon, M. A. Glaus, G. Deissmann, D. Bosbach, Impact of the heterogeneity of the sandy Opalinus clay facies at the Mont Terri underground research laboratory on radionuclide migration, *Safety of Nuclear Waste Disposal 1* (2021) 141–142. doi:10.5194/sand-1-141-2021.
URL <https://sand.copernicus.org/articles/1/141/2021/>
- [15] H. Cleaves, A. Alexanderian, B. Saad, Structure exploiting methods for fast uncertainty quantification in multiphase flow through heterogeneous media, *Computational Geosciences* 25 (2021) 2167–2189.
- [16] C.-F. Tsang, O. Stephansson, J. Hudson, A discussion of thermo–hydro–mechanical (THM) processes associated with nuclear waste repositories, *Int J Rock Mech Min Sci* 37 (1-2) (2000) 397–402.
- [17] C. F. Tsang, J. D. Barnichon, J. Birkholzer, X. L. Li, H. H. Liu, X. Sillen, 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* 49 (2012) 31–44.
- [18] F. Bernier, F. Lemy, P. De Cannière, V. Detilleux, Implications of safety requirements for the treatment of THMC processes in geological disposal systems for radioactive waste, *J Rock Mech Geotech Eng* 9 (3) (2017) 428–434.
- [19] X.-L. Li, J. Lanru, P. Blaser, Impact of Thermo-Hydro-Mechanical- Chemical (THMC) processes

- on the safety of underground radioactive waste repositories. Proceedings of the European Commission TIMODAZ-THERESA International Conference, Luxembourg, 29 September - 1 October 2009, in: Proceedings of the European Commission TIMODAZ-THERESA International Conference. Luxembourg: Publications Office, 2009, pp. 489–494.
- [20] H. R. Müller, B. Garitte, T. Vogt, S. Köhler, T. Sakaki, H. Weber, T. Spillmann, M. Hertrich, J. K. Becker, N. Giroud, V. Cloet, N. Diomidis, T. Vietor, Implementation of the full-scale emplacement (FE) experiment at the Mont Terri rock laboratory, in: Swiss Journal of Geosciences, Vol. 110, Springer, 2017, pp. 287–306.
- [21] J. Buchwald, O. Kolditz, T. Nagel, ogs6py and vtuinterface: streamlining opengeosys workflows in python, *J Open Source Softw* 6 (67) (2021) 3673. doi:10.21105/joss.03673. URL <https://doi.org/10.21105/joss.03673>
- [22] B. Li, E. W. Bhark, S. J. Gross, T. C. Billiter, K. Dehghani, Best practices of assisted history matching using design of experiments, *SPE J* 24 (4) (2019) 1435–1451.
- [23] B. J. Debusschere, D. T. Seidl, T. M. Berg, K. W. Chang, R. C. Leone, L. P. Swiler, P. E. Mariner, Machine learning surrogates of a fuel matrix degradation process model for performance assessment of a nuclear waste repository, *Nuclear Technology* 209 (9) (2023) 1295–1318. doi:10.1080/00295450.2023.2197666. URL <https://doi.org/10.1080/00295450.2023.2197666>
- [24] D. Brooks, L. Swiler, E. Stein, P. Mariner, E. Basurto, T. Portone, A. Eckert, R. Leone, Sensitivity analysis of generic deep geologic repository with focus on spatial heterogeneity induced by stochastic fracture network generation, *Advances in Water Resources* 169 (2022) 104310. doi:https://doi.org/10.1016/j.advwatres.2022.104310. URL <https://www.sciencedirect.com/science/article/pii/S0309170822001737>
- [25] F. Cadini, A. Gioletta, E. Zio, Improved metamodel-based importance sampling for the performance assessment of radioactive waste repositories, *Reliab Eng Syst Saf* 134 (2015) 188–197. doi:https://doi.org/10.1016/j.res.2014.10.018. URL <https://www.sciencedirect.com/science/article/pii/S0951832014002580>
- [26] N.-C. Xiao, K. Yuan, H. Zhan, System reliability analysis based on dependent kriging predictions and parallel learning strategy, *Reliab Eng Syst Saf* 218 (2022) 108083. doi:https://doi.org/10.1016/j.res.2021.108083. URL <https://www.sciencedirect.com/science/article/pii/S0951832021005810>
- [27] S. Saraygord Afshari, F. Enayatollahi, X. Xu, X. Liang, Machine learning-based methods in structural reliability analysis: A review, *Reliab Eng Syst Saf* 219 (2022) 108223. doi:https://doi.org/10.1016/j.res.2021.108223. URL <https://www.sciencedirect.com/science/article/pii/S0951832021007018>
- [28] J. Wang, G. Xu, Y. Li, A. Kareem, Akse: A novel adaptive kriging method combining sampling region scheme and error-based stopping criterion for structural reliability analysis, *Reliab Eng Syst Saf* 219 (2022) 108214. doi:https://doi.org/10.1016/j.res.2021.108214. URL <https://www.sciencedirect.com/science/article/pii/S095183202100692X>
- [29] A. Ameryan, M. Ghalehnovi, M. Rashki, Ak-sesc: a novel reliability procedure based on the integration of active learning kriging and sequential space conversion method, *Reliab Eng Syst Saf* 217 (2022) 108036. doi:https://doi.org/10.1016/j.res.2021.108036. URL <https://www.sciencedirect.com/science/article/pii/S0951832021005433>
- [30] T. Wang, Z. Chen, G. Li, J. He, C. Liu, X. Du, A novel method for high-dimensional reliability analysis based on activity score and adaptive kriging, *Reliab Eng Syst Saf* 241 (2024) 109643. doi:https://doi.org/10.1016/j.res.2023.109643. URL <https://www.sciencedirect.com/science/article/pii/S0951832023005574>
- [31] C. Zhou, Z. Shi, S. Kucherenko, H. Zhao, A unified approach for global sensitivity analysis based on active subspace and kriging, *Reliab Eng Syst Saf* 217 (2022) 108080. doi:https://doi.org/10.1016/j.res.2021.108080. URL <https://www.sciencedirect.com/science/article/pii/S0951832021005780>
- [32] Z. ming Jiang, D.-C. Feng, H. Zhou, W.-F. Tao, A recursive dimension-reduction method for high-dimensional reliability analysis with rare failure event, *Reliab Eng Syst Saf* 213 (2021) 107710. doi:https://doi.org/10.1016/j.res.2021.107710. URL <https://www.sciencedirect.com/science/article/pii/S0951832021002453>
- [33] E. Tosoni, A. Salo, J. Govaerts, E. Zio, Comprehensiveness of scenarios in the safety assessment of nuclear waste repositories, *Reliab Eng Syst Saf* 188 (2019) 561–573. doi:https://doi.org/10.1016/j.res.2019.04.012. URL <https://www.sciencedirect.com/science/article/pii/S0951832019300614>

- [34] A. Davila-Frias, N. Yodo, T. Le, O. P. Yadav, A deep neural network and bayesian method based framework for all-terminal network reliability estimation considering degradation, *Reliab Eng Syst Saf* 229 (2023) 108881. doi:<https://doi.org/10.1016/j.res.2022.108881>. URL <https://www.sciencedirect.com/science/article/pii/S0951832022004987>
- [35] G. Roma, F. Di Maio, A. Bersano, N. Pedroni, C. Bertani, F. Mascari, E. Zio, A bayesian framework of inverse uncertainty quantification with principal component analysis and kriging for the reliability analysis of passive safety systems, *Nuclear Engineering and Design* 379 (2021) 111230. doi:<https://doi.org/10.1016/j.nucengdes.2021.111230>. URL <https://www.sciencedirect.com/science/article/pii/S0029549321001825>
- [36] S. Kakhaia, P. Zun, D. Ye, V. Krzhizhanovskaya, Inverse uncertainty quantification of a mechanical model of arterial tissue with surrogate modelling, *Reliab Eng Syst Saf* 238 (2023) 109393. doi:<https://doi.org/10.1016/j.res.2023.109393>. URL <https://www.sciencedirect.com/science/article/pii/S0951832023003071>
- [37] Z. Qin, M. Naser, Machine learning and model driven bayesian uncertainty quantification in suspended nonstructural systems, *Reliab Eng Syst Saf* 237 (2023) 109392. doi:<https://doi.org/10.1016/j.res.2023.109392>. URL <https://www.sciencedirect.com/science/article/pii/S095183202300306X>
- [38] E. Saveleva, V. Svitelman, P. Blinov, D. Valetov, Sensitivity analysis and model calibration as a part of the model development process in radioactive waste disposal safety assessment, *Reliab Eng Syst Saf* 210 (2021) 107521. doi:[10.1016/J.RESS.2021.107521](https://doi.org/10.1016/J.RESS.2021.107521).
- [39] N. Grunwald, C. Lehmann, J. Maßmann, D. Naumov, O. Kolditz, T. Nagel, Non-isothermal two-phase flow in deformable porous media: systematic open-source implementation and verification procedure, *Geomech Geophys Geo-energ Resou* 8 (2022) 107–108. doi:[10.1007/s40948-022-00394-2](https://doi.org/10.1007/s40948-022-00394-2). URL <https://doi.org/10.1007/s40948-022-00394-2>
- [40] M. Pitz, S. Kaiser, N. Grunwald, V. Kumar, J. Buchwald, W. Wang, D. Naumov, A. A. Chaudhry, J. Maßmann, J. Thiedau, O. Kolditz, T. Nagel, Non-isothermal consolidation: A systematic evaluation of two implementations based on multiphase and richards equations, *Int J Rock Mech Min Sci* 170 (2023) 105534. doi:<https://doi.org/10.1016/j.ijrmm.2023.105534>. URL <https://www.sciencedirect.com/science/article/pii/S1365160923002083>
- [41] W. Wang, J. Rutqvist, U. J. Görke, J. T. Birkholzer, O. Kolditz, Non-isothermal flow in low permeable porous media: A comparison of richards' and two-phase flow approaches, *Environ Earth Sci* 62 (2011) 1197–1207. doi:[10.1007/S12665-010-0608-1](https://doi.org/10.1007/S12665-010-0608-1).
- [42] J. Buchwald, S. Kaiser, O. Kolditz, T. Nagel, Improved predictions of thermal fluid pressurization in hydro-thermal models based on consistent incorporation of thermo-mechanical effects in anisotropic porous media, *Int J Heat Mass Transf* 172 (2021) 121127. doi:<https://doi.org/10.1016/j.ijheatmasstransfer.2021.121127>. URL <https://www.sciencedirect.com/science/article/pii/S0017931021002301>
- [43] O. Kolditz, S. Bauer, L. Bilke, N. Böttcher, J. O. Delfs, T. Fischer, U. J. Görke, T. Kalbacher, G. Kosakowski, C. I. McDermott, C. H. Park, F. Radu, K. Rink, H. Shao, H. B. Shao, F. Sun, Y. Y. Sun, A. K. Singh, J. Taron, M. Walther, W. Wang, N. Watanabe, Y. Wu, M. Xie, W. Xu, B. Zehner, OpenGeoSys: An open-source initiative for numerical simulation of thermo-hydro-mechanical/chemical (THM/C) processes in porous media, *Environ Earth Sci* 67 (2) (2012) 589–599.
- [44] L. Bilke, B. Flemisch, T. Kalbacher, O. Kolditz, R. Helmig, T. Nagel, Development of open-source porous-media simulators: principles and experiences., *Transp Porous Media* (2019).
- [45] C. Schoutrop, M. Senders, L. Kuijpers, J. Wehner, J. van Dijk, A. Vijaykumar, Open source idr(s)stab(l) implementation (2021).
- [46] K. Aihara, K. Abe, E. Ishiwata, A variant of idrstab with reliable update strategies for solving sparse linear systems, *J Comput Appl Math* 259 (2014) 244–258. doi:[10.1016/J.CAM.2013.08.028](https://doi.org/10.1016/J.CAM.2013.08.028).
- [47] H. A. van der Vorst, Bi-cgstab: A fast and smoothly converging variant of bi-cg for the solution of nonsymmetric linear systems, *SIAM J Sci Statist Comput* 13 (2) (1992) 631–644. doi:[10.1137/0913035](https://doi.org/10.1137/0913035).
- [48] S. Wermeille, P. Bossart, In situ stresses in the mont terri region data compilation, Tech. Rep. 99-02, Institut Géotechnique SA St-Ursanne (November 1999).
- [49] C. Martin, G. Lanyon, Measurement of in-situ stress in weak rocks at mont terri rock laboratory, switzerland, *Int J Rock Mech Min Sci* 40 (7) (2003) 1077–1088, special Issue of the IJRMMS: Rock Stress Estimation ISRM Suggested Methods and Associated Supporting Papers. doi:[https://doi.org/10.1016/S1365-1609\(03\)00113-8](https://doi.org/10.1016/S1365-1609(03)00113-8). URL <https://www.sciencedirect.com/science/article/pii/S1365160903001138>

- [50] A. A. Chaudhry, J. Buchwald, T. Nagel, Local and global spatio-temporal sensitivity analysis of thermal consolidation around a point heat source, *Int J Rock Mech Min Sci* 139 (June 2020) (2021) 104662. doi:10.1016/j.ijrmms.2021.104662.
URL <https://doi.org/10.1016/j.ijrmms.2021.104662>
- [51] J. Buchwald, OliverPe, T. Fischer, D. Kern, D. Y. Naumov, O. Kolditz, *joergbuchwald/ogs6py*: (Nov. 2021). doi:10.5281/zenodo.5669262.
URL <https://doi.org/10.5281/zenodo.5669262>
- [52] J. Buchwald, D. Kern, *joergbuchwald/vtuinterface*: (Nov. 2021). doi:10.5281/zenodo.5669428.
URL <https://doi.org/10.5281/zenodo.5669428>
- [53] GPy, GPy: A gaussian process framework in python (2014).
URL <http://github.com/SheffieldML/GPy>
- [54] S. Seabold, J. Perktold, *statsmodels: Econometric and statistical modeling with python*, in: 9th Python in Science Conference, 2010, pp. 92–96.
- [55] H. Bock, RA Experiment: Updated Review of the Rock Mechanics Properties of the Opalinus Clay of the Mont Terri URL based on Laboratory and Field Testing, Mont Terri Technical Report TR 08-04, Federal Office of Topography (swisstopo), Wabern, Switzerland (2009).
- [56] P. Bossart, M. Thury, Characteristics of the Opalinus Clay at Mont Terri, Reports of the Swiss Geological Survey no. 3 (2008) 26.
- [57] W. Wang, H. Shao, T. Nagel, O. Kolditz, 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* 142 (February) (2021) 104683. doi:10.1016/j.ijrmms.2021.104683.
URL <https://doi.org/10.1016/j.ijrmms.2021.104683><https://linkinghub.elsevier.com/retrieve/pii/S1365160921000691>
- [58] M. M. Tamizdoust, O. Ghasemi-Fare, A fully coupled thermo-poro-mechanical finite element analysis to predict the thermal pressurization and thermally induced pore fluid flow in soil media, *Comput Geotech* 117 (2020) 103250. doi:10.1016/J.COMPGE0.2019.103250.
- [59] D. M. Seyedi, C. Plúa, M. Vitel, G. Armand, J. Rutqvist, J. Birkholzer, H. Xu, R. Guo, K. E. Thatcher, A. E. Bond, W. Wang, T. Nagel, H. Shao, O. Kolditz, Upscaling THM modeling from small-scale to full-scale in-situ experiments in the Callovo-Oxfordian claystone, *Int J Rock Mech Min Sci* 144 (2021) 104582. doi:<https://doi.org/10.1016/j.ijrmms.2020.104582>.
URL <https://www.sciencedirect.com/science/article/pii/S1365160920309485>
- [60] W. Wagner, H.-J. Kretzschmar, *Iapws industrial formulation 1997 for the thermodynamic properties of water and steam*, International steam tables: properties of water and steam based on the industrial formulation IAPWS-IF97 (2008) 7–150.
- [61] S. Chatterjee, A. Hadi, *Regression Analysis by Example*, Wiley Series in Probability and Statistics, Wiley, 2006.
- [62] M. D. Morris, Factorial sampling plans for preliminary computational experiments, *Technometrics* 33 (2) (1991) 161–174.
URL <http://www.jstor.org/stable/1269043>
- [63] C. F. Tsang, I. Neretnieks, Y. Tsang, Hydrologic issues associated with nuclear waste repositories, *Water Resour Res* 51 (2015) 6923–6972. doi:10.1002/2015WR017641.
URL <https://onlinelibrary.wiley.com/doi/full/10.1002/2015WR017641><https://onlinelibrary.wiley.com/doi/abs/10.1002/2015WR017641><https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2015WR017641>
- [64] J. Gonçalves, G. de Marsily, J. Tremosa, Importance of thermo-osmosis for fluid flow and transport in clay formations hosting a nuclear waste repository, *Earth Planet Sci Lett* 339-340 (2012) 1–10. doi:10.1016/J.EPSL.2012.03.032.