About the Trustworthiness of Physics-Based Machine Learning – Considerations for Geomechanical Applications

Denise Degen[1,2], Moritz Ziegler[3,2], Oliver Heidbach[2,4], Andreas Henk[5], Karsten Reiter[5], and Florian Wellmann[1,6]

[1]Computational Geoscience, Geothermics and Reservoir Geophysics (CG3), RWTH Aachen University, 52074 Aachen, Germany
[2]GFZ Helmholtz Centre for Geosciences, 14473 Potsdam, Germany
[3]Professorship of Geothermal Technologies, TU Munich, Arcisstr. 21, 80333 Munich, Germany
[4]Institute of Applied Geosciences, TU Berlin, 10587 Berlin, Germany
[5]Institute of Applied Geosciences, TU Darmstadt, 64287 Darmstadt, Germany
[6]Fraunhofer Research Institution for Energy Infrastructures and Geothermal Systems (IEG), 44801 Bochum, Germany

Correspondence: Denise Degen (denise.degen@cgre.rwth-aachen.de)

Abstract. Model predictions are important to assess the subsurface state distributions (such as the stress), which are essential to, for instance, determine the location of potential nuclear waste disposal sites. Providing these predictions with quantified uncertainties often requires a large number of simulations, which is difficult due to the high CPU time needed. One possibility for addressing the computational burden is to use surrogate models. Purely data-driven approaches face challenges when operating in data-sparse application fields such as geomechanical modeling or producing interpretable models. The latter aspect is critical for applications such as nuclear waste disposal, where it is essential to provide trustworthy predictions. To overcome the challenge of trustworthiness, we propose the usage of a novel hybrid machine learning method, namely the non-intrusive reduced basis method as a surrogate model. This method resolves both of the above challenges while being orders of magnitude faster than classical finite element simulations. In the paper, we demonstrate the usage of the non-intrusive reduced basis method for 3-D geomechanical-numerical modeling with a comprehensive sensitivity assessment. The usage of these surrogate geomechanical models yields a speed-up of six orders of magnitude while maintaining global errors in the range of less than 0.01%. Because of this enormous reduction in computation time, computationally demanding methods such as global sensitivity analyses, which provide valuable information about the contribution of the various model parameters to stress variability, become feasible. The opportunities of these added benefits are demonstrated with a benchmark example and a simplified study for a siting region for a potential nuclear waste repository in Nördlich Lägern (Switzerland).

1 Introduction

Knowledge of the crustal stress field is of key importance for the safe usage of the subsurface, for example, for geothermal exploitation or storage of energy, CO_2 and nuclear waste (Blöcher et al., 2018; Henk, 2008; Hergert et al., 2015; Smart et al., 2014). The importance is expressed in the unwanted release of stored elastic energy during anthropogenic utilization by means of failure of boreholes (Schmitt et al., 2012; Tingay et al., 2008), caverns and tunnels (Brady and Brown, 2006), subsidence or induced seismic events (Ellsworth, 2013; Segall and Fitzgerald, 1998; Ziegler et al., 2015). Thus, to develop strategies that reduce the risks of induced hazards and prevent failure due to man-made interventions in the subsurface, it is important to understand the undisturbed 3-D in-situ stress state (Gaucher et al., 2015). However, the in-situ stress state is challenging to quantify. Stress data are rare, often subject to large uncertainties, and describe, in most cases, only a subset of the six independent components of the symmetric 2^{nd} rank tensor that formally describes the stress state at an arbitrary point in the subsurface (Amadei and Stephansson, 1997; Desroches et al., 2021a; Heidbach et al., 2018; Morawietz et al., 2020).

To obtain a continuous description of the 3-D stress tensor in a given rock volume, geomechanical-numerical models are employed. These models usually use the finite element method to solve the partial differential equation that describes the equilibrium of forces (e.g. Ahlers et al., 2021; Fischer and Henk, 2013; Lecampion and Lei, 2010; Reiter and Heidbach, 2014; Singha and Chatterjee, 2015; van Wees et al., 2018). For the model input, we have to describe the rock properties, the boundary conditions, and the subsurface geological structures. However, knowledge about these parameters is usually incomplete and consequently associated with uncertainties (e.g. Hergert et al., 2015; Ziegler and Heidbach, 2020). This means that for providing reliable model predictions the information regarding these parameter variabilities needs to be included and assessed.

Such tasks are commonly achieved through global sensitivity analyses (SA, Degen et al., 2021b, a; Saltelli et al., 2019) and uncertainty quantification methods (Degen et al., 2022c, a; Saltelli et al., 2019). Both methods have the requirement of numerous model evaluations in common, which poses major challenges when each model evaluation is computationally costly. One way to circumvent the issue is the use of surrogate models (also referred to as meta or reduced models), i.e., low-order representations of the original model that are significantly faster to compute. Many surrogate model construction techniques exist, ranging from physics-based to data-driven approaches (e.g. Benner et al., 2015; Degen et al., 2023; Hesthaven et al., 2016; Jordan and Mitchell, 2015; Kotsiantis et al., 2007; Mahesh, 2020). Nonetheless, every surrogate model needs to be evaluated concerning its trustworthiness and explainability. This is important for several reasons: i) if we use the surrogate in subsequent analyses such as global sensitivity analyses we need to ensure that the surrogate represents the original model otherwise the obtained sensitivities are not representative, ii) if the model results are used for decision-making processes, they need to be reliable and robust.

Different techniques exist for the construction of surrogate models, generally subdivided into three categories. In the following we briefly present the various techniques, explaining their key advantages and disadvantages.

1. One class of techniques are data-driven machine learning approaches, which recently gained attention in the construction of surrogate models (e.g. Bergen et al., 2019; Degen et al., 2023; Li et al., 2023; Swischuk et al., 2019; Willcox et al., 2021). This is thanks to their capabilities of well approximating nonlinear applications, their straightforward usage because of their black box and non-intrusive nature (meaning that they do not need direct access to the numerical solver), and their availability in software frameworks such as PyTorch (Paszke et al., 2019) and TensorFlow (Abadi et al., 2015). Nonetheless, the black box nature has also the disadvantage of yielding non-explainable models that do not preserve the governing physical equations, making their utilization challenging in terms of reliability.

2. Physics-based methods such as the reduced basis method or the proper orthogonal decomposition have the advantage of preserving the physics (Benner et al., 2015; Degen et al., 2020; Hesthaven et al., 2016; Quarteroni et al., 2015; Rozza et al., 2008). The reduced basis method has been developed in the field of applied mathematics and has been used for various applications, such as transport and continuum mechanics (e.g. Benner et al., 2015; Hesthaven et al., 2016; Quarteroni et al., 2015; Rozza et al., 2008) and also for large-scale geothermal applications (Degen et al., 2021b, 2022c). However, they reach their limits in efficiently approximating highly nonlinear problems (Degen et al., 2023; Hesthaven and Ubbiali, 2018; Wang et al., 2019, see also Section 2).

3. To overcome the limitations of both individual approaches, physics-based machine learning methods are introduced, which combine data-driven and physics-based techniques. Several physics-based machine learning methods are available and they have different implications concerning the question of explainable surrogate models (e.g., Degen et al., 2023; Faroughi et al., 2022; Willard et al., 2022). In this study, we use the non-intrusive reduced basis (NI-RB) method (Hesthaven and Ubbiali, 2018; Swischuk et al., 2019) for the construction of the surrogate models for geomechanical applications and demonstrate how it fulfills the criteria of explainable and reliable surrogate models, in contrast to other physics-based machine learning techniques.

This paper focuses on two main aspects: i) the combined consideration of different sources of uncertainty such as stress magnitude data records used for calibration, material properties, and subsurface geometry, and ii) the consideration of rapid changes in the state distribution. Especially the later part distinguishes the work significantly from previous studies in geothermal applications (e.g., Degen et al., 2022a), where only smooth state variable distributions of the pore pressure and/or temperature have been considered so far. Depending on the permeability contrast, the pore pressure exhibits rapid changes as well. However, these scenarios have not yet been investigated with respect to the non-intrusive reduced basis method. Following a benchmark example the workflow is applied to a data set from northern Switzerland, where an underground repository for nuclear waste is planned.

2 Background and Methods

In the following, we present the governing equations for geomechanical-numerical modeling presented in this manuscript. Afterwards, we introduce the non-intrusive reduced basis method, which is responsible for constructing the surrogate models. Additionally, we explain the concept of global sensitivity analyses, which are used to investigate the influence of the model parameters.

2.1 Subsurface stress state

The stress state in the subsurface is usually described by the symmetric 2$^{\text{nd}}$ rank stress tensor σ_{ij}

$$\sigma_{ij} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{pmatrix} \tag{1}$$

with six independent components – the three normal stresses σ_{xx}, σ_{yy}, and σ_{zz} as well as three shear stresses $\sigma_{xy} = \sigma_{yx}$, $\sigma_{yz} = \sigma_{zy}$, and $\sigma_{zx} = \sigma_{xz}$. Commonly, instead of the full stress tensor, the stress state is referred to in its main axis system which is a rotation of the stress tensor in a way that all shear stresses dissipate and only normal stresses remain so that

$$S = \begin{pmatrix} S_1 & 0 & 0 \\ 0 & S_2 & 0 \\ 0 & 0 & S_3 \end{pmatrix}. \tag{2}$$

with the principal stress components S_1, S_2, and S_3 that are perpendicular to each other but arbitrarily oriented in space. A common assumption in upper crustal geomechanics is that one of the principal stress axes is vertical as a result of the overburden which can then be referred to as S_v (Zoback, 2007). The other two principal stress axes are then by definition horizontal and called the maximum and minimum horizontal stress S_{Hmax} and S_{hmin}, respectively. Then the stress state can be fully described by four variables only: The magnitudes of S_v, S_{Hmax}, and S_{hmin} as well as the orientation of one of the two horizontal stress components. This is then called the reduced stress tensor (Zoback, 2007).

Information on the orientation of S_{Hmax} is available from a variety of stress indicators that are e.g. compiled in the World Stress Map (Heidbach et al., 2018). Stress magnitude data is less frequently available, sparse, and often of little quality (Morawietz et al., 2020). However, in particular, the stress magnitudes are important for many subsurface applications. If detailed information on the in-situ stress state is required, 3-D geomechanical-numerical modeling is applied in order to estimate the stress state in a volume of interest based on few data records (e.g. Fischer and Henk, 2013; Lecampion and Lei, 2010; Singha and Chatterjee, 2015; van Wees et al., 2018; Ziegler et al., 2016).

2.2 Governing Equations

The modeling of the in-situ stress state is conducted under the assumption of a linear elastic upper crust as the governing constitutive equation (Reiter and Heidbach, 2014; Hergert et al., 2015; Singha and Chatterjee, 2015) and no acceleration except for gravity. The required material properties are thus the Young's modulus (E), the Poisson's ratio (ν), and the density (ρ). We derive the total stress σ from the momentum balance (Cacace and Jacquey, 2017):

$$\nabla \cdot \sigma + \rho g = 0, \tag{3}$$

where ρ is the density, and g the gravity acceleration. The numerical forward simulations are performed in GOLEM (Cacace and Jacquey, 2017), which is an open-source high-performance finite element software based on the MOOSE framework (Lindsay et al., 2022). The result is the full stress tensor σ_{ij} at discrete locations throughout the model volume. This information is used to derive scalar values such as individual components from the reduced stress tensor or the slip tendency of faults (Röckel et al., 2022). Furthermore, we consider the following constitutive linear elastic relationships:

$$E = \frac{\sigma}{\varepsilon_{axial}}, \qquad \text{and} \qquad \nu = -\frac{\varepsilon_{trans}}{\varepsilon_{axial}}. \tag{4}$$

Here, ε_{axial}, and ε_{trans}, are the axial and transverses strain, respectively.

The model geometry is usually oriented in a way that the boundaries are parallel and perpendicular to the orientation of S_{Hmax} and S_{hmin} (Fig. 1). The stress state is introduced to the model using displacement boundary conditions (Dirichlet–type) on the lateral boundaries of the model. The magnitude of displacements is adapted in a way that the resulting stress magnitudes (S_{Hmax} and S_{hmin}) match observed data records (Reiter and Heidbach, 2014; Ziegler and Heidbach, 2020). This process is referred to as the calibration of the model. It is of an iterative nature or aided by the software tool FAST Calibration (Ziegler et al., 2023).

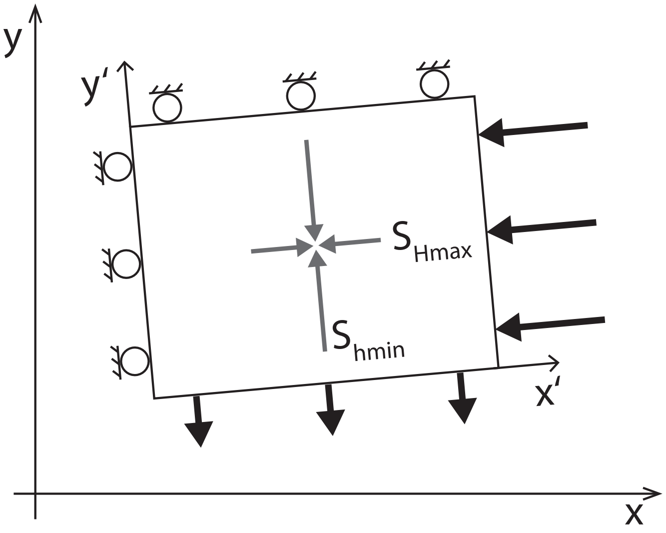

Figure 1. The model setup with the Cartesian model coordinate system (x and y axes) and the coordinate system used for application of boundary conditions (x´ and y´ axes) perpendicular to the model boundaries and the orientations of S_{Hmax} and S_{hmin}, respectively. The boundary conditions (Dirichlect-type) are indicated by the bold arrows (prescribed displacements) and rollers that indicate a zero displacement perpendicular to the model boundary.

2.3 Non-Intrusive Reduced Basis Method

In this work, we evaluate the trustworthiness of geomechanical surrogate models for use in applications such as nuclear waste disposal. Note that we focus on the construction of surrogate models for the S_{hmin}, S_{Hmax}, and S_{v} components. For better illustration of the general concepts, we assume that the material properties are homogeneous and isotropic within each layer.

However, the presented concepts are not restricted to these assumptions. Further details regarding the model setup are listed in section 3. For the purpose of the surrogate model construction, we use the non-intrusive reduced basis (NI-RB) method to construct the surrogate models (Hesthaven and Ubbiali, 2018; Swischuk et al., 2019; Wang et al., 2019). The NI-RB method combines physics-based and data-driven approaches to provide reliable predictions even for complex nonlinear and hyperbolic partial differential equations (PDEs).

The NI-RB method is most advantageous in the *many-query* or *real-time* context, so if either many and/or fast model evaluations are required (Benner et al., 2015; Hesthaven and Ubbiali, 2018; Hesthaven et al., 2016). To ensure this, the method is divided into two stages: the offline and online stages, as illustrated in Fig. 2. During the offline stage, the training data is generated and the surrogate model is constructed. This stage is computationally expensive since it requires several full-order solutions of the model, as we explain in the next paragraph. However, it needs to be performed only once. On the other hand, during the following online stage, the surrogate model is evaluated. This is a computationally fast procedure allowing numerous evaluations of the surrogate model in a short amount of time (Benner et al., 2015; Hesthaven and Ubbiali, 2018).

2.3.1 Offline Stage

The construction of the surrogate model in the offline stage is also a two-step procedure, which is preceded by step zero where the model parameters for the full-order solutions are determined (step 0 in Fig. 2). In our case, we use a quasi-random Latin-Hypercube sampling strategy (Loh, 1996) for the generation of the 100 to 200 training snapshots with the parameter ranges provided in Tab. 1. These snapshots are combinations of different geometries, rock properties, and boundary conditions, as shown in Fig. 2.

In the first step (Fig. 2) of the offline stage, the basis functions of the surrogate model are calculated by performing a proper orthogonal decomposition (POD). They capture the characteristic physical behaviour of the model by regarding the models response to different properties provided by the snapshots. In order to ensure an efficient execution of both the POD and machine learning stage, we scale the input parameters as well as the data sets. The input parameters are transformed using z-score normalization, resulting in a mean of zero and a standard deviation of one The training data is scaled with a min-max scaling, taking the maximum and minimum values which yields a data set being distributed between zero and one. The basis functions correspond to the most influential singular vectors (Hesthaven and Ubbiali, 2018). For defining this, the energy criterion is used (Guo and Hesthaven, 2019; Swischuk et al., 2019):

$$\frac{\sum_{i=1}^{r}\lambda_i^2}{\sum_{i=1}^{N}\lambda_i^2} \leq \epsilon, \tag{5}$$

where r corresponds to the dimension of the reduced model, λ to the singular value, N to the total number of samples, and ϵ to the desired accuracy of the surrogate. Since ϵ is a user-defined value, the quality of the surrogate is adjustable to the application. In our case studies, we set the tolerance ϵ to maintain 10^{-4} % of the information content to ensure surrogate models that well represent the full-order solutions. This tolerance is application-specific and based on the experience of previous studies (Degen et al., 2023, 2022a, 2020).

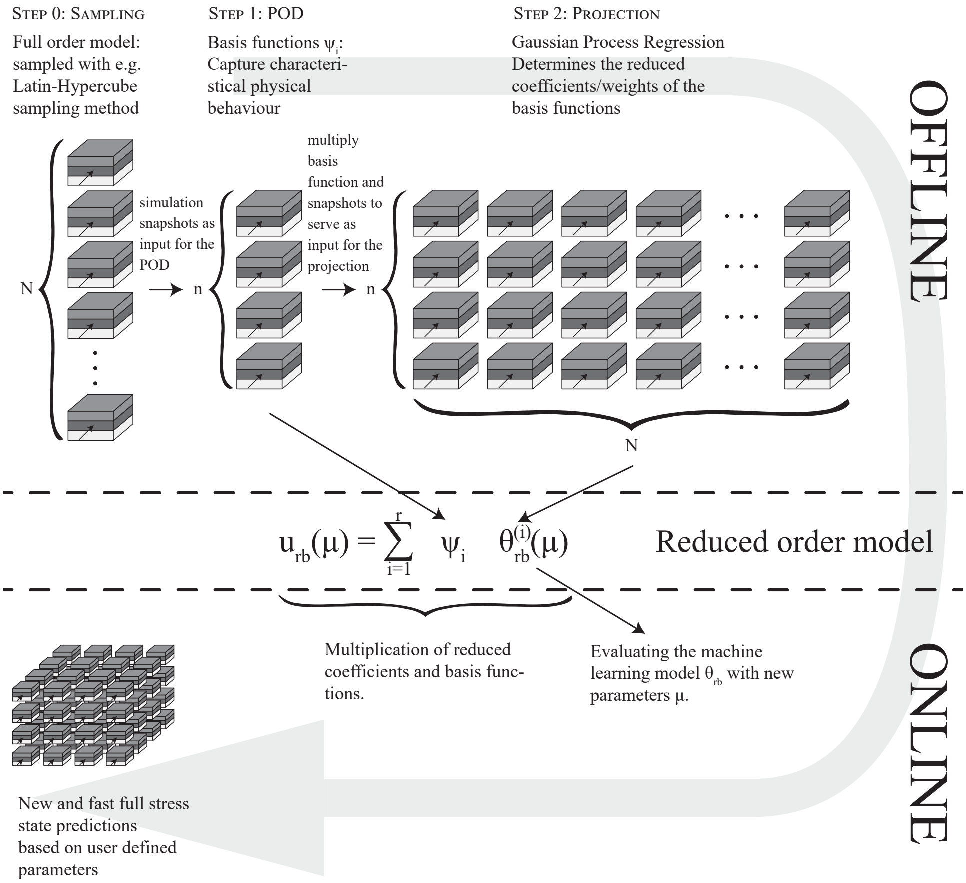

Figure 2. Schematic representation of the offline and online procedure for the NI-RB method, including the different steps required for the model construction in the offline phase. Note that u_{rb} denotes the reduced solution, μ the model parameters, r the reduced dimension, ψ the basis functions, θ_{rb} the reduced coefficients, N the total number of samples, and n the number of basis functions.

The resulting surrogate model u_{rb} is a linear combination of basis function ψ determined in step 1 and reduced coefficients θ_{rb}:

$$u_{\text{rb}}(\mu) = \sum_{i=1}^{r} \psi_i \, \theta_{\text{rb}}^{(i)}(\mu) \qquad \in V_{\text{rb}}. \tag{6}$$

Here, μ refers to the model parameters, and V_{rb} to the reduced space. The determination of these coefficients is the second step of the offline phase, which is classically performed by a Galerkin projection (Benner et al., 2015; Hesthaven et al., 2016; Quarteroni et al., 2015). However, to allow also for a later extension to nonlinear applications, we use in this study machine learning methods to determine the coefficients instead (step 2 in Fig. 2). This possible extension to the nonlinear setting is shown in several studies (Degen et al., 2022a; Hesthaven and Ubbiali, 2018; Swischuk et al., 2019). In the present study, we use Gaussian Process Regression (GPR, Schulz et al., 2018) as the machine learning method, which works better than neural networks in a linear setting since fewer hyperparameters need to be determined. We use anisotropic radial basis function kernels with an initial length scale of 1, which is optimized for each normalized model parameter μ within the ranges of 10^{-5} and 10^{5}. The GPR computations are executed with Scikit-learn (Pedregosa et al., 2011). The input for the GPR machine learning algorithm is the product of the basis functions and the training snapshots. The basis functions ψ provide the characteristic behavior of the model. The training snapshots are a controlled environment with known μ. This allows the GPR to derive the reduced coefficients θ_{rb} that complete the mapping between the input and output space since both are known for the training data.

2.3.2 Online Stage

Once the surrogate model is constructed, we have a flexible and fast performing replacement of the original high-dimensional problem, which can be used in subsequent analyses. Hence, new solutions for any desired combination of model parameters μ can be computed within the predefined parameter ranges (Fig. 2). To achieve this, we need to determine the corresponding reduced coefficients over the machine learning model and compute the associated solution by multiplying the coefficients and basis functions, as shown in Eq. 6. Note that only the coefficients have to be calculated for new combinations of model parameters. The basis functions remain the same throughout all realizations. This yields a great computational gain, allowing for rapid evaluations in, for instance, global sensitivity analyses or uncertainty quantification methods.

2.4 Global Sensitivity Analysis

The purpose of a sensitivity analysis (SA) is to investigate to which extent the model parameters influence the model response. This is helpful in determining which parameters to focus on in an in-depth study. In this study, the model parameters of interest are the value of the boundary conditions, the material properties, and the depth of the layer interfaces. The model response that we evaluate is the subsurface stress distribution.

In general two types of sensitivity analyses are distinguished: local and global analyses (e.g., Degen et al., 2021b; Degen and Wellmann, 2024; Sobol, 2001; Razavi and Gupta, 2015; Sarrazin et al., 2016; Song et al., 2015; Wainwright et al., 2014). The local SA has the advantage of requiring only very few model evaluations but the disadvantage of not considering parameter correlations, and investigating the influence of the model parameters only in the vicinity of a reference parameter. In contrast, a global SA investigates the entire parameter space and can determine parameter correlations if, for instance, a variance-based method is chosen. However, this comes at the cost of requiring numerous model evaluations, which makes the method computationally costly (Degen et al., 2021b; Degen and Wellmann, 2024; Saltelli et al., 2019; Sobol, 2001; Wainwright et al., 2014). This is the reason, why we propose the use of surrogate models for global sensitivity analyses.

We use the variance-based Sobol sensitivity analysis with a Saltelli sampler (Saltelli, 2002; Saltelli et al., 2010; Sobol, 2001). For the analyses of the benchmark example, we investigate the influence of up to 13 parameters (Young's modulus, density, Poisson's ratio of all three layers, the values of the two displacement boundary conditions, and the two interface positions, see Figure 3). In the case of the simplified case study of Nördlich Lägern, we investigate the Young's modulus of each layer yielding 15 parameters. For each of these up to 15 parameters, we generate 2^{17} samples, which is equal to 131,072. Note that we always need 2^n samples per parameter to ensure the convergence of the sampler. n can be any positive number and is as mentioned before for this study equal to 17. This number was determined through convergence tests. Because of the used sampler, the total number of forward evaluations per SA is $(n \cdot (2D+2))$, where D is the number of parameters. Since we perform 21 global SAs for up to 15 parameters, we require a total of 50,069,504 model evaluations for all analyses combined. Further details regarding the SAs are presented in the individual result sections.

For the different global SAs, we present first-order and total-order sensitivity indices. First-order indices describe the influence arising from the model parameters themselves, which are determined by the variance of the model parameter divided by the total variance that regards all model parameters variability. Total-order contributions contain the influence of the model parameters and their correlations to other parameters. For the sensitivity analysis, we need to define a threshold above which parameters are deemed influential. There is no general way of determining this threshold. For the purpose of this study, we set it to 10^{-2} since values below this threshold would be difficult to validate against typical in-situ stress measurement accuracies (Desroches et al., 2021b; Martin, 2007; Morawietz et al., 2020). This value is also in accordance with threshold values typically employed in literature (Cosenza et al., 2013; Degen et al., 2021a, b; Degen and Wellmann, 2024; Sin et al., 2011; Tang et al., 2006; Vanrolleghem et al., 2015). The threshold is a dimensionless number, which is determined by the division of two variances, for further details regarding the threshold itself and its determination, we refer to Degen and Wellmann (2024).

3 Synthetic Model

We consider a three-layer model to investigate systematically the potential of the NI-RB method to construct reliable surrogate models for representing the subsurface stress distribution. The model has an extent of 10 km in both the x- and y-direction and six kilometers in z-direction (Fig. 3). For the discretization, we consider hexahedral elements. In both the x- and y-direction, we have a model resolution of 400 m and in the z-direction a resolution of 10 m. The much higher resolution of the vertical component is chosen to allow considering geometrical uncertainties, i.e. uncertainties in the depth of geological horizons. The three layers are lateral homogenous and equally spaced, where the top layer consists of limestone, the middle layer of clay, and the lowest layer is a crystalline unit. Thus, S_{v}, S_{Hmax}, and S_{hmin} are indeed the principal stresses of the modeled total stress.

Throughout the study, we allow variations for the boundary conditions, the material properties, and the interface depths. For the displacement in the y-direction (northern boundary), we apply a Dirichlet boundary condition with values varying between four to six meters shortening, whereas the variations for the x-direction range from 0.2 m to 0.6 m (eastern boundary). Different boundary conditions are needed to ensure reasonable variations of the stress magnitudes as observed in data records. The top boundary is assigned with a zero Neumann boundary condition, and all remaining boundaries are subjected to zero Dirichlet boundary conditions normal to the boundary (roller boundary conditions). The material properties are varied according to the expected uncertainties for the corresponding lithology. The depth of geological horizons can be varied by $\pm$ 200m.

We consider five cases: i) changing boundary conditions, ii) changing material properties, iii) changing interface positions, iv) simultaneously changing boundary conditions and material properties, and iv) changing all three sources of uncertainty.

Table 1. Variation ranges of the input parameters for all five cases. Note that $\| x$ and $\| y$ denote the displacement for the boundary condition in the x- and y-axis, respectively. Furthermore, “upper” denotes the interface between the limestone and clay layer, and “lower” the interface between the clay and crystalline layer. The parameter ranges have been chosen to be representative for nuclear waste disposal applications (Crisci et al., 2022a, b; Gonus et al., 2022a, b, 2023; Spillmann et al., 2022), the thickness data are from Crisci et al. (2022a). The Young’s modulus of the crystalline layers was extended beyond that range to account also for other subsurface engineering applications.

Layer	**E [GPa]**	**ν [-]**	**ρ [kg/m^3]**	**$\|$ x [m]**	**$\|$ y [m]**	**upper [m]**	**lower [m]**
Case 1: Changing Boundary Conditions							
Limestone	25	0.21	2600	0.2-0.6	4.0-6.0	-2000	-4000
Clay	15	0.34	2400				
Crystalline (Basement)	60	0.25	2800				
Case 2: Changing Material Properties							
Limestone	20-40	0.18-0.23	2500-2700	0.5	5.0	-2000	-4000
Clay	10-25	0.30-0.38	2300-2500				
Crystalline (Basement)	30-80	0.20-0.30	2700-2900				
Case 3: Changing Geometry							
Limestone	25	0.21	2600	0.5	5.0	-1800 to -2200	-3800 to-4200
Clay	15	0.34	2400				
Crystalline (Basement)	60	0.25	2800				
Case 4: Changing Boundary Conditions & Material Properties							
Limestone	20-40	0.18-0.23	2500-2700	0.2-0.6	4.0-6.0	-2000	-4000
Clay	10-25	0.30-0.38	2300-2500				
Crystalline (Basement)	30-80	0.20-0.30	2700-2900				
Case 5: Changing Boundary Conditions & Material Properties & Geometry							
Limestone	20-40	0.18-0.23	2500-2700	0.2-0.6	4.0-6.0	-1800 to -2200	-3800 to -4200
Clay	10-25	0.30-0.38	2300-2500				
Crystalline (Basement)	30-80	0.20-0.30	2700-2900				

The different scenarios including their input parameters are listed in Tab. 1. The uncertainties of the parameters are chosen in a way comparable to those encountered in datasets or case studies (e.g. Bär et al., 2020; Bond et al., 2015; Ziegler and Heidbach, 2024).

4 Results

In the following, we present the construction of a geomechanical surrogate model. To understand the different requirements and impacts of the different sources of uncertainties, we first vary them individually and then investigate the combined effects.

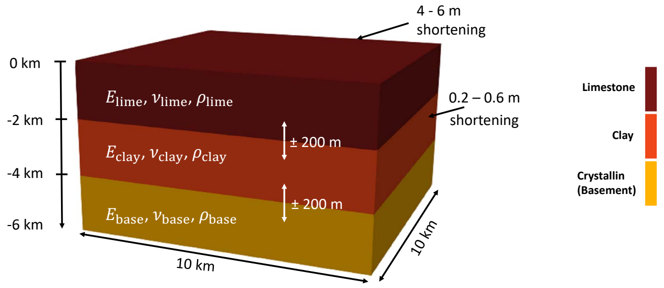

Figure 3. Schematic representation of synthetic model and all variations investigated in this study. The model ($10\times10\times6\,\text{km}^3$) consists of three different lithological units having purely elastic material properties, which are variable. Mechanical parameters are the Young's modulus (E), the Poisson's ration (ν) and the density (ρ). The horizontal stratigraphic boundaries can be varied by ± 200 m. The applied lateral boundary conditions ranges between 4 to 6 m shortening in y-direction, driving S_{Hmax} magnitude and orientation and 0.2 to 0.6 m shortening driving the S_{hmin} magnitude and orientation.

In order to validate the NI-RB results, for each test case randomly chosen parameters are used in several full-order model runs which are then compared to the NI-RB results.

4.1 Case 1: Changing Boundary Conditions

We first investigate the potential of using surrogate models to determine the influence of uncertainties in the stress magnitude data records that are available. This is achieved by changing the displacement boundary conditions in a way that corresponds to uncertainties in the stress magnitude data. Therefore, we use the model displayed in Fig. 3 and construct a training set of 100 simulations, where we allow a variation of the boundary condition but keep all the material properties and the interface positions fixed.

Since we only vary the Dirichlet boundary conditions along the x- and y-axis, we evaluate the surrogates only for the horizontal components of the stress tensor, which are displayed in Fig. 4a and b. We observe that both the S_{hmin}- and S_{Hmax}-contributions of the stress tensor are well represented by the surrogate model. The global model errors for the scaled training and validation data set are for the S_{hmin}-component in the order of 10^{-15} (about 10^{-12} MPa^2) and for the S_{Hmax}-component in the order of 10^{-8} (about 10^{-4} MPa^2), which means that visually no difference between the full and reduced solutions are detectable (Fig. 4). To calculate the global model errors, we use the mean squared error (MSE).

$$MSE(u_{fe}, u_{rb}) = \frac{1}{N}\sum_{i=0}^{N-1}(u_{fe} - u_{rb})^2, \tag{7}$$

where u_{fe} refers to the finite element solution.

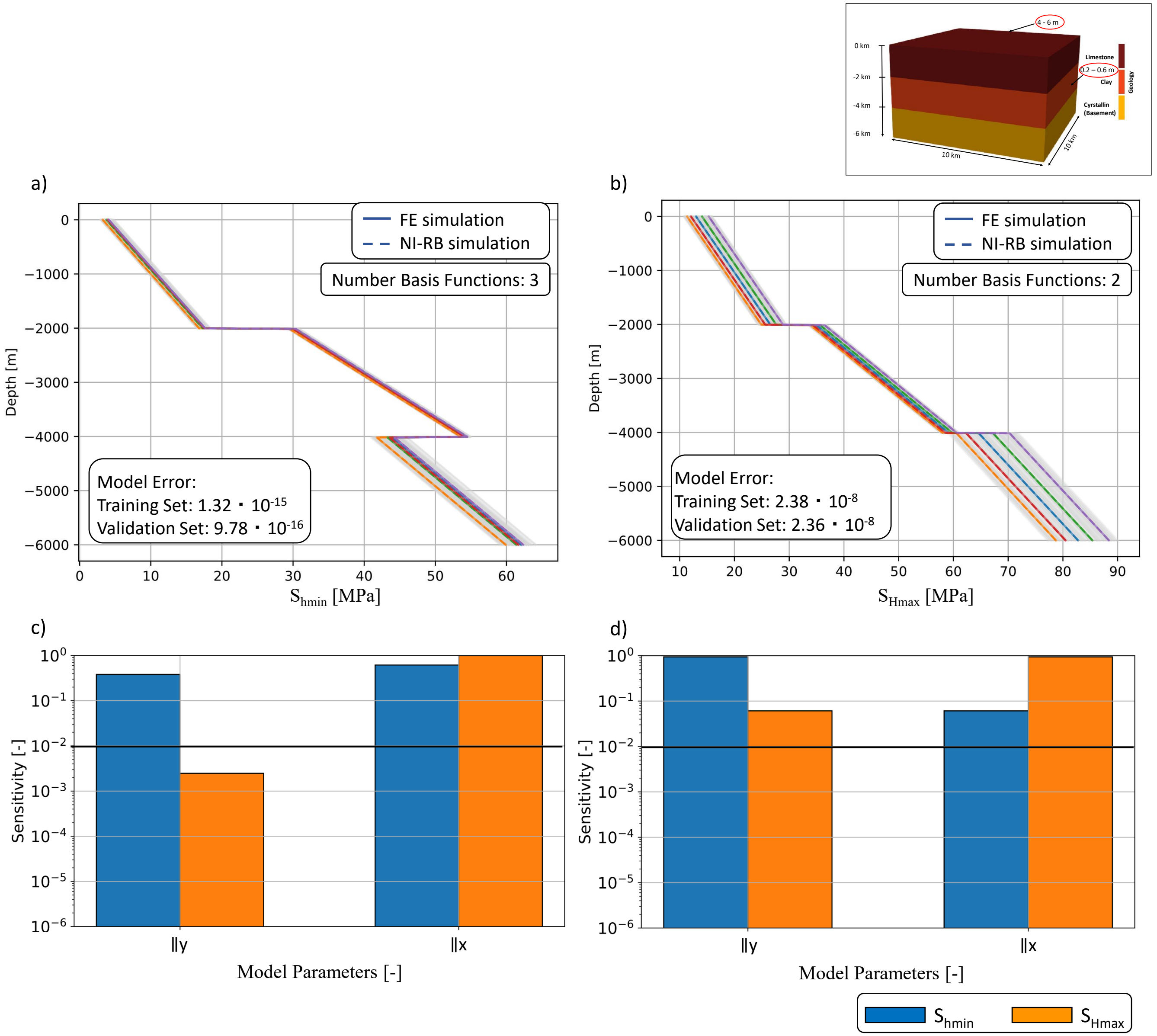

Figure 4. Investigation of the potential of surrogate models to determine the influence of varying boundary conditions. Comparison of the surrogate model accuracy for five randomly chosen realization of the validation data set for a) the S_{hmin}- and b) the S_{Hmax}-component of the stress tensor. Note that "Sensitivity" denotes the value of the sensitivity index (section 2.4), and $\parallel x$ and $\parallel y$ denote displacements parallel to the x-axis and y-axis, respectively. The full-order solution are denoted by colored solid lines and the reduced solutions by colored dashed lines. The different colors represent the stress response for different boundary condition values. The training data set consisting of 100 samples is indicated with light gray lines. Furthermore, we show the global SA results for c) non-equal x- and y-strains d) equal x- and y-strains.

The results for the surrogates are obtained with only a few basis functions: two and three for the S_{Hmax}- and S_{hmin}-contribution, respectively. This demonstrated a general low complexity induced by the changes in model parameters and serves as a good illustration of how the NI-RB method operates. In contrast to other machine learning techniques, the method does not try to learn the state behavior of the stress directly. Instead, it investigates the changes in the state distribution because of different model parameters. In this example, this change yields a simple shift in the response, where the amount of shift varies for the different geological layers. This is also the reason why the abrupt shift in stress magnitude at interfaces is perfectly captured in all reduced solutions since only the magnitude of the shift but not the position changes.

In general, we observe the highest variability of the stress response for the S_{Hmax}-component. This is caused by the boundary conditions since we assigned higher displacement values parallel to the x-axis than parallel to the y-axis. The global SA, displayed in Figure 4c and d, confirms these findings.

In Figure 4c, we observe a significantly higher influence of the boundary condition applying a displacement parallel to the x-axis for both stress components. In the case of the S_{Hmax}-component the boundary condition along the y-axis does not impact the response notably. To offer a final proof, we repeated the surrogate model construction and the sensitivity analysis, considering equal parameter ranges of four to six meters for both boundaries. This analysis with equal parameter ranges is referred to as “equal strains”, whereas the original setup is denoted as “non-equal strains”. The results of the corresponding SA are seen in Figure 4d. In comparison to the previous analysis, we obtain increasing influences of the boundary conditions along the y-axis. This yields precisely mirrored behaviors for the S_{Hmax}- and S_{hmin}-components.

4.2 Case 2: Changing Material Properties

We now investigate the variation of material properties. In this geomechanical study, we consider the variations of the Young’s modulus, the Poisson’s ratio, and the density for each layer individually. This results in nine parameters that can change, whereas for the previous example of the boundary conditions, we only had two parameters. Therefore, we increase the size of the training data set from 100 to 200 full-order solutions to compensate for the expected increase in complexity.

Although we have an increase in the amount of parameters that we consider, we still fit the full solution very well with the reduced order model. Note, that we investigate the stress distribution for the S_{hmin}-, S_{Hmax}-, and S_v-component, which are displayed in Figure 5a to c. The S_{hmin}- and S_{Hmax}-contributions show similar behaviors regarding the global errors, where we obtain errors in the order of 10^{-9} for the scaled validation set (corresponding to 10^{-5} MPa^2), and up to 10^{-17} for the scaled training data set (corresponding to 10^{-14} MPa^2). The global errors for the S_v-component are in the order of 10^{-15} (about 10^{-10} MPa^2) for both the scaled training and validation data set. However, we observe an increase in the number of basis functions from two or three to six for the horizontal stress responses. This increase in the number of basis functions is caused by the increased complexity, which shows the general scaling behavior of the NI-RB method. The dimension of the surrogate model is scaling with the complexity induced by the variability of the parameters. But this does not mean that an increase in the amount of variable parameters automatically yields are larger reduced dimension. Relevant for the dimension is the amount of variable parameters that lead to a change in the model response, so in our case the stress distribution.

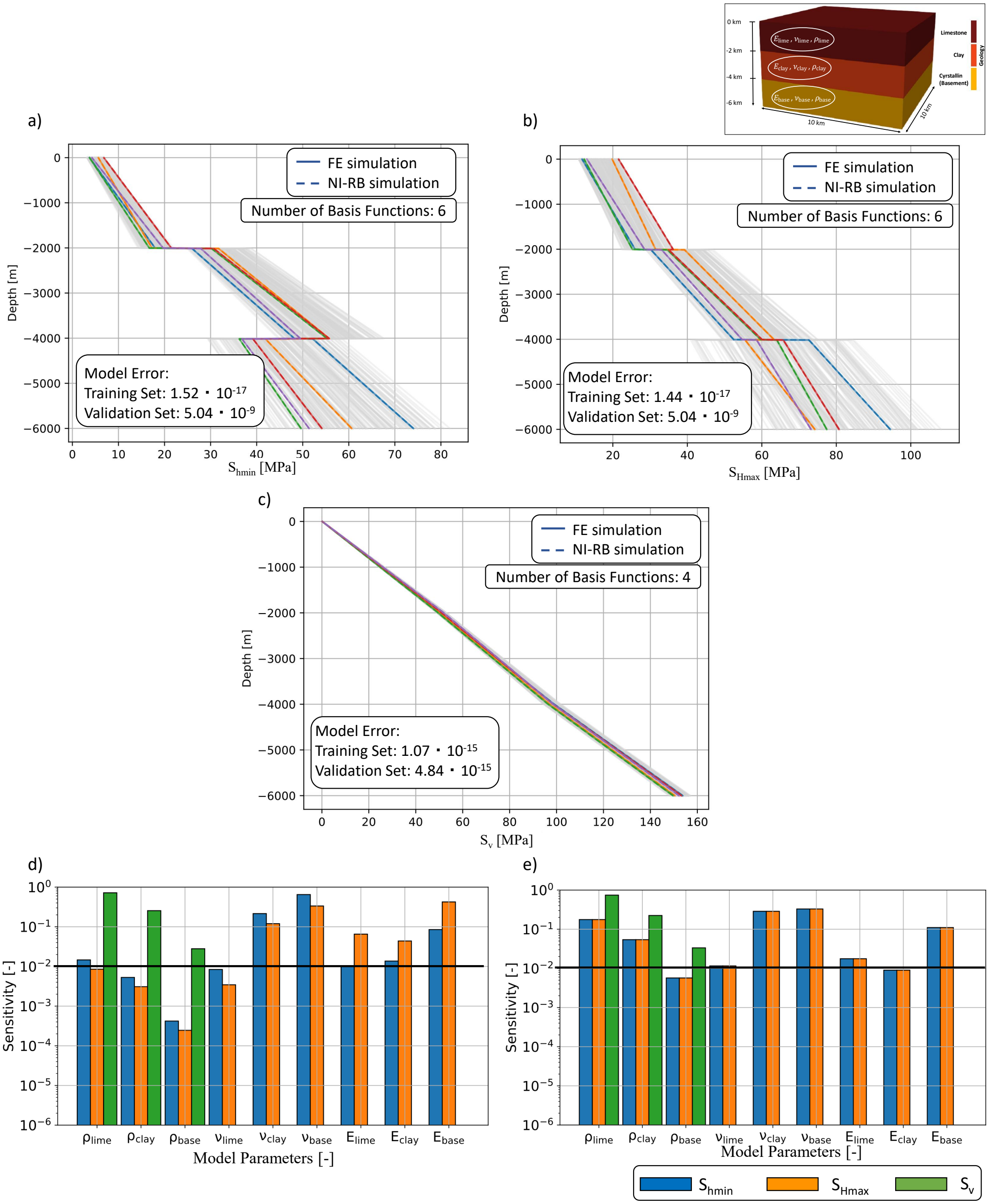

Figure 5. Investigation of the potential of surrogate models to determine the influence of varying material properties. Comparison of the surrogate model accuracy for five randomly chosen realization of the validation data set for a) the S_{hmin}-, b) the S_{Hmax}-, and c) the S_v-component of the stress tensor. Note that "Sensitivity" denotes the value of the sensitivity index (section 2.4). The full-order solution are denoted by colored solid lines and the reduced solutions by colored dashed lines. The different colors represent the stress response for different material property values. The training data set consisting of 200 samples is indicated with light gray lines. Furthermore, we show the global SA results for d) non-equal x- and y-strains and material parameter ranges and e) equal x- and y-strains and material parameter ranges.

This is nicely demonstrated by comparing the number of basis functions of the S_{hmin}- or S_{Hmax}-contribution with the S_v-component. For the vertical stress response, we only obtain four basis functions, which is the first indication that fewer parameters influence the stress response. This is later also confirmed by the SA. Still, also six basis functions of the horizontal components are a low number taking into account that we vary in total nine parameters. So, as for the vertical component, this yields the assumption that not all parameters are influential with respect to the stress. As for the case of the changing boundary conditions, we obtain the highest variability for the S_{Hmax}-component of the stress distribution. This is again caused by the higher boundary value of the displacement parallel to the x-axis. Both the results regarding the variability of the stress distribution and the amount of basis function are confirmed by the sensitivity analysis, which we present in the following.

The results of the SA are shown in Figure 5d and e. Here, we investigate again two possible scenarios. The first scenario chooses the variation range of the material properties depending on the typically physically plausible variation range (Figure 5d). This has the consequence that certain parameters have a wider relative variation range than others. To recover the underlying process behavior, we conduct the same analysis with a variation range of 3 % for all parameters and a displacement value of 5 m for both boundaries (Figure 5e).

Focusing first on the results of the SA in Figure 5d for the S_{Hmax}-component of the stress (orange bars), we obtain the highest influences for the Young's modulus of the crystalline layer, as it is the stiffest unit. This is followed by the Poisson's ratios of the crystalline and clay layers, and then by the Young's moduli of the limestone and clay layers. All other model parameters are non-influential. Consequently, we have a high impact of the Young's modulus and the Poisson's ratio but no influence of the density. This low influence is the cause of the in relation much lower variation range of the density. Assuming equal variation ranges for all model parameters, as seen in Figure 5e, we obtain the increased influence of the density especially for the limestone and clay layers. At the same time, we retrieve decreasing importance for the Young's modulus, such that the density is ranked in between the Poisson's ratio and the Young's modulus.

Continuing with the S_{hmin}-component (blue bars in Figure 5d), we observe a higher influence of the Poisson's ratios and a significantly decreased impact of the Young's moduli in contrast to the S_{Hmax}-component. The reason is again the different boundary values for the displacement. The differences between the two horizontal stress responses nearly vanish for the analysis with equal boundary conditions (Figure 5e).

For the S_v-component, we obtain in both scenarios only influences of the density, which demonstrates that the vertical stress component is predominantly driven by gravity, what corresponds to the definition. Furthermore, we have increasing influences of the density with decreasing depth values. Hence, we get the highest influence for the limestone layer, followed by the clay and crystalline layers, which is in agreement with a previous model study (Ziegler, 2022).

4.3 Case 3: Changing Geometry

For the scenario of a changing geometry, we consider two parameters, which are the interface depths between the limestone and clay layer (denoted as upper interface), and the interface between the clay and crystalline layer (denoted as lower interface). The upper interface is by default at a depth of two kilometers, and the lower interface is at a depth of four kilometers. For both interfaces a variation range of ± 200 m is considered. As before, the training data set is constructed using a Latin Hypercube

Sampling but for the geometry, we have a training data set size of again 200 full-order solutions instead of the initial 100 due to the expected higher complexity compared to Case 1.

In addition to a higher complexity comparable to Case 2 a change in interface depths means that the geometry of the model is changed, which is different from the material properties or boundary condition changes. The challenges for the surrogate modeling approach become apparent from Figure 6a. The surrogate model behaves differently in approximating the stress tensor compared to the previous two cases. The stress components (S_{hmin}-component visualized in Figure 6) are well approximated within the three layers such that again no visible differences are notable. However, at the transition between layers, the approximation quality is significantly decreased (marked with black ellipses). So, the surrogate has major issues resolving the position of the stress shift due to the changes in geological horizon depth. The difference to the previous examples is that the shift is no longer stationary and the number of elements per unit changes. In the case of changing boundary conditions and material properties only the magnitude of the shift changes. But for the geometrical parameters, the magnitude stays roughly the same but the position of the shift can vary by $\pm 200\,\mathrm{m}$.

To overcome this issue, the number of elements needs to remain the same for each lithology. This is achieved using GMSH (Geuzaine and Remacle, 2009) for discretization which allows for a higher flexibility during the construction of the structured meshes. Most importantly it allows for a fixed number of elements per lithology to, which is in our case 200. That means that the maximum vertical length of each element is $10\,\mathrm{m}$, but that also elements with smaller vertical length are produced depending on the interface location. This results in a better quality at the interface (Figure 6b).

These two representations of the geometry result in the same depth vs. stress relationship (Figure 6c) but for the surrogate model construction, we provide the information about the relationship of the number of elements vs. the stress distribution (Figure 6d). This relationship is different for the two scenarios. For scenario 1 (Figure 6a), we obtain the changing position of the shift. However, for scenario 2 (Figure 6b), we have the shift again stationary since we have always 200 elements per subdomain independent of the interface positions. Consequently, the challenge associated with the geometrical characterization can be addressed by reformulating the original problem.

The difference in behavior is also clearly visible by the amount of basis function required to approximate the solution. We obtain 79 basis functions for the S_{hmin}-component of the stress tensor, whereas the version with a fixed number of elements per subdomain only requires three basis functions. This means that we obtain 25 times more basis functions for the version with a changing number of elements per subdomain. Since the scaling behavior of the reduced model is tied to the complexity of the parameter space and not to the number of changing parameters, this demonstrates the change in complexity between the two model versions.

We now move to demonstrate the physical effects of changing the geometry, as presented in Figure 7a-c. Regarding the horizontal component, we obtain no difference within the layer, which is in accordance with our expectations since neither the material properties nor the boundary condition values are changed throughout the simulations. Differences occur at the interface positions, which vary from realization to realization. Consequently, the transition in the stress values between one unit to the other is changing dependent on the corresponding interface depth. Furthermore, we observe as in the previous example nearly no variations in the vertical component.

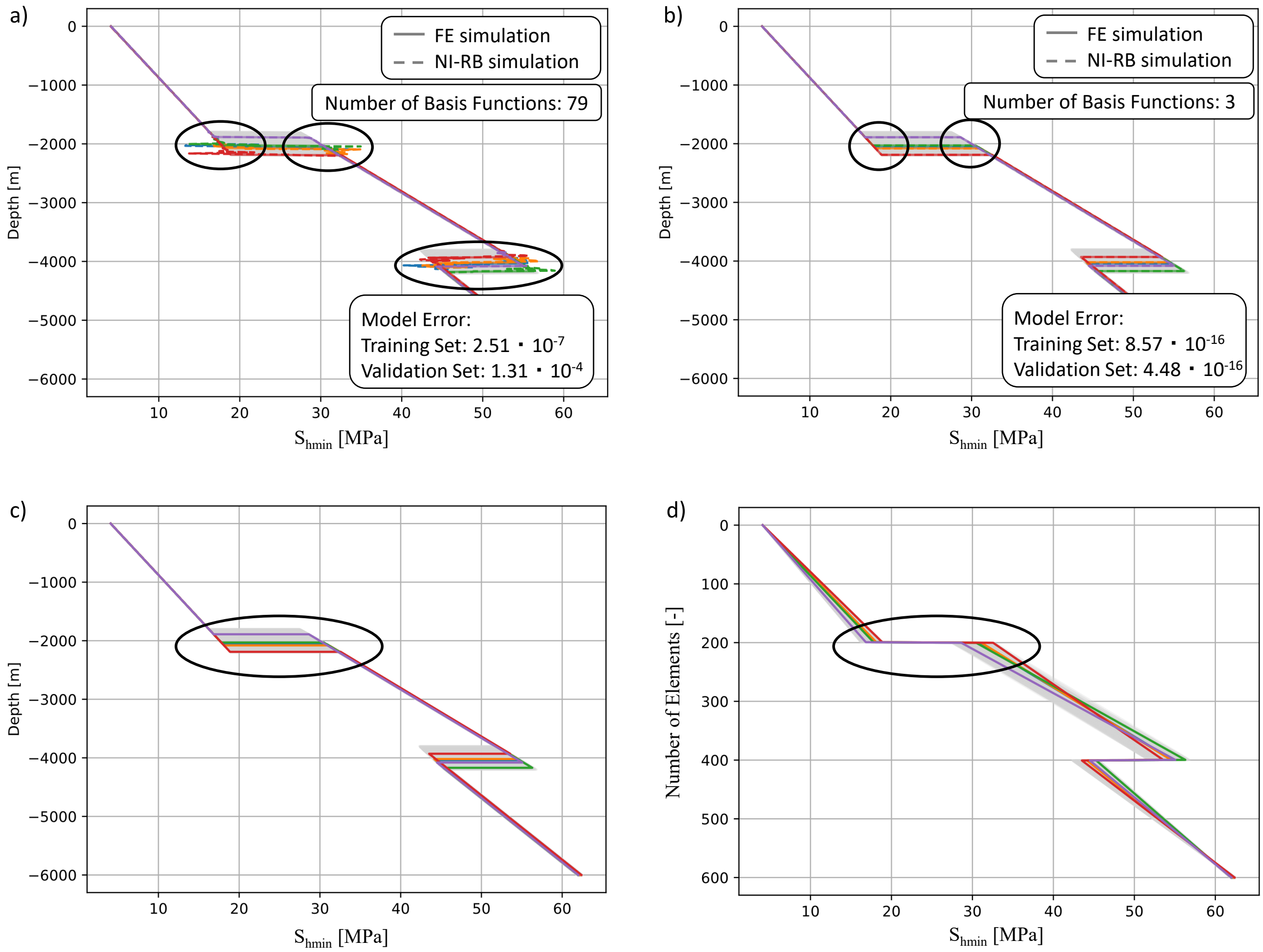

Figure 6. Challenge of accurately predicting the solution because of varying geometrical parameters. Comparison of the surrogate model accuracy for five randomly chosen realization of the validation data set for the S_{hmin}-component of the stress tensor for a) the GMSH model with varying number of elements per layer, b) the GMSH model with a fixed number of elements per layer. Furthermore, we display few realization from the validation data set over c) the depth, and d) the number of elements of the mesh.

Regarding, the sensitivity analysis (Figure 7c), we observe identical behavior for changes of both interface depths for the horizontal stress tensor components. For the vertical contribution, we observe a slightly higher influence of the location of the lower interface on the stress response. This higher influence is caused by the higher density contrast of the clay and crystalline layer with respect to the limestone and clay layers. Note that the lack of parameter correlations observable throughout all three scenarios is induced by the linearity of the application, this image significantly changes once nonlinear problems are considered.

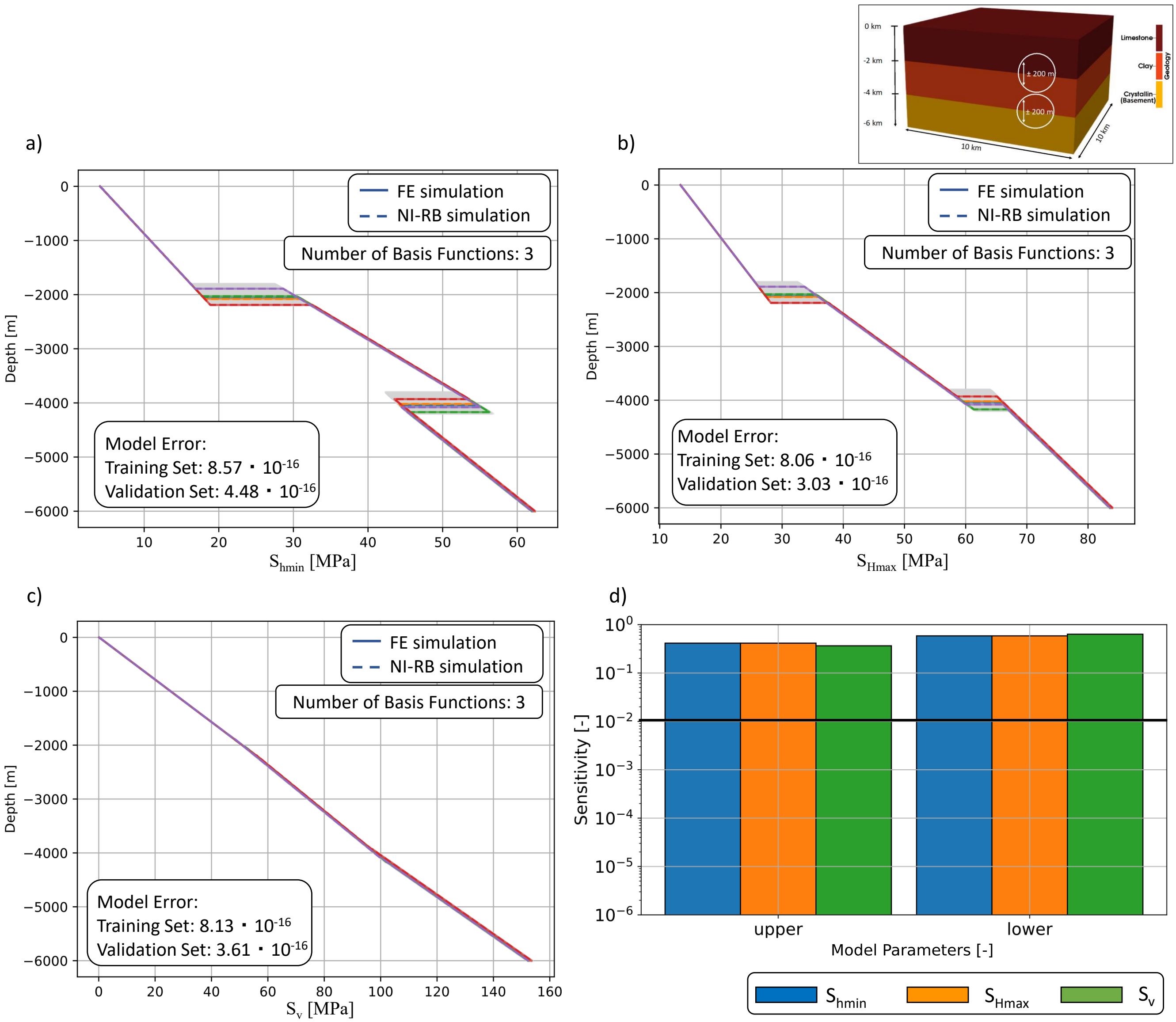

Figure 7. Investigation of the potential of surrogate models to determine the influence of varying interface depths. Comparison of the surrogate model accuracy for a) the S_{hmin}-, b) the S_{Hmax}-, and c) the S_v-component of the stress tensor. Note that "Sensitivity" denotes the value of the sensitivity index (section 2.4). The full-order solution are denoted by colored solid lines and the reduced solutions by colored dashed lines. The different colors represent the stress response for different interface depths. The training data set is indicated with light gray lines. Furthermore, we show in d) the global SA results.

4.4 Case 4: Simultaneously Changing Boundary Conditions and Material Properties

In the previous three scenarios (case 1-3), we investigated the behavior of the surrogate model individually for the three sources of variability considered in this manuscript. However, in a real-case application, it is unlikely that only one source of uncertainty is present. Typically all sources of uncertainty occur and can differ in their degree of variability. To evaluate which consequences this poses in terms of the surrogate, we consider two additional cases. In this section, we account for a simultaneous variation of both the boundary conditions and the material properties. The case of varying all three sources of uncertainty at the same time is presented in the next section.

Analyzing the combined effect of varying the boundary condition and the material properties is of interest for two main reasons: i) does it yield an increase in complexity and consequently an increase in the dimension of the surrogate, and ii) how is the SA affected by the increase in variable parameters?

Starting with the first aspect of complexity, we obtain for all three components of the stress tensor the same number of basis functions (Figure 8a-c) as for the case of changing only the material properties (case 2). So considering additionally the boundary conditions next to the material properties does not increase the complexity notable. This might seem counter-intuitive at first glance since we consider two additional parameters yielding a total amount of 11 parameters. The reason for this behavior is two-fold: i) the simplicity of the geological model, and ii) the threshold behavior of the POD. We use a relatively simple benchmark with horizontal horizons to enable a better understanding of the surrogate construction method. However, this also means that the model has an underlying low complex behavior which is mirrored in the low number of basis functions obtained. For more complex models at least a slight increase in the amount of basis functions is expected if the additional parameters influence the model response. That this is the case, we see later on in the SA. The second aspect is about the POD threshold. We defined an error tolerance of 10^{-4} % for the POD. This means that the POD can disregard at most 10^{-4} % of information content. So, it could happen that for the less complex scenario, the error decreases with the addition of the sixth basis function to 10^{-5} %, whereas it decreases only to $2 \cdot 10^{-4}$ % for the more complex scenario. In that case, both scenarios would have the same number of basis functions although the complexity impacts the accuracy of the model. Note, that this is not occurring in our example. Both the scenario of changing material properties and the scenario of varying material and boundary conditions have comparable accuracies. This means that here, the complexity is not increased by considering in addition to the material properties also the boundary conditions. However, this is caused by the simplicity of the model and will not occur for more complex studies. The reduced solution well captures the full-order response for all three stress tensor components and we retrieve similar global errors in the case of changing only the material properties.

Switching focus to the sensitivity analysis (Figure 8d), we see no correlations between the boundary values and the different material properties. This is expected since we consider a linear physical process. This would in general justify a separate consideration of the two sources of uncertainties. Still, a joint consideration has several advantages. The first benefit concerns generalizability. Only for linear applications, no correlations are expected, for nonlinear problems correlations are likely to occur. So, introducing the joint approach allows for a straightforward extension to nonlinear applications. Furthermore, even

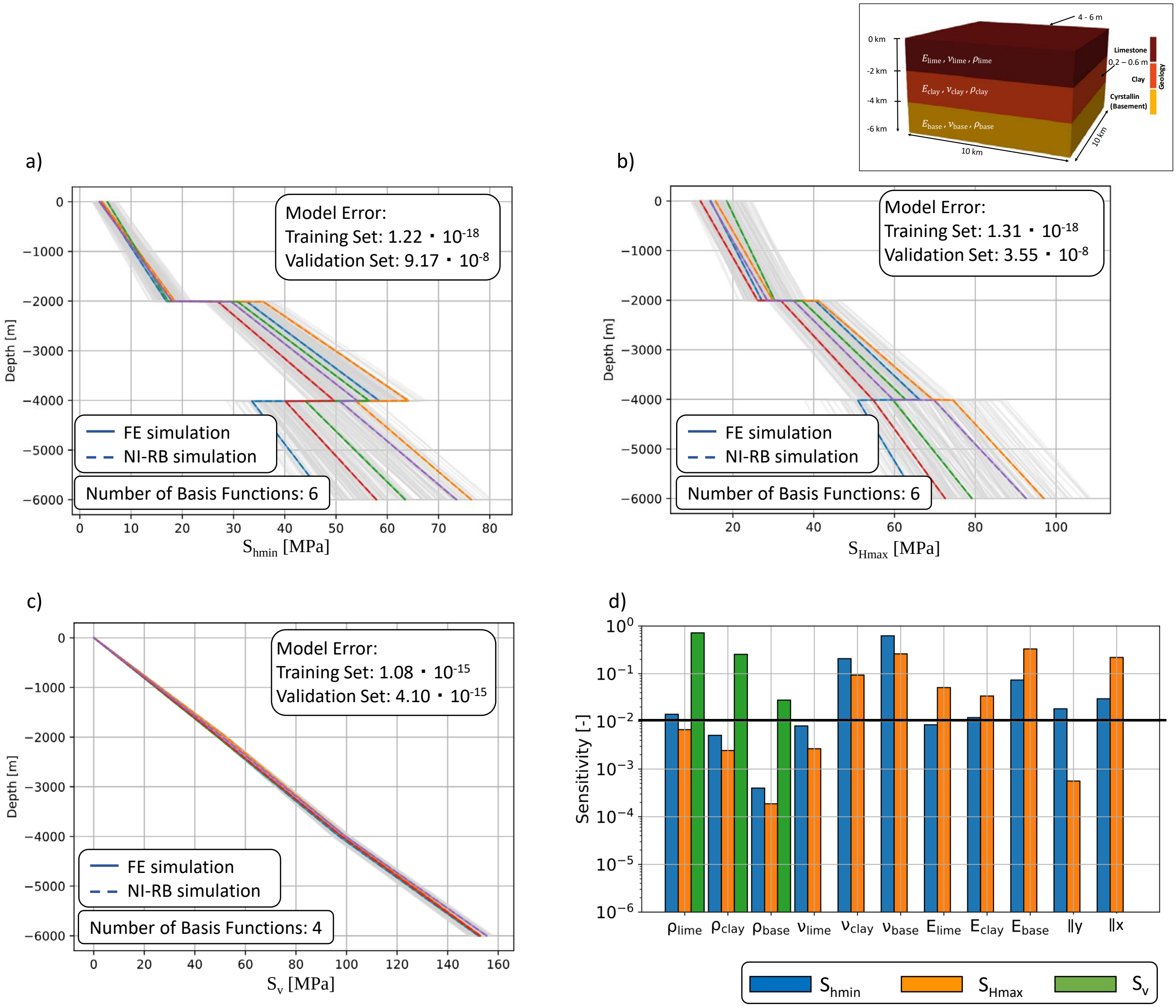

Figure 8. Investigation of the potential of surrogate models to determine the influence of varying boundary conditions and material properties. Comparison of the surrogate model accuracy for five randomly chosen realization of the validation data set for a) the S_{hmin}-, b) the S_{Hmax}-, and c) the S_v-component of the stress tensor. Note that "Sensitivity" denotes the value of the sensitivity index (section 2.4). The full-order solution are denoted by colored solid lines and the reduced solutions by colored dashed lines. The different colors represent the stress response for different boundary conditions and material properties. The training data set consisting out of 200 samples is indicated with light gray lines. Furthermore, we show in d) the global SA results.

for the linear case, the joint approach has the benefit of being able to determine which relative impact the boundary conditions have with regard to the different material properties.

This is seen in the SA results of Figure 8d, where we obtain a very similar result to the analysis considering only material properties. However, we can now determine the rank of the boundary condition values with respect to the Young's modulus, the Poisson's ratio, and the density. The boundary condition along the x-axis has a similar importance than the Young's modulus of the crystalline layer for both horizontal components of the stress tensor. The same is true of the other boundary condition in the case of the S_{Hmax}-component of the stress. For the S_{hmin}-component the boundary condition along the y-axis is non-influential with regard to the stress response. This matches the previously observed behaviors of the scenarios changing only the boundary conditions or the material properties. The S_v-component remains entirely gravity-driven and hence no influence of changes in the boundary conditions is detectable.

4.5 Case 5: Simultaneously Changing Boundary Conditions, Material Properties, and Geometry

Considering the case of simultaneously changing the boundary conditions (using the setup presented in Fig. 6b), material properties, and the depth of the interfaces, we first observe a slight increase in the number of basis functions compared to the case of only changing the geometry (case 3) and the same number of basis function as for the case of changing both the material properties and boundary conditions (case 4). Also, the global errors are with about 10^{-18} for the training data set and 10^{-7} for the validation data set for the horizontal components of the stress tensor of similar order of magnitude than the previous case (Figure 9a-c). Only the global error of the validation data set for the vertical stress component is with 10^{-12} slightly increased because of the higher complexity of the response. This higher complexity arises from the combined influence of the density and interface position as shown in the SA (Figure 9d).

Furthermore, the SA demonstrates that for S_{hmin}, we obtain only minor influences of the interface positions, and for S_{Hmax} we obtain negligible impacts. So for the given variation ranges and model geometry the material properties and boundary conditions have a higher impact than the geometrical parameters.

4.6 Case Study – Nördlich Lägern

So far, we focused on a benchmark problem to better illustrate the concepts of surrogate modeling for geomechanical applications. To demonstrate that the methodology is also applicable to a more realistic setting, we extend our considerations to a model based on the lithological variability of a potential siting region for radioactive waste Nördlich Lägern in northern Switzerland. The model is a simplified version of the model of Hergert et al. (2015) with the stratigraphy and their depth inspired by the well Stadel-3-1 (Crisci et al., 2022a). The model does not have any lateral variation of the topography or layer thickness, nor does the units have an inclination, as observed in that region. Therefore, we construct a model (Fig. 10) with an extend of $2500 \times 2500 \times 1400\,\mathrm{m}^3$ with five elements in each horizontal direction (x- and y-direction) and 140 elements in the vertical direction for each of the 15 layers to ensure a vertical resolution of minimum 3 m. We on purpose use the same number of elements for each layer to avoid a bias in the sensitivity analysis caused by the layer thickness. To elaborate, consider a model consisting of two layers, where the top layer has a thickness of 500 m and the base layer has a thickness of 200 m. If we

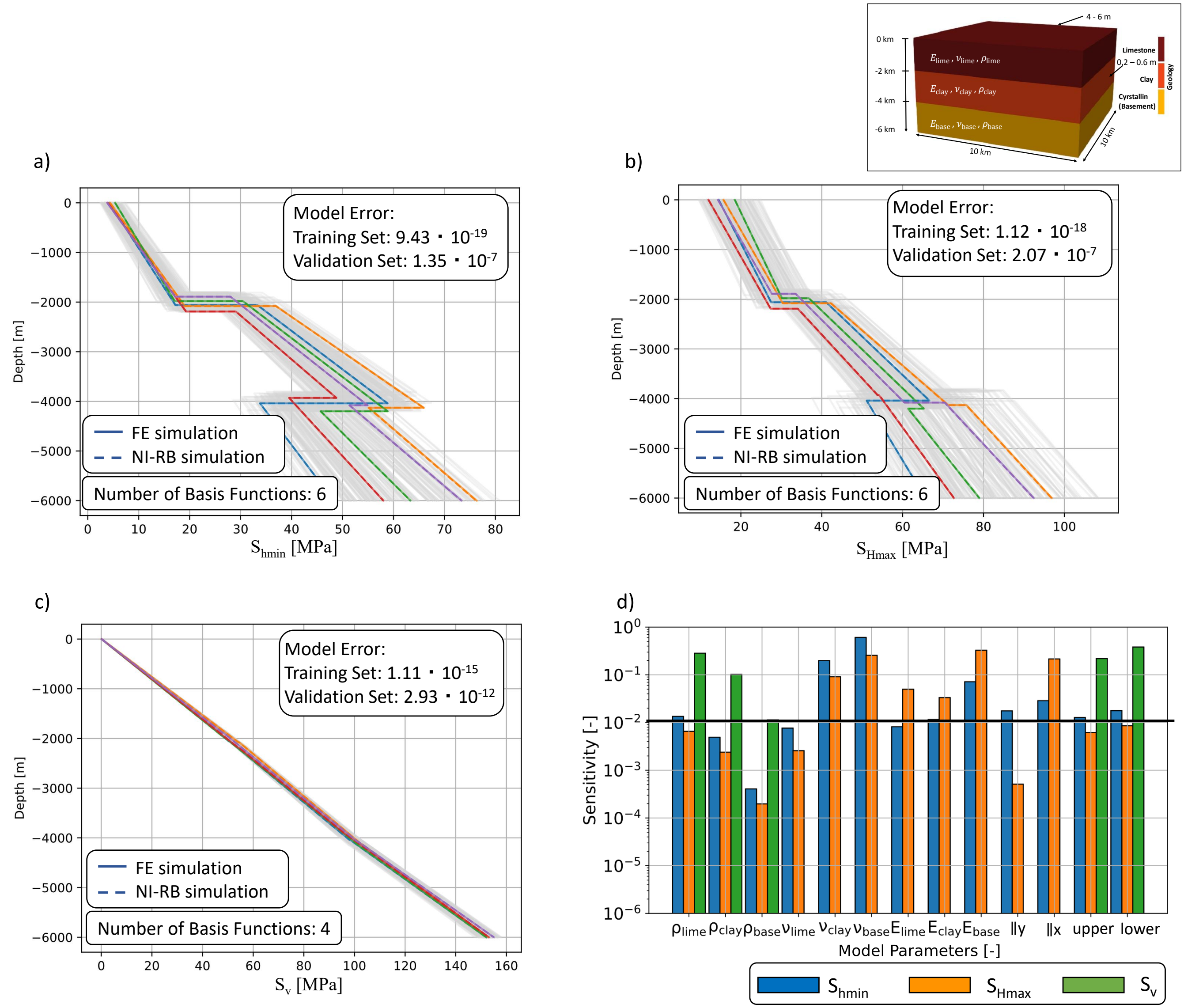

Figure 9. Investigation of the potential of surrogate models to determine the influence of varying boundary conditions, material properties and the interface depths. The figure shows comparison of the surrogate model accuracy for five randomly chosen realization of the validation data set for a) the S_{hmin}-component, b) the S_{Hmax}-component, and c) the S_v-component of the stress tensor. Note that "Sensitivity" denotes the value of the sensitivity index (section 2.4). The full-order solution are denoted by colored solid lines and the colored reduced solutions by dashed lines. The different colors represent the stress response for different boundary conditions, material properties, and interface depths. The training data set consisting of 200 samples for is indicated with light gray lines. Furthermore, we show in d) the results of the global SA.

discretize the model with a vertical resolution of 100 m, we require five elements for the first layer and two elements for the second layer. Calculating the sensitivity indices takes the variation of the stress distribution in each element into consideration. Consequently, the top layer will have more elements contributing and, because of that, likely a higher impact. By discretizing the model with an equal number of elements per layer (e.g., five elements for both the top and base layer), we can avoid this effect. For further information regarding potential biases, we refer to Degen and Wellmann (2024). For this model, we consider only variations of the Young's modulus. We do not consider the variability of the other elastic material properties since the density is well-known from density logs of boreholes, and both the density and the Poisson's ratio demonstrated little impact in previous studies (e.g. Gölke and Coblentz, 1996; Reiter, 2021; Ziegler, 2022). All linear elastic input parameters are listed in Tab. 2. Regarding the boundary conditions, we follow a similar approach to the previous models. The eastern boundary has a Dirichlet boundary condition of 0.14 m extension, and the southern boundary has a Dirichlet boundary condition of 1.8 m shortening. The top boundary is associated with a zero Neumann boundary condition, and all other boundaries are subjected to zero Dirichlet boundary conditions normal to the boundary.

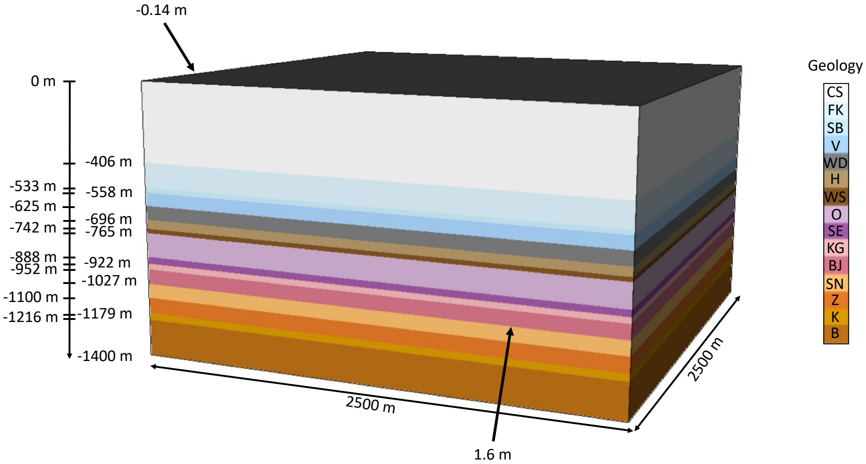

Figure 10. Schematic representation of simplified model based on the well Stadel-3-1, representing the case study of Nördlich Lägern in Switzerland. The geomechnaical units are the Cenozoic Sediments (CS), Felsenkalke (FK), Schwarzbach Fm. (SB), Villigen Fm (V), Wildegg Fm. (WD), Herrenwis Fm. (H), Wedelsandstein Fm. (WS), Opalinus clay (O), Staffelegg Fm. (SE), Klettgau Fm. (KG), Bänkerjoch Fm. (BJ), Schinznach Fm. (SN), Zeglingen Fm. (Z), Kaiseraugust Fm. (K), and the Basement incl. Buntsandstein and Rotliegend (B). The lateral boundary conditions of the model are a shortening of 1.6 m from the south and an extension of 0.14 m in eastern direction.

The stratigraphic column is inspired by the well Stadel-3-1 and the material properties are based on data from the four boreholes Bachs-1-1, Bülach-1-1, Stadel-2-1, and Stadel-3-1 that are presented in Dossiers VI and IX which can be downloaded from the NAGRA (download at https://nagra.ch/downloads). These are in particular the NAGRA Working Reports NAB 22-04 (Gonus et al., 2023, Bachs-1), NAB 22-26 (Spillmann et al., 2022, Bülach-1), NAB 22-02 (Crisci et al., 2022b; Gonus et al., 2022b, Stadel-2) and NAB 22-01 (Crisci et al., 2022a; Gonus et al., 2022a, Stadel-3). Consequently, also the geometrical rep-

Table 2. Variation ranges of the Young's modulus (E) and values for the other elastic material properties, as well as the layer thicknesses, for the case study of Nördlich Lägern. The parameter ranges have been obtained through personal communication and are publicly available in the technical report NAGRA (2024) by the 15th of May 2025. Note that Fm. denotes Formation. The data on the Young's moduls is based on information jointly inferred from four boreholes: Bachs-1-1, Bülach-1-1, Stadel-2-1, and Stadel-3-1 (Crisci et al., 2021, 2022a, b), the thickness data are derived from well Stadel-3-1 Crisci et al. (2022a).

Layer	**E [GPa]**	**E_{mean} [GPa]**	**Varability of E [GPa]**	**ν [-]**	**ρ [kg/m^3]**	**Layer Thickness [m]**
Cenozoic Sediments (CS)	10-20	15	10	0.30	2500	406
Felsenkalke (FK)	19-40	35	21	0.19	2686	127
Schwarzbach Fm. (SB)	15-36	24	21	0.25	2656	25
Villigen Fm. (V)	27-51	39	24	0.21	2685	67
Wildegg Fm. (WD)	14-31	20	17	0.25	2625	71
Herrenwis Fm. (H)	10-23	19	13	0.26	2676	46
Wedelsandstein Fm. (WS)	6-20	10	14	0.29	2306	23
Opalinus clay (O)	10-23	12	13	0.35	2544	123
Staffelegg Fm. (SE)	11-41	23	30	0.24	2573	34
Klettgau Fm. (KG)	11-39	17	28	0.23	2544	30
Bänkerjoch Fm. (BJ)	13-57	23	44	0.20	2625	75
Schinznach Fm. (SN)	18-38	27	20	0.23	2683	73
Zeglingen Fm. (Z)	15-46	26	31	0.20	2800	79
Kaiseraugst Fm. (K)	13-55	19	42	0.27	2664	37
Basement (B)	15-35	25	20	0.23	2479	184

resentation is subjected to low uncertainties, meaning that we do not vary the interface positions for this case study. Lastly, the boundary conditions are derived from a previous study by Hergert et al. (2015) and remain the same throughout all realizations.

As for the previous section, we first present the approximation quality and other characteristics of the surrogate models themselves. For both the surrogate model of the S_{hmin}- (Fig. 11a) and S_{Hmax}-component (Fig. 11b), we observe overall a similar approximation quality as for the benchmark study demonstrating the ability of the approach to be extended to more complex settings. The highest inaccuracies in the predictions are observed for realizations yielding maximum stress values.

Another interesting aspect is the number of basis functions, indicating the complexity of the parameter space. For both contributions of the stress tensor, we require 15 basis functions to achieve a POD tolerance of 10^{-4} %. Taking into account that we vary in total 15 Young's modulus values, this demonstrates generally a low complexity of the models and holds promises for the extension to higher dimensional parameter spaces. Note that we did not construct a surrogate model for the vertical stress component since we do not vary the density and hence we obtain no variations of the vertical stress for all realizations.

Shifting to the results of the global SA, presented in Fig. 11c, we note that the Young's modulus values of all 15 layers impact the responses of both horizontal stress components.

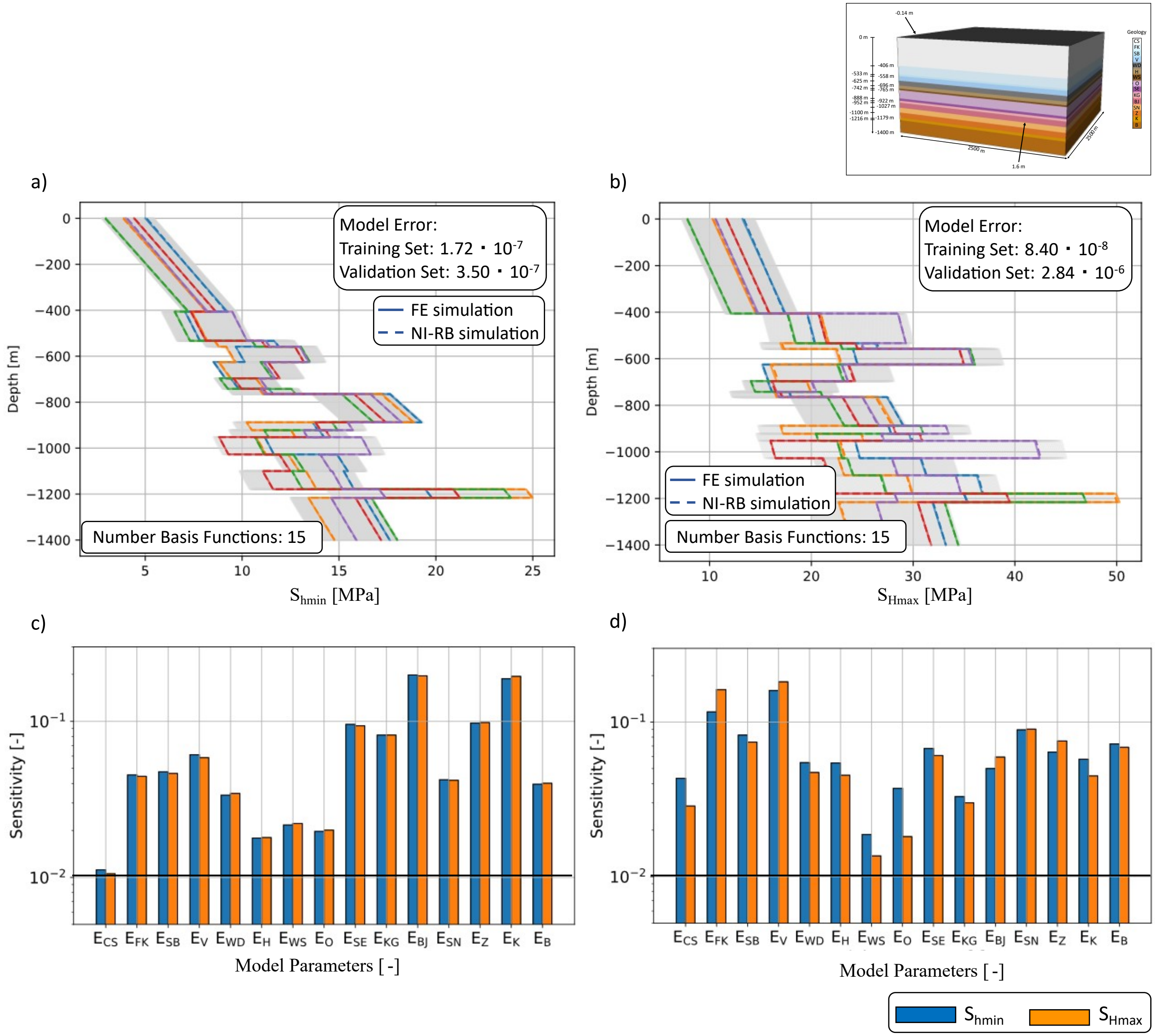

Figure 11. Investigation of the potential of surrogate models for the case study of Nördlich Lägern in Switzerland. Comparison of the surrogate model accuracy for five randomly chosen realization of the validation data set for a) the S_{hmin}- and b) the S_{Hmax}-component of the stress tensor. Note that "Sensitivity" denotes the value of the sensitivity index (section 2.4). The full-order solution are denoted by colored solid lines and the reduced solutions by colored dashed lines. The different colors represent the stress response for different Young's modulus values. The training data set consisting of 200 samples is indicated with light gray lines. Furthermore, we show the global SA results in c) non-equal, and d) for equal parameter ranges.

Consequently, for all following analyses, we would need to take all parameters into account. Although all parameters are deemed influential, we still observe differences in the amount of influence they have on the model response. For the S_{hmin}-component, the Young's modulus of the Bänkerjoch Fm. has the highest impact, followed by the Young's modulus of the Kaiseraugst Fm. Also, for the S_{Hmax}-component these two formations yield the highest impact on the stress distribution, although their influences are reversed with respect to the S_{hmin}-component. The main reason for the high influence of these two layers is the large variation range in the Young's module value. Another minor reason is the possible high contrast in the Young's module to the adjacent layers partly because of this variation range. To better illustrate the impact of the variations range, the variability is shown in Tab. 2.

The high influences of the Bänkerjoch and Kaiseraugst formations are followed by the high influences of the Staffelegg, Klettgau, and Zeglingen formations. All three formations are characterized by variation ranges of about 30 GPa. So, slightly lower variation ranges than the previously mentioned formations but higher variation ranges than for the units. This pattern is also observable for the remaining parameters within the global SA. So, within this global SA, the level of impact is mainly determined by the variation ranges of the Young's modulus values. Consequently, the SA results are highly impacted by our prior knowledge since the different variation ranges are a result of heterogeneities or uncertainties within the layers. However, naturally, a wrong estimate of these parameter ranges yields a bias in the global SA. To investigate the potential differences in the SA, assuming similar heterogeneities and levels of knowledge for the individual layers, we repeated the SA with equal variation ranges for the Young's moduli. Therefore, we used the mean values provided in Tab. 2 and applied variation of $\pm 10\%$. This yields a shift in the relevance of the different layers for the stress distributions. For both horizontal stress components, the Villigen Fm. has the highest impact, followed by the Felsenkalke and Schwarzbach formation. The lowest influences arise from the Wedelsandstein formation. Furthermore, low influences of the Opalinus clay are observed for the S_{hmin}-component. Overall the influence of the individual layers in this configuration is mainly determined by the contrast in the Young's modulus of the current layer and its adjacent layers.

5 Discussion

Within this study, we present how physics-based machine learning can be used to construct reliable surrogate models to predict the full stress tensor for a geomechanical model, e.g. for a nuclear waste disposal application. We start the discussion with general implications and challenges of surrogate modeling. Afterwards, we provide a comparison of physics-based vs. data-driven machine learning methods to emphasize the importance of appropriately choosing a surrogate model technique. Finally, we present application-specific aspects.

5.1 General Implications and Challenges for Surrogate Modeling

An important aspect is the explainability of the generated surrogate models. In the case of the NI-RB method, we perform first the POD which produces a number of basis functions. These basis functions are singular vectors characterizing the dominant physical behavior of the system. The machine learning part is only responsible for determining the weights of these basis functions. This has several implications. First, the non-provable part of the model and consequently the non-provable part of the error is solely restricted to the calculation of the reduced coefficients and is captured by a combination of scalar values. In addition, the errors that occur are explainable by considering both the physical relationship and the model geometry. As an example, in Figure 11b, we observe very small visible deviations between the reduced and full-order models (e.g., for the orange lines at a depth 1,200 m). Still, these deviations occur in the form of shifts and yield smooth error distributions and are not very pronounced. On the other hand, the errors of the NN models (as we present in the next section) do not appear in the entirety of the model and introduce a noisy behavior in the error distribution. This means, they do not follow any physical relationship, which has been discussed in further detail in Degen et al. (2023).

In literature (e.g., Degen et al., 2023; Faroughi et al., 2022; Raissi et al., 2019; Willard et al., 2022) physics-based machine learning is often used to reduce the amount of training data. The idea is that through introducing physical knowledge the amount of admissible solutions can be decreased and hence less data is required to train the surrogate models. Due to the simplicity of the presented example, this is a phenomenon that we do not observe. But note that this will become important for real-case studies.

In this manuscript, we investigate three sources of variability: i) the boundary conditions, ii) the material properties, and iii) the geometry. As demonstrated the NI-RB method performs very well in constructing physically consistent surrogate models for considering variations of boundary conditions and material properties. The variability in the layer depth proves to be more challenging. This is especially pronounced for the presented linear elastic models since the high contrast in the material properties yields shifts in the stress distribution. However, this phenomenon occurs also for other applications, for instance, for hydrological studies, where the permeability contrast is large. So, this demonstrates the importance to reformulate the problem as presented in Section 4.3.

Reduced basis methods using Proper Orthogonal Decomposition and Support Vector Machines (Zhao, 2021), Recurrent Neural Networks (Kumar et al., 2021), or the Empirical Cubature method (Guo et al., 2024) as the projection method have already been used for mechanical applications, even considering the stress distribution. However, in the example presented in this manuscript, the high contrast in the linear elastic material properties (mainly Young's modulus, but also Poission's ratio and density) yields non-smooth distributions which result in consequences not discussed before. Furthermore, we also consider the variation of the geometrical parameters, which with the exception of Guo et al. (2024) have not been taken into account before. Note that Guo et al. (2024) consider geometrical variations for the context of two-scale simulations yielding very different requirements in terms of the stress conditions.

Another aspect we want to highlight is the global error. As seen for the case of changing geometrical parameters with changing number of elements per layer (Figure 6a), the global errors are still low with 10^{-7} for the scaled training and 10^{-4} for the validation data set of the S_{hmin}-component despite the inaccuracies at the interface location. This demonstrated the general problem of global error estimators. If there is a small region that has large deviations and the rest of the model regions are well fitted, this does not become apparent in the global error since this error considers an average over all realizations and the entire model domain. The other question is why the problem is more pronounced at the upper than the lower interface. Looking at the stress distributions, we see larger gaps between the different stress realizations at the upper than the lower interface. Hence, in the case of the lower interface, better distributed information is available to determine the reduced solutions.

A major advantage of the RB method is the computational gain by using surrogate models instead of full-order model evaluations. Therefore, we briefly want to discuss the aspect of computational costs. The computational cost for all five scenarios is comparable with an execution time of about 3 min for a single finite element solution, in contrast, an evaluation of the surrogate required 0.1 ms yielding a speed-up of six orders of magnitude. The speed-up does not consider the computational time during the offline stage. So, we want to extend the discussion by providing the computational cost for the construction of the surrogate. In this work, we used between 100 to 200 realizations for the training data set, and up to 20 realizations for the validation data set. Hence, the most expensive surrogate model construction requires 220 realizations, with 3 min per realization this yields of cost of about 11 h. In addition to this cost comes the cost of performing the POD and the GPR, which are in the order of seconds. Since this is negligible with respect to the computational cost of constructing the data sets, we will discard this cost in the following. Consequently, the total construction cost of the surrogate model is equal to the cost of about 220 model evaluations. Comparing this to the over 50 million model evaluations required for the sensitivity analyses combined, the surrogate quickly pays off. It is also worth noting that the construction of the data sets is “embarrassingly” parallel, meaning that all model runs are completely independent. Therefore, parallel computing can be used to reduce the time. This is important because methods such as Markov chain Monte Carlo, used to quantify uncertainties, are not fully parallelizable. So, we not only save computation time but also shift the expensive computation to a stage where we can fully exploit parallel hardware infrastructures. This aspect, as well as how we can exploit graphics processing units is already discussed in detail (Degen et al., 2023). Another advantage is the non-intrusive nature of the algorithm, which allows for straightforward inclusion of varying forward solvers.

The geometry considered in this study is simple, to focus on the surrogate model construction itself and its implications for geomechanical modeling. Both the non-intrusive and intrusive versions of the reduced basis method have been applied to geothermal real-case studies with a complex geometrical setup, demonstrating a similar performance as presented in this study (Degen and Cacace, 2021; Degen et al., 2021a, b, 2022a, b, c). This implies, that especially for fixed geometries, no degradation in the surrogate model quality or performance is expected. Also, considering variations for complex geometries is possible. However, one should keep in mind that the base assumption for the method is the existence of a low-dimensional parameter space. Hence, if too many geometrical parameters are varied at once, this assumption breaks down, and the approach will become inefficient. As mentioned before, the non-intrusive reduced basis method has been developed in order to provide an efficient extension of the RB method for nonlinear applications (Hesthaven and Ubbiali, 2018). In contrast to this study, Degen et al. (2022a) consider fully coupled thermo-hydro-mechanical simulations, demonstrating that the current approach is also extendable to nonlinear and coupled relationships. The implications for nonlinear applications are further detailed in Degen et al. (2023) for different subsurface applications, including the highly nonlinear Richard’s equations. The results of these studies demonstrate great promise also for nonlinear geomechanical applications, which is especially important when considering potential extensions to elasto-plasticity.

5.2 Physics-based vs. Data-Driven Machine Learning Techniques

To illustrate the differences between the data-driven and the presented physics-based machine learning method, we constructed another surrogate model for the case of simultaneously changing boundary conditions and material properties (case 4) using a neural network (NN). The results are shown in Figure 12 and yield several observations. First, the global errors of both approaches differ by orders of magnitude for both the training and the validation data set. For the training data set, we achieve an error in the order of 10^{-18} for the non-intrusive RB method and 10^{-6} for the NN. We set for the POD a tolerance of 10^{-4} %. Since we use the root mean square error for the evaluation of the surrogate that means the NN approach is two orders of magnitude less accurate than originally desired. This behavior is – albeit not as extreme – also observable for the validation data set, where we have errors in the order of 10^{-7} for the NI-RB method and 10^{-6} for the NN.

The difference in the error behavior is especially prominent when focusing not on the global errors but on the local distributions. Therefore, we plot in Figure 12b the reduced solutions of the data-driven surrogate and the full solutions. At first glance, the approximation quality seems to be similar between the two surrogate models. However, by having a closer look major differences are observable. For the NI-RB surrogates the solutions are exactly on top of the full-order solutions making them non-distinguishable by eye. In contrast, for the NN surrogate differences are directly visible. This can be best observed in the base layer for the red and blue lines. For the blue line, we obtain an underestimation of the stress response in the base layer, whereas for the red case the stress response is overestimated. This becomes more prominent by observing the differences between the full-order realizations and the NN solutions in Fig. 12d. But more interesting are the different overall responses of both surrogate models. The NI-RB reduced solutions are straight lines corresponding to the physical response. For the NN surrogate, we obtain mostly straight lines but with superimposed oscillations. To better highlight the differences, we plot in addition the differences between the non-intrusive RB and the full-order solutions in Fig. 12c, and the differences between the NN and full-order solutions in Fig. 12d. The oscillations caused by the NN are not explainable from a physical perspective and are a result of treating the solutions in a purely data-driven manner within a statistical method. This demonstrates the fundamental difference between both approaches even without considering aspects such as the amount of available data. The NI-RB method produces solutions that follow the physical behavior whereas the NN model produces results that statistically vary around this general physical response. In a presented trivial case, this does not matter. However, in more complex cases, these differences can be significant and crucial.

A comparison between the different physics-based machine learning methods, for instance, between the non-intrusive reduced basis method and physics-informed neural networks (PINNs), is also interesting. In their original formulation, PINNs are designed for state estimation problems (Raissi et al., 2019). This makes a direct comparison of both approaches challenging since our study focuses on parameter estimation applications. Nonetheless, the error behavior of the PINN is expected to be similar to the one of the NN because of the architectural design of PINNs, as discussed in detail in Degen et al. (2023). A detailed comparison of the NI-RB method with other surrogate modeling approaches is detailed in previous studies (Degen et al., 2023; Santoso et al., 2022), and since no conceptual changes are expected not repeated in this study.

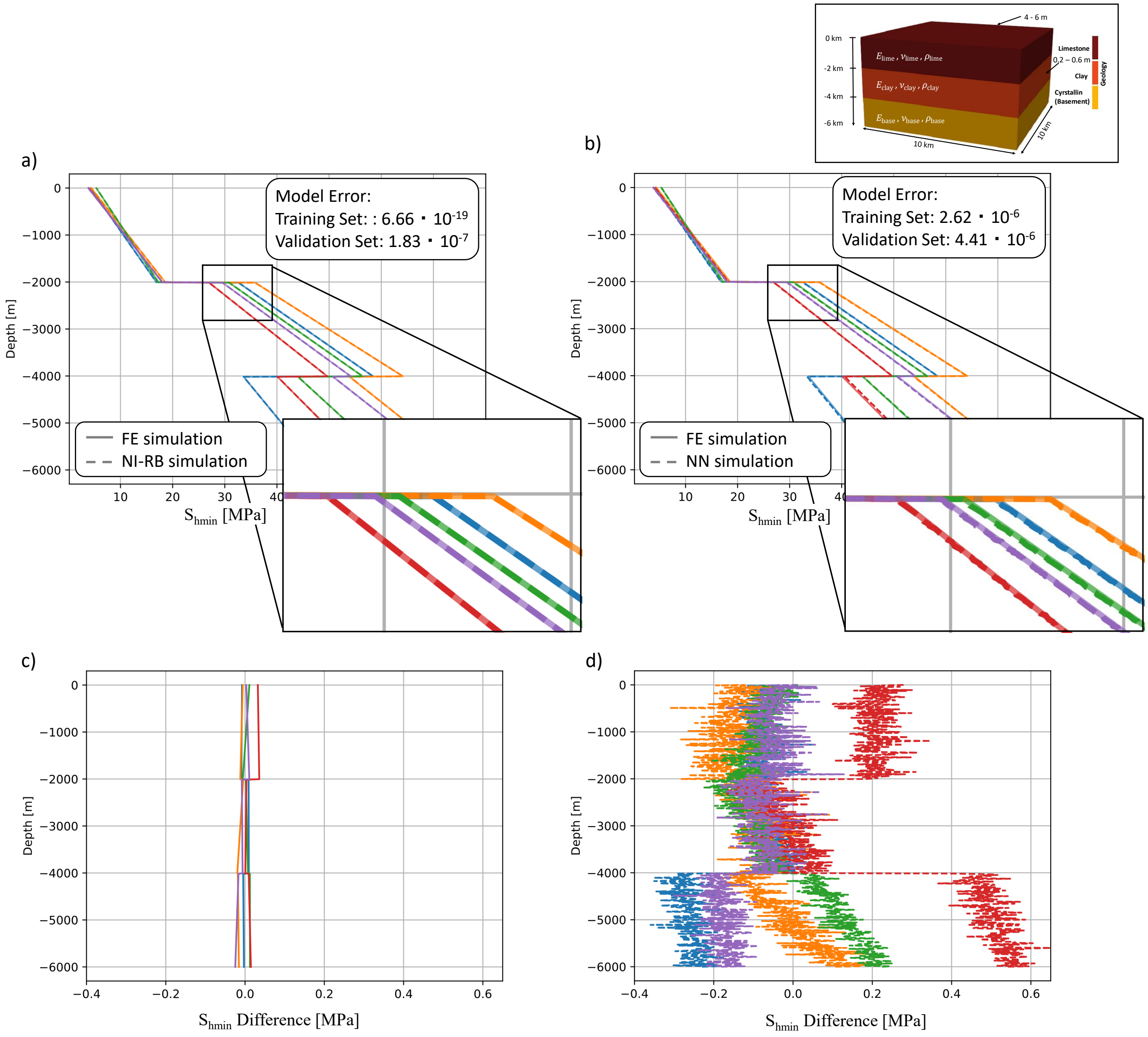

Figure 12. Comparison of the surrogate models constructed with a) the non-intrusive RB method, b) a neural network, c) the difference between the non-intrusive RB solutions and the full-order solutions, and d) the difference between the NN solutions and the full-order solutions for five randomly chosen realization of the validation data set for the case of changing simultaneously the boundary conditions and material properties. The full-order solution are denoted by colored solid lines and the reduced solutions by colored dashed lines. The different colors represent the stress response for different boundary conditions and material properties.

5.3 Implications for Nuclear Waste Disposal Applications

The subsurface stress distribution is of major importance in applications such as nuclear waste disposal, geothermal energy production, or subsurface (interim) storage of H_2 or CO_2. For these applications not only the cost of the numerical simulations but also their trustworthiness is essential since predictions have to be made over long periods of time and – in particular in the case of nuclear waste storage – false predictions can have major environmental impacts. The sparsity and uncertainty of available stress magnitude data, geological structures, and rock properties is a major challenge for the in-situ stress prediction (e.g., Hergert et al., 2015; Morawietz et al., 2020; Ziegler and Heidbach, 2020). The uncertainties in the in-situ stress state will then be transmitted to the results of forward models that predict the evolution of the stress state during e.g., the excavation stage of the repository. In particular, long-term subsurface usage such as nuclear waste storage has to account for a broad range of different future scenarios which in turn drive the uncertainties. Therefore, the uncertainties of the initial stress prediction should be as small as possible.

We demonstrated the possibility to extend the methodology to more complex settings through the case study of Nördlich Lägern. An interesting aspect of the global SAs for this case study is the relative low impact of the Opalinus clay for the S_{Hmax}-component, which is the target horizon for the future nuclear waste disposal site. This low impact means, that within the tested variations ranges, changes in the Young's modulus of the Opalinus clay do impact the maximum stress contribution less than most other layers. The largest impact arises from the stiff units. This is an important finding since it provides information about the impact of uncertainties on the planning phase of the nuclear waste disposal site. In the current study, we focus on the conceptual analyses of dominant physical processes and parameters in the form of sensitivity analyses. We do not consider real data on the stress magnitudes originating, for instance, from downhole measurements. This presents an interesting extension for future studies, where these measurements can be incorporated in the form of probabilistic uncertainty quantification methods, such as Markov chain Monte Carlo. This is interesting because many current studies consider best-case and worst-case scenarios only. A probabilistic uncertainty quantification approach allows to provide not only the most likely scenario but also the associated range of uncertainties and their probability to be encountered. This is important for the planning of nuclear waste disposal sites since best and worst-case analyses tend to estimate extreme values that have a low probability to be encountered. With a revised estimate of the uncertainties, the planning and construction might be improved, for instance, in terms of resource usage. Other potential future extensions include the incorporation of both global sensitivity analyses and uncertainty quantification into decision-making processes.

The last aspect, we want to discuss concerns the sensitivity analysis. We investigate two scenarios, sensitivity analysis using i) non-equal parameter ranges and ii) equal parameter ranges. To understand why these different analyses were conducted it is important to note that as with every model, sensitivity analyses are designed with a specific purpose in mind. This purpose is defined in the quantity of interest and means that an SA is not easily reusable if the scientific question changes. In a typical application, we are interested in which model parameters impact the responses the most. So, in our case, whether the Young's modulus, the Poisson's ratio, or the density has the highest impact on the stress distribution. In a realistic case study, the relative variations ranges of these various material properties may differ. That is the case because some properties can be determined with greater accuracy than others and properties such as the density have lower physically plausible variation ranges than, for instance, the Young's modulus. This is a naturally occurring phenomenon and we want to incorporate these into the SA. Therefore, typically the possible variation ranges are determined with respect to this prior knowledge yielding an SA with unequal parameter ranges.

Still, we are not always aiming to perform a SA to explain the behavior of a specific test site. In some cases, we are interested in what physical processes are driving the system. In this case, we would like to change the setup of the SA. To understand the driving forces independent of a specific site it might be beneficial to consider equal parameter ranges. This removes the bias that is introduced by the unequal parameter distributions. However, note that in most cases this bias is desired since it reflects the effects encountered in nature.

6 Conclusion

To conclude, we presented a physics-based machine learning model that can efficiently construct trustworthy surrogate models. The non-intrusive RB method enables evaluations that take less than a millisecond, whereas the full-order solves are in the range of minutes. This yields a speed-up of six orders of magnitude demonstrating the use of these surrogate models for probabilistic uncertainty quantification, global sensitivity analysis, and parameter estimation studies in general.

In contrast to data-driven surrogates, the model follows the physical relationship and is physically explainable. This is of great importance for predictions and decision-making processes, as demonstrated in the highly sensitive case of nuclear waste disposal.

The surrogate models allow a simultaneous evaluation of both the boundary conditions and the material properties without losing relevant precision with respect to the full-order solutions. In addition, the incorporation of geometrical variations is possible without degrading the quality of the surrogate model. Therefore, this work lays the foundation for a joint consideration of all three sources of uncertainty for robust stress predictions.

Code and data availability. The training and validation datasets, their associated model parameters, and the non-intrusive RB and neural network code for the construction of all surrogate models are published in the following Zenodo repository: https://doi.org/10.5281/zenodo.13767010. For the construction of the datasets the software package GOLEM (https://doi.org/10.5281/zenodo.999401; Jacquey and Cacace, 2017; Cacace and Jacquey, 2017) has been used. This software is based on the finite-element solver MOOSE (Permann et al., 2020), which is freely available on Zenodo (https://doi.org/10.5281/zenodo.4074777).

Author contributions. All authors discussed and interpreted the presented work. DD carried out the simulations and drafted the paper. The "Subsurface Stress State" section were jointly drafted by MZ and KR. Furthermore, MZ generated Figure 1 and 2.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. The authors greatly acknowledge the funding provided by the Bundesministerium für Umwelt, Naturschutz, nukleare Sicherheit und Verbraucherschutz through the project SQuaRe (project number: 02E12062C). Contributions by MZ were made under his German Research Foundation grant (DFG grant PHYSALIS 523456847).

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., and Zheng, X.: TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, https://www.tensorflow.org/, software available from tensorflow.org, 2015.

Ahlers, S., Henk, A., Hergert, T., Reiter, K., Müller, B., Röckel, L., Heidbach, O., Morawietz, S., Scheck-Wenderoth, M., and Anikiev, D.: 3D crustal stress state of Germany according to a data-calibrated geomechanical model, Solid Earth, 12, 1777–1799, 2021.

Amadei, B. and Stephansson, O.: Rock stress and its measurement, Springer Science & Business Media, 1997.

Bär, K., Reinsch, T., and Bott, J.: The PetroPhysical Property Database (P3) – a global compilation of lab-measured rock properties, Earth System Science Data, 12, 2485–2515, https://doi.org/10.5194/essd-12-2485-2020, 2020.

Benner, P., Gugercin, S., and Willcox, K.: A Survey of Projection-Based Model Reduction Methods for Parametric Dynamical Systems, SIAM Review, 57, 483–531, https://doi.org/10.1137/130932715, 2015.

Bergen, K. J., Johnson, P. A., de Hoop, M. V., and Beroza, G. C.: Machine learning for data-driven discovery in solid Earth geoscience, Science, 363, eaau0323, 2019.

Blöcher, G., Cacace, M., Jacquey, A. B., Zang, A., Heidbach, O., Hofmann, H., Kluge, C., and Zimmermann, G.: Evaluating micro-seismic events triggered by reservoir operations at the geothermal site of Groß Schönebeck (Germany), Rock Mechanics and Rock Engineering, 51, 3265–3279, 2018.

Bond, C. E., Johnson, G., and Ellis, J. F.: Structural model creation: the impact of data type and creative space on geological reasoning and interpretation, Geological Society, London, Special Publications, 421, 83–97, https://doi.org/10.1144/SP421.4, 2015.

Brady, B. H. and Brown, E. T.: Rock mechanics: for underground mining, Springer science & business media, 2006.

Cacace, M. and Jacquey, A. B.: Flexible parallel implicit modelling of coupled thermal–hydraulic–mechanical processes in fractured rocks, vol. 8, pp. 921–941, Copernicus GmbH, 2017.

Cosenza, A., Mannina, G., Vanrolleghem, P. A., and Neumann, M. B.: Global sensitivity analysis in wastewater applications: A comprehensive comparison of different methods, Environmental modelling & software, 49, 40–52, 2013.

Crisci, E., Laloui, L., and Giger, S.: Arbeitsbericht NAB 20-08 TBO Bülach Dossier IX Rock-mechanical and Geomechanical Laboratory Testing, Arbeitsbericht, National Cooperative for the Disposal of Radioactive Waste (NAGRA), Wettingen, 2021.

Crisci, E., Laloui, L., and Giger, S.: Nagra technical report 22-01, TBO Stadel-3-1: Data Report, Dossier IX: Rock-mechanical and Geomechanical Laboratory Testing, 2022a.

Crisci, E., Laloui, L., and Giger, S.: Nagra technical report 22-02, TBO Stadel-2-1: Data Report, Dossier IX: Rock-mechanical and Geomechanical Laboratory Testing, 2022b.

Degen, D. and Cacace, M.: Effects of transient processes for thermal simulations of the Central European Basin, Geoscientific Model Development, 14, 1699–1719, 2021.

Degen, D. and Wellmann, F.: Guidelines for Sensitivity Analyses in Process Simulations for Solid Earth Geosciences, 2024.

Degen, D., Veroy, K., and Wellmann, F.: Certified reduced basis method in geosciences, Computational Geosciences, 24, 241–259, 2020.

Degen, D., Spooner, C., Scheck-Wenderoth, M., and Cacace, M.: How biased are our models? – a case study of the alpine region, Geoscientific Model Development, 14, 7133–7153, https://doi.org/10.5194/gmd-14-7133-2021, 2021a.

Degen, D., Veroy, K., Freymark, J., Scheck-Wenderoth, M., Poulet, T., and Wellmann, F.: Global sensitivity analysis to optimize basin-scale conductive model calibration–A case study from the Upper Rhine Graben, Geothermics, 95, 102 143, 2021b.

Degen, D., Cacace, M., and Wellmann, F.: 3D multi-physics uncertainty quantification using physics-based machine learning, Scientific Reports, 12, 17 491, 2022a.

Degen, D., Veroy, K., Scheck-Wenderoth, M., and Wellmann, F.: Crustal-scale thermal models: Revisiting the influence of deep boundary conditions, Environmental Earth Sciences, 81, 1–16, 2022b.

Degen, D., Veroy, K., and Wellmann, F.: Uncertainty quantification for basin-scale geothermal conduction models, Scientific reports, 12, 1–10, 2022c.

Degen, D., Caviedes Voullième, D., Buiter, S., Hendricks Franssen, H.-J., Vereecken, H., González-Nicolás, A., and Wellmann, F.: Perspectives of physics-based machine learning strategies for geoscientific applications governed by partial differential equations, Geoscientific Model Development, 16, 7375–7409, https://doi.org/10.5194/gmd-16-7375-2023, 2023.

Desroches, J., Peyret, E., Gisolf, A., Wilcox, A., Di Giovanni, M., de Jong, A. S., Sepehri, S., Garrard, R., and Giger, S.: Stress measurement campaign in scientific deep boreholes: focus on tool and methods, in: SPWLA Annual Logging Symposium, p. D021S018R003, SPWLA, 2021a.

Desroches, J., Peyret, E., Gisolf, A., Wilcox, A., Giovanni, M., Jong, A., Milos, B., Gonus, J., Bailey, E., Sepehri, S., Garitte, B., Garrard, R., and Giger, S.: Stress-Measurement Campaign in Scientific Deep Boreholes: Focus on Tools and Methods, in: SPWLA 62nd Annual Logging Symposium, p. 21, 2021b.

Ellsworth, W. L.: Injection-induced earthquakes, science, 341, 1225 942, 2013.

Faroughi, S. A., Pawar, N., Fernandes, C., Das, S., Kalantari, N. K., and Mahjour, S. K.: Physics-guided, physics-informed, and physics-encoded neural networks in scientific computing, arXiv preprint arXiv:2211.07377, 2022.

Fischer, K. and Henk, A.: A workflow for building and calibrating 3-D geomechanical models &ndash a case study for a gas reservoir in the North German Basin, Solid Earth, 4, 347–355, 2013.

Gaucher, E., Schoenball, M., Heidbach, O., Zang, A., Fokker, P. A., van Wees, J.-D., and Kohl, T.: Induced seismicity in geothermal reservoirs: A review of forecasting approaches, Renewable and Sustainable Energy Reviews, 52, 1473–1490, 2015.

Geuzaine, C. and Remacle, J.-F.: Gmsh: A 3-D finite element mesh generator with built-in pre-and post-processing facilities, International journal for numerical methods in engineering, 79, 1309–1331, 2009.

Gölke, M. and Coblentz, D. D.: Origins of the European regional stress field, Tectonophysics, 266, 11–24, https://doi.org/10.1016/S0040-1951(96)00180-1, 1996.

Gonus, J., Bailey, E., Desroches, J., and Garrard, R.: Nagra technical report 22-02, TBO Stadel-3-1: Data Report, Dossier VI: Wireline Logging, Micro-hydraulic Fracturing and Pressure-meter Testing, 2022a.

Gonus, J., Bailey, E., Desroches, J., and Garrard, R.: Nagra technical report 22-02, TBO Stadel-2-1: Data Report, Dossier VI: Wireline Logging, Micro-hydraulic Fracturing and Pressure-meter Testing, 2022b.

Gonus, J., Bailey, E., Desroches, J., and Garrard, R.: Nagra technical report 22-04, TBO Bachs-1-1: Data Report, Dossier VI: Wireline Logging, Micro-hydraulic Fracturing and Pressure-meter Testing, 2023.

Guo, M. and Hesthaven, J. S.: Data-driven reduced order modeling for time-dependent problems, Computer methods in applied mechanics and engineering, 345, 75–99, 2019.

Guo, T., Rokoš, O., and Veroy, K.: A reduced order model for geometrically parameterized two-scale simulations of elasto-plastic microstructures under large deformations, Computer Methods in Applied Mechanics and Engineering, 418, 116 467, 2024.

Heidbach, O., Rajabi, M., Cui, X., Fuchs, K., Müller, B., Reinecker, J., Reiter, K., Tingay, M., Wenzel, F., Xie, F., et al.: The World Stress Map database release 2016: Crustal stress pattern across scales, Tectonophysics, 744, 484–498, 2018.

Henk, A.: Perspectives of Geomechanical Reservoir Models–Why Stress is Important, European Magazine, 4, 1–5, 2008.

Hergert, T., Heidbach, O., Reiter, K., Giger, S., and Marschall, P.: Stress field sensitivity analysis in a sedimentary sequence of the Alpine foreland, northern Switzerland, Solid Earth, 6, 533–552, 2015.

Hesthaven, J. and Ubbiali, S.: Non-intrusive reduced order modeling of nonlinear problems using neural networks, Journal of Computational Physics, 363, 55–78, https://doi.org/10.1016/j.jcp.2018.02.037, 2018.

Hesthaven, J. S., Rozza, G., Stamm, B., et al.: Certified reduced basis methods for parametrized partial differential equations, SpringerBriefs in Mathematics, Springer, 2016.

Jacquey, A. B. and Cacace, M.: GOLEM, a MOOSE-based application, https://doi.org/10.5281/zenodo.999401, 2017.

Jordan, M. I. and Mitchell, T. M.: Machine learning: Trends, perspectives, and prospects, Science, 349, 255–260, 2015.

Kotsiantis, S. B., Zaharakis, I., Pintelas, P., et al.: Supervised machine learning: A review of classification techniques, Emerging artificial intelligence applications in computer engineering, 160, 3–24, 2007.

Kumar, A., Hu, R., and Walsh, S. D.: Development of Reduced Order Hydro-mechanical Models of Fractured Media, Rock Mechanics and Rock Engineering, pp. 1–14, 2021.

Lecampion, B. and Lei, T.: Reconstructing the 3d initial stress state over reservoir geo-mechanics model from local measure-ments and geological priors: a bayesian approach, Schlumberger J Model Des Simul, 1, 100–104, 2010.

Li, Y. E., O'Malley, D., Beroza, G., Curtis, A., and Johnson, P.: Machine Learning Developments and Applications in Solid-Earth Geosciences: Fad or Future?, 2023.

Lindsay, A. D., Gaston, D. R., Permann, C. J., Miller, J. M., Andrš, D., Slaughter, A. E., Kong, F., Hansel, J., Carlsen, R. W., Icenhour, C., Harbour, L., Giudicelli, G. L., Stogner, R. H., German, P., Badger, J., Biswas, S., Chapuis, L., Green, C., Hales, J., Hu, T., Jiang, W., Jung, Y. S., Matthews, C., Miao, Y., Novak, A., Peterson, J. W., Prince, Z. M., Rovinelli, A., Schunert, S., Schwen, D., Spencer, B. W., Veeraraghavan, S., Recuero, A., Yushu, D., Wang, Y., Wilkins, A., and Wong, C.: 2.0 - MOOSE: Enabling massively parallel multiphysics simulation, SoftwareX, 20, 101 202, https://doi.org/https://doi.org/10.1016/j.softx.2022.101202, 2022.

Loh, W.-L.: On Latin hypercube sampling, The annals of statistics, 24, 2058–2080, 1996.

Mahesh, B.: Machine learning algorithms-a review, International Journal of Science and Research (IJSR).[Internet], 9, 381–386, 2020.

Martin, D.: Quantifying in situ stress magnitudes and orientations for Forsmark – Formsark stage 2.2, Tech. Rep. SKB R-07-26, Swedish Nuclear Fuel and Waste Management Co, Stockholm Sweden, 2007.

Morawietz, S., Heidbach, O., Reiter, K., Ziegler, M., Rajabi, M., Zimmermann, G., Müller, B., and Tingay, M.: An open-access stress magnitude database for Germany and adjacent regions, Geothermal Energy, 8, 1–39, 2020.

NAGRA: Arbeitsbericht NAB 24-19 In-situ stress field in the siting regions Jura Ost, Nördlich Lägern and Zürich Nordost, Arbeitsbericht, National Cooperative for the Disposal of Radioactive Waste (NAGRA), Wettingen, 2024.

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., and Chintala, S.: PyTorch: An Imperative Style, High-Performance Deep Learning Library, in: Advances in Neural Information Processing Systems 32, edited by Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., and Garnett, R., pp. 8024–8035, Curran Associates, Inc., http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf, 2019.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E.: Scikit-learn: Machine Learning in Python, Journal of Machine Learning Research, 12, 2825–2830, 2011.

Permann, C. J., Gaston, D. R., AndrÅ¡, D., Carlsen, R. W., Kong, F., Lindsay, A. D., Miller, J. M., Peterson, J. W., Slaughter, A. E., Stogner, R. H., and Martineau, R. C.: MOOSE: Enabling massively parallel multiphysics simulation, SoftwareX, 11, 100 430, https://doi.org/https://doi.org/10.1016/j.softx.2020.100430, 2020.

Quarteroni, A., Manzoni, A., and Negri, F.: Reduced Basis Methods for Partial Differential Equations: An Introduction, UNITEXT, Springer International Publishing, 2015.

Raissi, M., Perdikaris, P., and Karniadakis, G. E.: Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, Journal of Computational physics, 378, 686–707, 2019.

Razavi, S. and Gupta, H. V.: What do we mean by sensitivity analysis? The need for comprehensive characterization of “global” sensitivity in E arth and E nvironmental systems models, Water Resources Research, 51, 3070–3092, 2015.

Reiter, K.: Stress rotation – impact and interaction of rock stiffness and faults, Solid Earth, 12, 1287–1307, https://doi.org/10.5194/se-12-1287-2021, 2021.

Reiter, K. and Heidbach, O.: 3-D geomechanical–numerical model of the contemporary crustal stress state in the Alberta Basin (Canada), Solid Earth, 5, 1123–1149, https://doi.org/10.5194/se-5-1123-2014, 2014.

Röckel, L., Ahlers, S., Müller, B., Reiter, K., Heidbach, O., Henk, A., Hergert, T., and Schilling, F.: The analysis of slip tendency of major tectonic faults in Germany, Solid Earth, 13, 1087–1105, https://doi.org/10.5194/se-13-1087-2022, 2022.

Rozza, G., Huynh, D. B. P., and Patera, A. T.: Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations: application to transport and continuum mechanics, Archives of Computational Methods in Engineering, 15, 229–275, 2008.

Saltelli, A.: Making best use of model evaluations to compute sensitivity indices, Computer physics communications, 145, 280–297, 2002.

Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S.: Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index, Computer Physics Communications, 181, 259–270, 2010.

Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., and Wu, Q.: Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices, Environmental modelling & software, 114, 29–39, 2019.

Santoso, R., Degen, D., Cacace, M., and Wellmann, F.: State-of-the-art physics-based machine learning for hydro-mechanical simulation in geothermal applications, in: European Geothermal Congress, pp. 1–10, 2022.

Sarrazin, F., Pianosi, F., and Wagener, T.: Global Sensitivity Analysis of environmental models: Convergence and validation, Environmental Modelling & Software, 79, 135–152, 2016.

Schmitt, D. R., Currie, C. A., and Zhang, L.: Crustal stress determination from boreholes and rock cores: Fundamental principles, Tectonophysics, 580, 1–26, 2012.

Schulz, E., Speekenbrink, M., and Krause, A.: A tutorial on Gaussian process regression: Modelling, exploring, and exploiting functions, Journal of Mathematical Psychology, 85, 1–16, 2018.

Segall, P. and Fitzgerald, S. D.: A note on induced stress changes in hydrocarbon and geothermal reservoirs, Tectonophysics, 289, 117–128, 1998.

Sin, G., Gernaey, K. V., Neumann, M. B., van Loosdrecht, M. C., and Gujer, W.: Global sensitivity analysis in wastewater treatment plant model applications: prioritizing sources of uncertainty, Water research, 45, 639–651, 2011.

Singha, D. K. and Chatterjee, R.: Geomechanical modeling using finite element method for prediction of in-situ stress in Krishna–Godavari basin, India, International Journal of Rock Mechanics and Mining Sciences, 73, 15–27, 2015.

Smart, K. J., Ofoegbu, G. I., Morris, A. P., McGinnis, R. N., and Ferrill, D. A.: Geomechanical modeling of hydraulic fracturing: Why mechanical stratigraphy, stress state, and pre-existing structure matter, AAPG bulletin, 98, 2237–2261, 2014.

Sobol, I. M.: Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Mathematics and computers in simulation, 55, 271–280, 2001.

Song, X., Zhang, J., Zhan, C., Xuan, Y., Ye, M., and Xu, C.: Global sensitivity analysis in hydrological modeling: Review of concepts, methods, theoretical framework, and applications, Journal of hydrology, 523, 739–757, 2015.

Spillmann, T., Schoenball, M., Tanner, R., and Diehl, T.: Nagra technical report 22-26, Ergänzung Schwachbebennetz in der Nordschweiz: Installation Bohrlochseismometer Bülach-1-1, 2022.

Swischuk, R., Mainini, L., Peherstorfer, B., and Willcox, K.: Projection-based model reduction: Formulations for physics-based machine learning, Computers & Fluids, 179, 704–717, https://doi.org/10.1016/j.compfluid.2018.07.021, 2019.

Tang, T., Reed, P., Wagener, T., and Van Werkhoven, K.: Comparing sensitivity analysis methods to advance lumped watershed model identification and evaluation, Hydrology and Earth System Sciences Discussions, 3, 3333–3395, 2006.

Tingay, M., Heidbach, O., Davies, R., and Swarbrick, R.: Triggering of the Lusi mud eruption: Earthquake versus drilling initiation, Geology, 36, 639–642, 2008.

van Wees, J.-D., Osinga, S., Van Thienen-Visser, K., and Fokker, P. A.: Reservoir creep and induced seismicity: inferences from geomechanical modeling of gas depletion in the Groningen field, Geophysical Journal International, 212, 1487–1497, 2018.

Vanrolleghem, P. A., Mannina, G., Cosenza, A., and Neumann, M. B.: Global sensitivity analysis for urban water quality modelling: Terminology, convergence and comparison of different methods, Journal of Hydrology, 522, 339–352, 2015.

Wainwright, H. M., Finsterle, S., Jung, Y., Zhou, Q., and Birkholzer, J. T.: Making sense of global sensitivity analyses, Computers & Geosciences, 65, 84–94, 2014.

Wang, Q., Hesthaven, J. S., and Ray, D.: Non-intrusive reduced order modeling of unsteady flows using artificial neural networks with application to a combustion problem, Journal of Computational Physics, 384, 289–307, https://doi.org/10.1016/j.jcp.2019.01.031, 2019.

Willard, J., Jia, X., Xu, S., Steinbach, M., and Kumar, V.: Integrating scientific knowledge with machine learning for engineering and environmental systems, ACM Computing Surveys, 55, 1–37, 2022.

Willcox, K. E., Ghattas, O., and Heimbach, P.: The imperative of physics-based modeling and inverse theory in computational science, Nature Computational Science, 1, 166–168, 2021.

Zhao, H.: A reduced order model based on machine learning for numerical analysis: An application to geomechanics, Engineering Applications of Artificial Intelligence, 100, 104 194, 2021.

Ziegler, M., Reiter, K., Heidbach, O., Zang, A., Kwiatek, G., Stromeyer, D., Dahm, T., Dresen, G., and Hofmann, G.: Mining-Induced Stress Transfer and Its Relation to a M _w M w 1.9 Seismic Event in an Ultra-deep South African Gold Mine, Pure and Applied Geophysics, 172, 2557–2570, 2015.

Ziegler, M., Heidbach, O., Morawietz, S., and Wang, Y.: Manual of the Matlab Script FAST Calibration v2.4, WSM Technical Report ; 23-02, https://doi.org/10.48440/WSM.2023.002, 2023.

Ziegler, M. O.: Rock properties and modelled stress state uncertainties: a study of variability and dependence, Rock Mechanics and Rock Engineering, 55, 4549–4564, 2022.

Ziegler, M. O. and Heidbach, O.: The 3D stress state from geomechanical–numerical modelling and its uncertainties: a case study in the Bavarian Molasse Basin, Geothermal Energy, 8, https://doi.org/10.1186/s40517-020-00162-z, 2020.

Ziegler, M. O. and Heidbach, O.: Increasing accuracy of 3-D geomechanical-numerical models, Geophysical Journal International, 237, 1093–1108, https://doi.org/10.1093/gji/ggae096, 2024.

Ziegler, M. O., Heidbach, O., Reinecker, J., Przybycin, A. M., and Scheck-Wenderoth, M.: A multi-stage 3-D stress field modelling approach exemplified in the Bavarian Molasse Basin, Solid Earth, 7, 1365–1382, 2016.

Zoback, M. D.: Reservoir Geomechanics, Cambridge University Press, https://doi.org/10.1017/cbo9780511586477, 2007.