Surrogate models in convection-dominated fault systems: Considerations for efficient and reliable realizations

Thomas Simader[1,2,3*], Denise Degen[4,5], Tobias Meier[1], Horst Rüter[3], Florian Wellmann[2,6]
geomecon GmbH, 10553 Berlin, Germany[1*], RWTH Aachen University, 52062 Aachen, Germany[2], HarbourDom GmbH, 44287 Dortmund, Germany[3],
TU Darmstadt, 64287 Darmstadt, Germany[4] ,GFZ Helmholtz Center for Geosciences, 14473 Potsdam, Germany[5], Fraunhofer IEG, 44801 Bochum, Germany[6]
thomas.simader@geomecon.de*

INTRODUCTION

Accurate temperature predictions in geothermal systems are essential but **computationally demanding**, especially when fluid flow and heat transport are strongly coupled. Exploring uncertainties or multiple scenarios is therefore often **infeasible. Surrogate models** offer a promising way to **drastically reduce computation time**.

Fault zones can be particularly attractive geothermal targets when they exhibit **strong thermal anomalies** through **natural convection**. However, convection introduces non-linearity, multistability, and potentially chaotic behavior, which **challenges surrogate modeling** approaches.

In this study, we investigate how the **non-intrusive reduced basis method (NI-RB)** can be used to efficiently construct surrogate models of the **thermal state at fault zones** and analyze the associated physical and numerical challenges.

METHODS

Full-order model

- Synthetic thermo-hydraulic model generated with COMSOL Multiphysics® (FEM)
- Temperature- and pressure-dependent fluid density and viscosity (Zyvoloski et al., 1997)

Surrogate model (NI-RB)

- The NI-RB combines physics-based and data-driven approaches to provide reliable predictions of complex partial differential equations (Hesthaven et al., 2018)
- Consists of a computationally expensive offline stage which involves the a priori generation of 100-200 full-order model snapshots
- Next, r basis functions ψ are selected via Proper Orthogonal Decomposition. A feed-forward neural network is then trained to map the parameters $\vec{\mu}$ onto the reduced coefficients θ_{rb}. The resulting reduced solution u_{rb} is:

$$u_{rb}(\vec{\mu}) = \sum_{i=1}^{r} \theta_{rb}^{(i)}(\vec{\mu})\psi_i$$

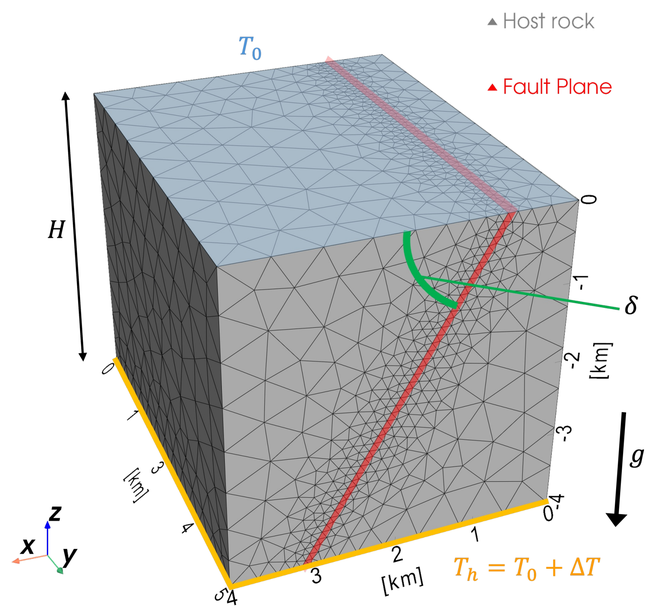


Figure 1. Full-order model consisting of a high-permeable fault (1.5x10^{-13} m^2) embedded within a low-permeable host rock (10^{-17} m^2).

RESULTS

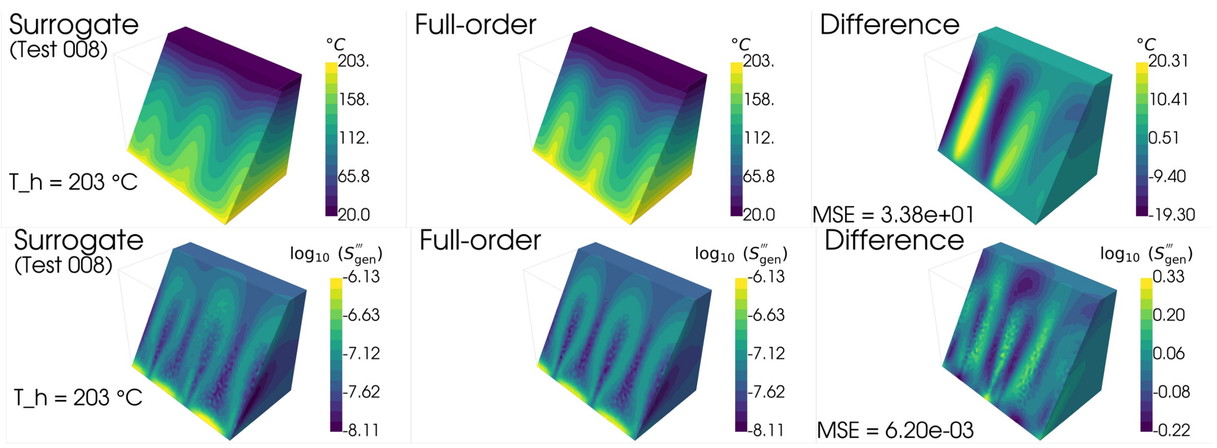


Figure 2. Comparison of surrogate and full-order model results for a varying bottom temperature T_h evaluated at 203 °C for a temperature- and entropy-based surrogate.

- Entropy Generation Evaluation: The local entropy generation rate is defined as:

$$S'''_{gen} = \frac{\lambda}{T_0^2}(\nabla T)^2 \ \text{[W K}^{-1}\text{ m}^{-3}\text{ s}^{-1}\text{]}$$

- To analyze the thermal state of the system, we use the dimensionless entropy generation number:

$$N_s = \frac{\int_V S'''_{gen} dV}{\int_V \left[\frac{\lambda(\Delta T)^2}{H^2 T_0^2}\right] dV} \ [-]$$

- Constructed entropy- and temperature-based surrogate models for single-parameter variations: varying bottom temperature T_h, dip angles δ, and fault permeabilities k_F
- Expanded models to multi-parameter spaces varying T_h alongside thermal conductivity λ, porosity φ, and permeability k for fault and host rock.
- Evaluating N_s against the free parameter can reveal a bifurcation diagram
- Overlapping stable solutions for similar parameter values can lead to unreliable surrogate models
- Physical preconditioning through a suitable temperature initial condition leads to fewer bifurcations:

$$T(x,y,z) = A\sin\left(\frac{\pi z}{H}\right)\cos\left(\frac{m\pi x}{Hh_x}\right)\cos\left(\frac{n\pi y}{Hh_y}\right) + 1 - \frac{z}{H}$$

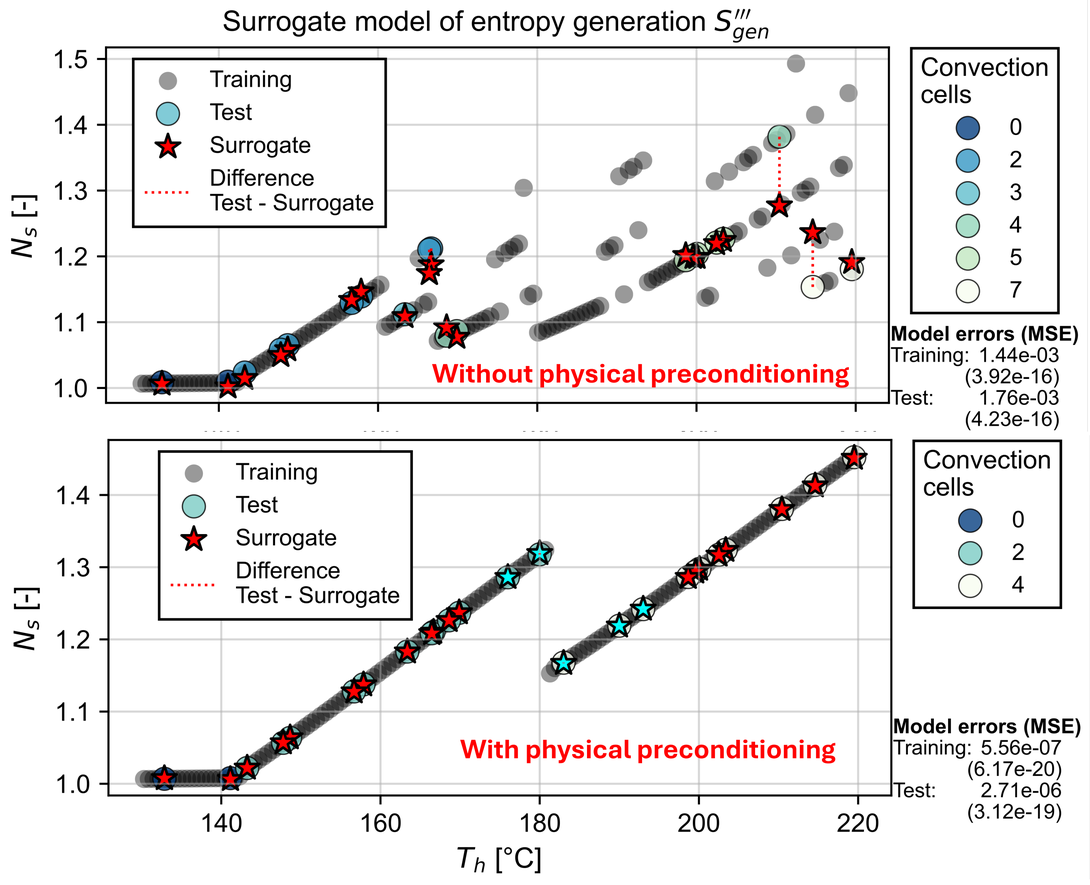


Figure 3. Surrogate model results for the entropy generation field for varying bottom boundary temperatures T_h. The derived N_s values are shown for the full-order training and test data sets as well as the surrogate. Mean squared errors are provided for scaled entropy generation; values in brackets denote the unscaled MSE. The upper and lower graph correspond to models without and with physical preconditioning, respectively.

DISCUSSION

- Surrogate modeling of **convection-dominated fault systems** is challenged by **strong nonlinearity, multistability, and chaotic behavior.**
- **Entropy generation–based surrogates** were **more robust** than temperature-based surrogates in this study.
- **Physical preconditioning** significantly **improves surrogate accuracy** by reducing the number of competing solution branches in the full-order model.
- Surrogate performance consistently **deteriorates near bifurcation** points and numerical instability is unavoidable.
- Surrogate **models provide speed-ups of up to seven orders of magnitude**, enabling uncertainty quantification and sensitivity analyses that are otherwise computationally infeasible.

CONCLUSION

- Surrogate models can accelerate simulations of convective fault zone–host rock systems by **several orders of magnitude**, but their reliability is challenged by **coexisting solutions** and **regime transitions**.
- **Entropy generation analysis** facilitates the **identification of bifurcation points**, which coincide with regions of **degraded surrogate performance**.
- N_s **serves as an effective additional performance metric** for surrogate model validation beyond conventional error measures.
- Due to the inherently **non-unique and oscillatory** nature of free convection, **stochastic surrogate evaluations** may be required. When system behavior is poorly constrained a priori, extending the surrogate parameter space with **wavenumbers** m, n enables exploration of **multiple physically possible solution branches**.

Acknowledgements

The authors greatly acknowledge the funding provided by the Ministry of Economic Affairs, Industry, Climate Action and Energy of the State of North Rhine-Westphalia and the European Union (grant KarboEx2 EFRE-20800290). Contributions by Denise Degen were made under the grant from the German Federal Ministry of Research, Technology and Space (grant PBML-HM 16|S24062). Furthermore, the COMSOL Multiphysics® simulations could be carried out via the Python API developed by Henning et. al (2024).

KarboEx^2

Kofinanziert von der Europäischen Union

Gefördert durch:
Ministerium für Wirtschaft, Industrie, Klimaschutz und Energie des Landes Nordrhein-Westfalen

REFERENCES

Bejan, A.: Convection Heat Transfer, **4th ed**. (2013), Wiley, New Jersey.
Börsing, N., Wellmann, J. F., Niederau, J., and Regenauer-Lieb, K.: Entropy production in a box: Analysis of instabilities in confined hydrothermal systems, Water Resources Research, **53** (2017), 7716–7739.
Degen, D. et al.: Perspectives of physics-based machine learning strategies for geoscientific applications governed by partial differential equations, Geosci. Model Dev., **16** (2023), 7375–7409.
Guillou-Frottier, L. et al.: Heat flow, thermal anomalies, tectonic regimes and high-temperature geothermal systems in fault zones, Comptes Rendus. Géoscience, **356** (2024), 389–421.
Hesthaven, J. S. and Ubbiali, S.: Non-intrusive reduced order modeling of nonlinear problems using neural networks, Journal of Computational Physics, **363** (2018), 55–78.