Proceedings World Geothermal Congress 2026
Calgary, Canada, June 8 – 11, 2026

Bayesian Data Assimilation for High-Dimensional Subsurface Parameter Inference in Modeling of Borehole Heat Exchanger Fields

Hesam Soltan Mohammadi[1,2], Lisa Maria Ringel[3], Christoph Bott[1], Selçuk Erol[4] and Peter Bayer[1*]

Department of Applied Geology, Institute of Geosciences and Geography, Martin Luther University Halle-Wittenberg, Von-Seckendorff-Platz 3, 06120 Halle (Saale), Germany[1], Climate Protection Office, City Administration of Viersen, Bahnhofstraße 23-29, 41747 Viersen, Germany[2], Univ. Rennes, CNRS, Géosciences Rennes, UMR 6118, 35000 Rennes, France[3], Department of Energy Systems Engineering, Engineering Faculty, Petroleum and Natural Gas Engineering Department, Middle East Technical University (METU), ODTÜ Üniversiteler Mah.Dumlupınar Blv.No:1 06800 Çankaya Ankara, Türkiye[4]

[*]E-mail: peter.bayer@geo.uni-halle.de

Keywords: Borehole heat exchanger, Uncertainty quantification, Bayesian inversion, Data assimilation

ABSTRACT

Accurate prediction of subsurface temperature evolution in borehole heat exchanger (BHE) systems relies on robust characterization of spatiotemporal model parameters. This study presents a Bayesian data assimilation framework to infer the statistical distribution of highly correlated parameters using a low-fidelity but computationally efficient modeling approach. The framework is tailored for complex, multi-layered subsurface environments with groundwater flow. By sequentially assimilating monthly temperature measurements, the method systematically explores and updates the model input parameter space - such as thermal conductivity, Darcy velocity, and heat extraction rates - synthetic case study simulating a five-year BHE field operation in COMSOL Multiphysics® is used to validate the framework. The setup includes four active and one inactive BHE, with temperature measurements collected monthly from the inactive BHE. These observations are used to update the model and identify parameter sets that best reproduce the thermal response of the system. Updated parameters are then propagated forward in time for improved predictions. After 32 months of sequential inference, the framework reduces uncertainty in predicted subsurface temperatures to 8%, while maintaining flexibility to accommodate evolving thermal conditions. This work underscores the potential of Bayesian data assimilation and data-driven approaches to enhance long-term thermal modeling of BHE systems by adaptively tuning model parameters in response to observed data, thereby avoiding the need for computationally intensive simulation tools.

1 INTRODUCTION

Borehole heat exchangers (BHEs) are widely used for sustainable heating and cooling, and their long-term performance depends on accurately modeling the coupled conductive–advective heat transport processes in the subsurface (Benz et al. 2022, Chen et al. 2021, Gil et al. 2022, Stauffer et al. 2013). Although a variety of analytical and numerical models exist for system design, many of these approaches rely on simplified or static subsurface assumptions and depend heavily on thermal response tests that characterize the system only at the start of operation (Guo et al. 2023, Heim et al. 2024, Li and Lai 2015, Molina-Giraldo et al. 2011, Rivera et al. 2015). As the system continues to operate, factors such as seasonal load variations, groundwater flow, and geological heterogeneity gradually alter the subsurface thermal regime. These evolving conditions can introduce significant prediction errors and, in some cases, lead to thermal imbalance or regulatory non-compliance (Bayer et al. 2019, Blum et al. 202), Chen et al. 2021, Haehnlein et al. 2010, Rivera et al. 2017, Soltan Mohammadi et al. 2024a, Tsagarakis et al. 2020).

As the thermal state of the subsurface changes dynamically during long-term operation, calibration of model parameters becomes essential for maintaining reliable predictions (Puttige et al. 2020, Shin et al. 2024). Classical calibration approaches and optimization methods can improve parameter estimates but often struggle with high-dimensional, ill-posed inverse problems and may fail to capture nonlinear or time-varying uncertainties (Heim et al. 2022, Soltan Mohammadi et al. 2025). Alternatively, probabilistic techniques such as Kalman filters provide real-time updates but remain constrained by Gaussian assumptions and tend to lose robustness in nonlinear settings (Shoji et al. 2023).

In this context, Bayesian inference, implemented through Markov Chain Monte Carlo (MCMC) methods, offers a more flexible framework by characterizing full posterior parameter distributions without imposing restrictive

distributional assumptions. Previous studies have successfully used Bayesian techniques to estimate subsurface thermal conductivity, borehole thermal resistance, or model structural bias (Choi et al. 2018, Menberg et al. 2019, Pasquier and Marcotte 2020, Spitler and Gehlin 2015). However, these studies have been limited to short-term testing periods and have not addressed long-term operational data assimilation, which is critical for tracking the evolving thermal behavior of BHE systems (Soltan Mohammadi et al. 2024b).

To address this gap, this study presents a Bayesian inversion framework that continuously updates subsurface thermal parameters during operation of a BHE field (Soltan Mohammadi et al. 2025). The approach incorporates temperature measurements over the operational lifespan and infers statistical distributions of key parameters, including layer-specific thermal conductivities, Darcy velocity, and heat extraction rates. A computationally efficient moving finite line source model with anisotropy (MFLSA) is used as the forward model, enabling repeated evaluation of parameter proposals in a realistic multi-layer, advective subsurface (Erol and François 2018). By updating parameter estimates as new data become available, the framework improves predictive accuracy, reflects real thermal system evolution, and provides a quantified assessment of uncertainty in long-term performance predictions.

The goal of this work is to demonstrate that Bayesian data assimilation can effectively manage high-dimensional parameter spaces in BHE modeling and that it offers a pathway toward dynamic, data-driven thermal state prediction in complex geothermal environments.

2 METHODOLOGY

2.1 Forward modeling

To predict temperature changes around the borehole heat exchanger (BHE) field over long operational periods, a computationally efficient forward model is required that captures conductive and advective heat transport in heterogeneous subsurface conditions. In this study, the moving finite line source model with anisotropy (MFLSA) is employed as the forward solver, following the analytical developments in Erol and François (2018). The MFLSA accounts for layered geological media, anisotropic thermal conditions, and groundwater flow, and is therefore well suited for repeated evaluation within a Bayesian inversion framework.

The subsurface is represented as a multilayered porous medium, in which each layer is assigned independent hydraulic and thermal properties. Temperature changes at a given observation location are computed using a composite computational approach, where the domain is divided into a real segment (the layer containing the observation point) and imaginary segments in adjacent layers. Boundary conditions at layer interfaces are handled analytically via the method of images, allowing the layer-specific thermal responses to be superposed (Figure 1).

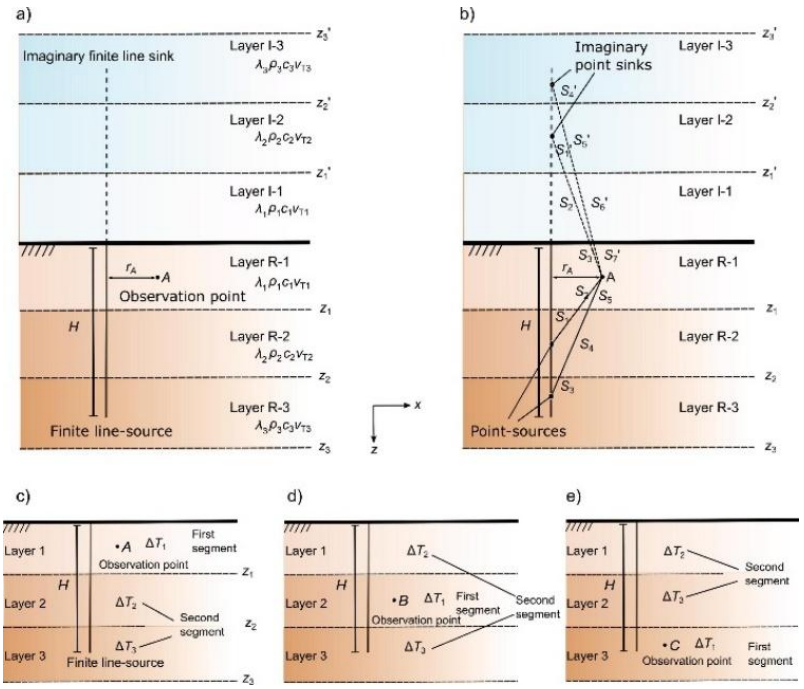


Figure 1. Illustration of the composite model approach in MFLSA for a BHE with length H (after Soltan Mohammadi et al. 2025).

For an observation point located in layer 1, the temperature change is

$$\Delta T_1(x,y,z,t) = \frac{q_L}{2\pi\lambda_{y1}} \exp\left[\frac{x v_{T1}}{2a_{x1}}\right]\left[\int_0^{z_1} f(x,y,z,t)dz' - \int_{-z_1}^{0} f(x,y,z,t)dz'\right] \quad (1)$$

where q_L is the heat extraction rate, λ_{y1} the effective transverse thermal conductivity, and a_{x1} the longitudinal thermal diffusivity. The kernel $f(\cdot)$ represents the combined effect of conduction, advection, and complementary error functions, and depends on the radial distance to the line source and on time.

The thermal transport velocity is

$$v_{T1} = \frac{u_x \rho_w c_w}{\rho_m c_m} \quad (2)$$

where u_x is Darcy velocity, $\rho_w c_w$ is the volumetric heat capacity of water, and $\rho_m c_m$ is the bulk volumetric heat capacity of the layer.

To incorporate dispersion, the effective thermal conductivities are

$$\lambda_x = \lambda_m + \alpha_l \rho_w c_w u_x \quad (3)$$

$$\lambda_y = \lambda_z = \lambda_m + \alpha_t \rho_w c_w u_x \quad (4)$$

with α_l and α_t denoting longitudinal and transverse dispersivities.

Using analogous expressions for layers 2 and 3, the total temperature change at an observation point A is obtained by summing contributions from all layers (including real and imaginary parts required by the method of images)

$$\Delta T_A(x,y,z,t) = \Delta T_1 + \Delta T_2 + \Delta T_3 \quad (5)$$

The comprehensive derivation and further details on the MFLSA can be found in Erol and François (2018).

2.2 Inverse Modeling

The objective of the inversion is to estimate the layer-specific heat extraction rates, thermal conductivities, and Darcy velocities, comprising a nine-dimensional parameter vector $\theta = \{q_1, q_2, q_3, \lambda_1, \lambda_2, \lambda_3, u_1, u_2, u_3\}$.

Following Bayes' rule, the posterior distribution of parameters given the monthly temperature measurements d along the inactive BHE is

$$p(\theta|d) \propto p(d|\theta)p(\theta) \quad (6)$$

where $p(\theta)$ denotes the prior distribution and $p(d|\theta)$ the likelihood. A Gaussian likelihood is chosen

$$\log p(d|\theta) \propto -\frac{1}{2\sigma^2 N}\sum_{i=1}^{N}(d_i - F(\theta)_i)^2 \quad (7)$$

with $F(\theta)$ denoting the MFLSA-predicted temperature changes and σ^2 representing the data error variance, updated sequentially based on differences between predicted and measured values in the previous timestep.

Priors for all parameters are uniform within physically reasonable bounds, consistent with limited prior information

$$p(\theta) = \frac{1}{\beta - \alpha}, \quad \alpha \le \theta \le \beta \quad (8)$$

As the parameter space is high-dimensional and strongly correlated, an affine-invariant ensemble sampler (AIES) is used, which efficiently explores stretched or rotated posterior distributions (Goodman and Weare 2010).

The sampler proposes parameter updates through the stretch move

$$z = \left((a-1)U(0,1)+1\right)^2/a \tag{9}$$

with step-size parameter $a = 2$. The acceptance of a proposed state is determined by the Metropolis-Hastings criterion,

$$\log\left(\frac{p(d|\tilde{\theta})}{p(d|\theta)}\right) > \log(s), \quad s \sim U(0,1) \tag{10}$$

A schematic overview of the sampler is shown in Figure 2.

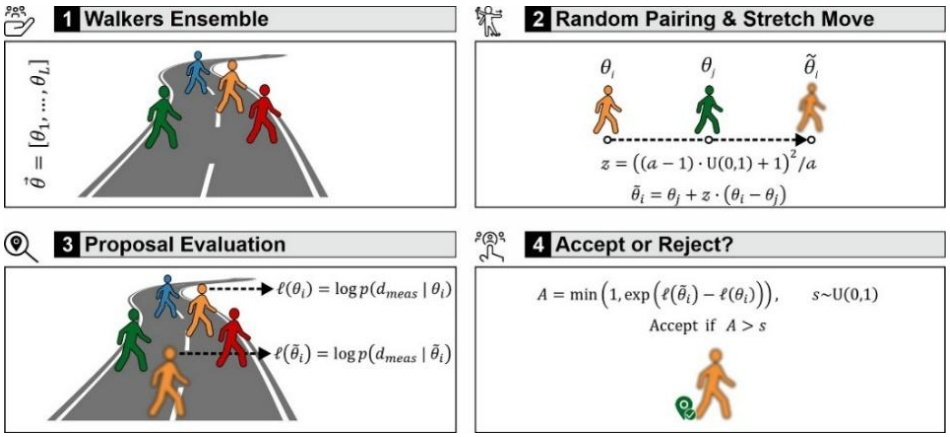


Figure 2. Schematic illustration of the affine-invariant ensemble MCMC sampler (after Soltan Mohammadi et al. 2025).

At each monthly timestep, 180 walkers and 50,000 iterations generate 5,040 posterior samples per parameter. The posterior median is used as the updated parameter estimate for the next timestep, and newly measured temperatures replace model-predicted temperatures from the same timestep, ensuring consistent sequential data assimilation.

2.3 Synthetic Model Setup

To evaluate the framework under controlled conditions, a synthetic BHE field is constructed using a high-resolution COMSOL Multiphysics® model (Figure 3). The field comprises four active BHEs operating in heating mode and one inactive BHE used exclusively for temperature monitoring. Sensors record monthly temperature changes along the inactive BHE with 1-m vertical resolution over a five-year period. The subsurface consists of three 20-m-thick layers, each with distinct hydraulic and thermal properties.

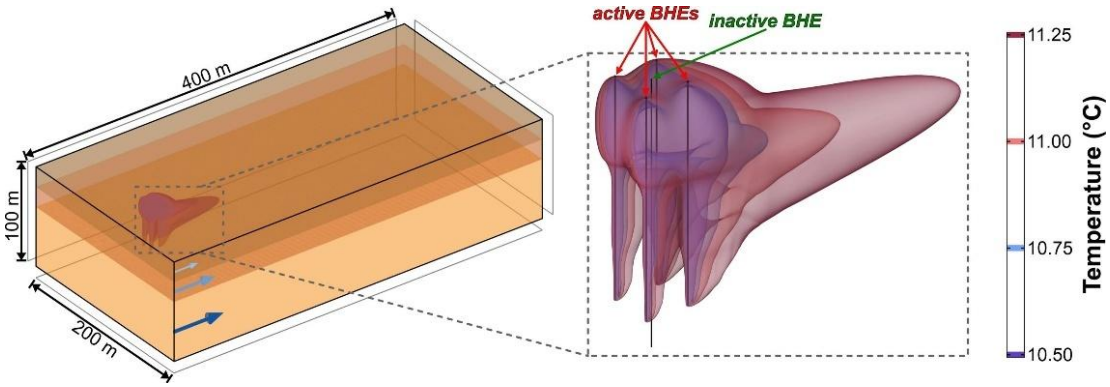


Figure 3. Numerical model with three layers (left) and the thermal plume after five years of operation (right), based on Soltan Mohammadi et al. 2025).

The numerical model incorporates advective–conductive heat transfer, realistic pipe-flow physics via the Pipe Flow Module, and appropriate boundary conditions. The resulting monthly temperature-change profiles serve as synthetic measurements for the Bayesian inversion.

2.4 Inversion Implementation

The inversion jointly estimates nine unknown parameters across 60 monthly timesteps. Parameter ranges follow physical constraints (e.g., heat extraction rate 5–40 $\mathrm{W\,m^{-1}}$; thermal conductivity 0.5–4 $\mathrm{W\,m^{-1}K^{-1}}$; Darcy velocity $6 \times 10^{-8} - 6 \times 10^{-6}\,\mathrm{m\,s^{-1}}$). At each timestep, the AIES sampler generates posterior distributions, whose medians are propagated forward. The computed temperatures for each new month are replaced with the measured temperatures from the numerical model, ensuring that the inversion consistently assimilates updated field information. This sequential approach enables the framework to capture the temporal evolution of parameter uncertainties and their effect on long-term temperature predictions.

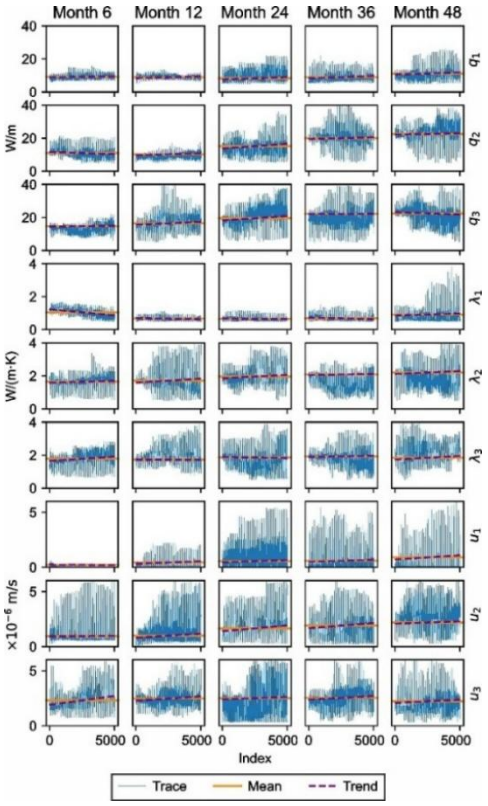


Figure 4. Trace plots for the nine model parameters (based on Soltan Mohammadi et al. 2025).

3 RESULTS AND DISCUSSION

3.1 MCMC Diagnostics

Sampling behavior assessed via trace plots shows well-mixed, stationary chains without drift for most parameters, indicating convergence in the sampling procedure (Figure 4).

3.2 Posterior Structure and Correlations

Posterior distributions reveal clear differences in identifiability across layers. Layer 1 parameters (q_1, λ_1, u_1) become tightly constrained, showing unimodal distributions due to pronounced temperature changes and limited groundwater advection (Figure 5). Layers 2 and 3 exhibit broader distributions because temperatures reach steady-state more quickly, reducing information for an efficient parameter identification.

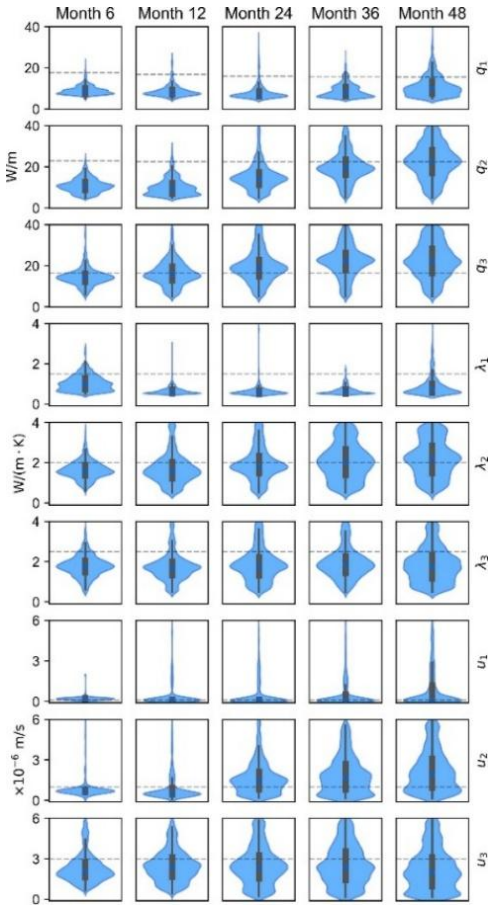


Figure 5. Violin plots of the nine model parameters (based on Soltan Mohammadi et al. 2025).

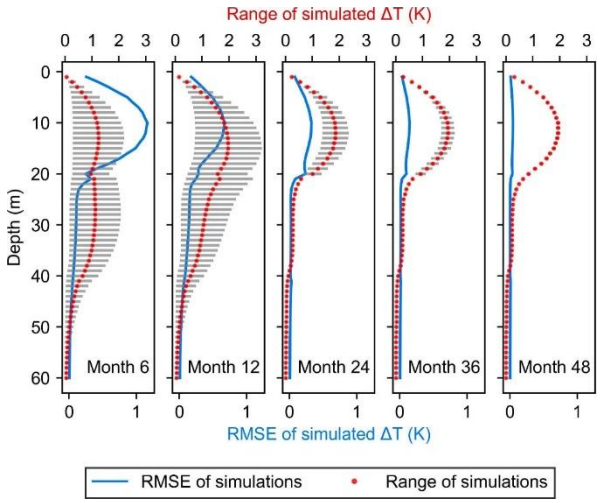


Figure 6. Comparison of the root mean square error and the range of temperature change predictions at month 60 (based on Soltan Mohammadi et al. 2025).

3.3 Prediction of Temperature Evolution

Posterior samples obtained from selected months are propagated to predict temperature changes at month 60. Early estimates (e.g., month 6) lead to underprediction of temperatures and large uncertainty, especially in the first layer. As more temperature data are assimilated, prediction uncertainty decreases substantially. The RMSE at month 60 decreases from 1.17 K (month-6 inference) to 0.05 K (month-48 inference). The spread in predicted temperatures decreases from 1.91 K to 0.16 K, representing >90 % reduction in uncertainty (Figure 6).

Temporal predictions at 10 m depth show that early posteriors incorrectly predict fast stabilization, whereas later posteriors capture the nonlinear transient behavior more accurately. Layers with higher groundwater flow (layers 2 and 3) reach steady-state faster, resulting in smaller long-term uncertainty but lower sensitivity early in the operation (Figure).

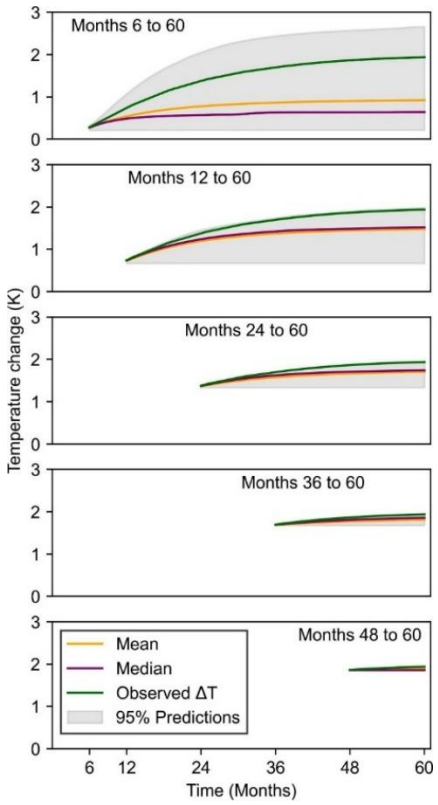


Figure 7. Comparison of the 95 % confidence interval, mean, and median of predicted temperature changes with observed temperature changes at a depth of 10 m, using inferred parameters from months 6, 12, 24, 36, and 48 until the end of the operation (after Soltan Mohammadi et al. 2025).

Overall, sequential Bayesian updating effectively refines parameter estimates and improves long-term thermal predictions.

4 CONCLUSIONS

This study is based on Soltan Mohammadi et al. (2025) and presents a Bayesian uncertainty quantification framework for long-term probabilistic prediction of temperature changes around BHE fields in stratified geological settings with groundwater flow. The combination of an efficient anisotropic line-source model with the affine invariant ensemble sampler enables dynamic inference of nine correlated thermal and hydraulic parameters using monthly temperature measurements.

Application to a synthetic five-year case demonstrates substantial reduction in predictive uncertainty as more data are assimilated. After approximately 2-3 years, uncertainty in temperature predictions decreases by more than 90 %, and posterior distributions stabilize for most parameters. The framework also extends the applicability of analytical line-source models by allowing layer-specific, time-varying heat extraction behavior.

The approach provides a practical and computationally efficient tool for improving long-term geothermal system assessments. Future work will apply the method to real field data, incorporate transient operational conditions, and investigate machine-learning to further improve sampling efficiency to further enhance robustness and applicability.

Acknowledgements

This work was supported by the Federal Ministry for Economic Affairs and Climate Action of Germany (BMWK) based on grant number 03EE4039A, within the GEOTHERMICA project RECOIN (ID: 455).

REFERENCES

Bayer, P., Attard, G., Blum, P., and Menberg, K.: The geothermal potential of cities, Renew Sustain Energy Rev, 106, (2019), 17–30.

Benz, S.A., Menberg, K., Bayer, P., and Kurylyk, B.L.: Shallow subsurface heat recycling is a sustainable global space heating alternative, Nat Commun, 13, (2022), 3962.

Blum, P., Menberg, K., Koch, F., Benz, S.A., Tissen, C., Hemmerle, H., et al.: Is thermal use of groundwater a pollution?, J Contam Hydrol, 239, (2021), 103791.

Chen, S., Cai, W., Witte, F., Wang, X., Wang, F., Kolditz, O., et al.: Long-term thermal imbalance in large borehole heat exchangers array – A numerical study based on the Leicester project, Energy Build, 231, (2021), 110518.

Choi, W., Menberg, K., Kikumoto, H., Heo, Y., Choudhary, R., Ooka, R.: Bayesian inference of structural error in inverse models of thermal response tests, Appl Energy, 228, (2018), 1473–1485.

Erol, S., and François, B.: Multilayer analytical model for vertical ground heat exchanger with groundwater flow, Geothermics, 71, (2018), 294–305.

Gil, A.G., Schneider, E.A.G., Moreno, M.M., and Cerezal, J.C.S.: Shallow Geothermal Energy: Theory and Application, Springer Nature, (2022).

Goodman, J., and Weare, J.: Ensemble samplers with affine invariance, Commun Appl Math Comput Sci, 5, (2010), 65–80.

Guo, Y., Huang, G., and Liu, W.V.: A new semi-analytical solution addressing varying heat transfer rates for U-shaped vertical borehole heat exchangers in multilayered ground, Energy, 274, (2023), 127373.

Haehnlein, S., Bayer, P., and Blum, P.: International legal status of the use of shallow geothermal energy, Renew Sustain Energy Rev, 14, (2010), 2611–2625.

Heim, E., Stoffel, P., Düber, S., Knapp, D., Kümpel, A., Müller, D., et al.: Comparison of simulation tools for optimizing borehole heat exchanger field operation, Geotherm Energy, 12, (2024), 24.

Heim, E., Laska, M., Becker, R., and Klitzsch, N.: Estimating the subsurface thermal conductivity and Its uncertainty for shallow geothermal energy use—A workflow and geoportal based on publicly available data, Energies, 15, (2022), 3687.

Li, M., and Lai, A.C.K.: Review of analytical models for heat transfer by vertical ground heat exchangers (GHEs): A perspective of time and space scales, Appl Energy, 151, (2015), 178–191.

Menberg, K., Heo, Y., and Choudhary, R.: Influence of error terms in Bayesian calibration of energy system models, J Build Perform Simul, 12, (2019), 82–96.

Molina-Giraldo, N., Blum, P., Zhu, K., Bayer, P., and Fang, Z.: A moving finite line source model to simulate borehole heat exchangers with groundwater advection, Int J Therm Sci, 50, (2011), 2506–2513.

Pasquier, P., and Marcotte, D.: Robust identification of volumetric heat capacity and analysis of thermal response tests by Bayesian inference with correlated residuals, Appl Energy, 261, (2020), 114394.

Puttige, A.R., Andersson, S., Östin, R., and Olofsson, T.: Improvement of borehole heat exchanger model performance by calibration using measured data, J Build Perform Simul, 13, (2020), 430–442.

Rivera, J.A., Blum, P., and Bayer, P.: Increased ground temperatures in urban areas: Estimation of the technical geothermal potential, Renew Energy, 103, (2017), 388–400.

Rivera, J.A., Blum, P., and Bayer, P.: Analytical simulation of groundwater flow and land surface effects on thermal plumes of borehole heat exchangers, Appl Energy, 146, (2015), 421–433.

Shoji, Y., Katsura, T., and Nagano, K.: Improvement of accuracy with uncertainty quantification in the simulation of a ground heat exchanger by combining model prediction and observation, Geothermics, 107, (2023), 102611.

Shin, E., Kim, Y., Kim, Y.-S., Lee, S., and Choi, W.: Dynamic management of ground thermal response uncertainty through temporal analysis of parameter sensitivity, Appl Energy, 376, (2024), 124267.

Soltan Mohammadi, H., Ringel, L.M., Bott, C., Erol, S., and Bayer, P.: Bayesian uncertainty quantification in temperature simulation of borehole heat exchanger fields for geothermal energy supply, Appl Therm Eng, 265, (2025), 125210.

Soltan Mohammadi, H., Ringel, L.M., de Paly, M., and Bayer, P.: Sequential long-term optimization of shallow geothermal systems under descriptive uncertainty and dynamic variation of heating demand, Geothermics, 121, (2024), 103021.

Soltan Mohammadi, H., Ringel, L.M., Bott, C., Bayer, P.: Adaptive management of borehole heat exchanger fields under transient groundwater flow conditions, Renew Energy, (2024), 121060.

Spitler, J.D., and Gehlin, S.E.A.: Thermal response testing for ground source heat pump systems - An historical review, Renew Sustain Energy Rev, 50, (2015), 1125–1137.

Stauffer, F., Bayer, P., Blum, P., Giraldo, N.M., and Kinzelbach, W.: Thermal use of shallow groundwater, CRC Press, (2013).

Tsagarakis, K.P., Efthymiou, L., Michopoulos, A., Mavragani, A., Anđelković, A.S., Antolini, F., et al.: A review of the legal framework in shallow geothermal energy in selected European countries: Need for guidelines, Renew Energy, 147, (2020), 2556–2571.