Data Worth Analysis for a Fluvial Geothermal System using an Ensemble Smoother with Multiple Data Assimilation

G. Song[1], S. Geiger[1], D. Voskov[1,2], H. Abels[1], P. Vardon[1]

[1] Delft University of Technology; [2] Stanford University

Summary

Long-term geothermal production is subject to considerable uncertainty due to limited data availability and inherent geological heterogeneity. While observation and data acquisition improve our understanding of the reservoir, they also contribute significantly to project costs. It is essential to identify the most informative observation strategy. In this study, we apply a previously developed scenario-based data assimilation framework that integrates rapid geological modelling, efficient numerical simulation, and Ensemble Smoother with Multiple Data Assimilation (ESMDA) to constrain uncertainties in reservoir properties and production forecasts to a synthetic but geologically realistic fluvial geothermal system and conduct a data worth analysis to evaluate the impact of different observations (production temperature and injection pressure, well temperature and pressure profiles, etc.) on uncertainty reduction. Results show that production temperature and injection pressure alone, though cost-effective, are insufficient to significantly reduce uncertainties in reservoir performance forecasts. In contrast, well temperature and pressure profiles exhibit substantially higher data worth, leading to much better-constrained predictions. Moreover, incorporating a monitoring borehole further constrains uncertainty by capturing subsurface dynamics between the injector and producer. These findings underscore the importance of monitoring pressure and temperature profiles in the wells of a geothermal doublet.

Data Worth Analysis for a Fluvial Geothermal System using an Ensemble Smoother with Multiple Data Assimilation

Introduction

Long-term production in Geothermal reservoirs is faced with considerable uncertainty due to limited data availability and intrinsic geological reservoir heterogeneity, which constrains our understanding of reservoir structures and systems. Inspired by the Delft campus geothermal project, a low-enthalpy project for district heating at TU Delft (Voskov et al., 2024), we carry out a systematic investigation of the value of different data collection strategies to improve our ability to predict production metrics. We propose a scenario-based data assimilation framework that integrates efficient geological modelling, high-fidelity production simulation, and ensemble-based data assimilation to explore, quantify, and constrain uncertainties in reservoir properties (porosity and permeability) as well as in reservoir performance forecasts for fluvial geothermal systems (Song et al., 2025a, b). The method combines prior equiprobable geological realizations, forward simulations, and field observations to generate revised posterior realizations for uncertainty estimation. Observation data can enhance our understanding of the fluvial geothermal system. However, acquiring such data would significantly increase the budget for data collection (Dausman et al., 2010) and can be very disruptive for the local population in the vicinity for a geothermal system. This raises a crucial question: how can we identify the most informative data to collect to ensure that reservoir performance forecasts for geothermal systems are reliable?

Building on our existing scenario-based data assimilation framework, this study evaluates the impact of different data types, including wellhead temperatures and pressures, as well as temperature and pressure profiles measured along a well, on reducing uncertainty in a doublet fluvial geothermal system. This process, known as data worth analysis, aims to quantify the amount of information that a data observation scheme is expected to provide. The evolution of uncertainty in reservoir performance forecasts is assessed at two randomly selected validation points based on simulated temperature and pressure. We design four observation schemes that combines wellhead temperatures and pressures with downhole temperature and pressure profiles along the injection well, production well, and a monitoring borehole. The comparative data worth analysis ranks these schemes according to their effectiveness in reducing predictive uncertainty, thereby offering constructive guidance for data collection in the field.

Method

To investigate the impact of different observations on uncertainty reduction in reservoir forecasts, we applied the scenario-based data assimilation framework to one specific scenario of a channelised fluvial geothermal system (Figure 1). The geological model is inspired by the Delft Sandstone reservoir, which is a fluvial Lower Cretaceous formation located approximately 2 km beneath the campus. The model domain has a dimension of 3 km × 2 km × 120 m. A standard doublet system—comprising one injection vertical well and one vertical production well—is used. We also consider a monitoring borehole, which is planned to be drilled at the Delft campus geothermal project. Detailed settings for the reference model are provided in Song et al. (2025a).

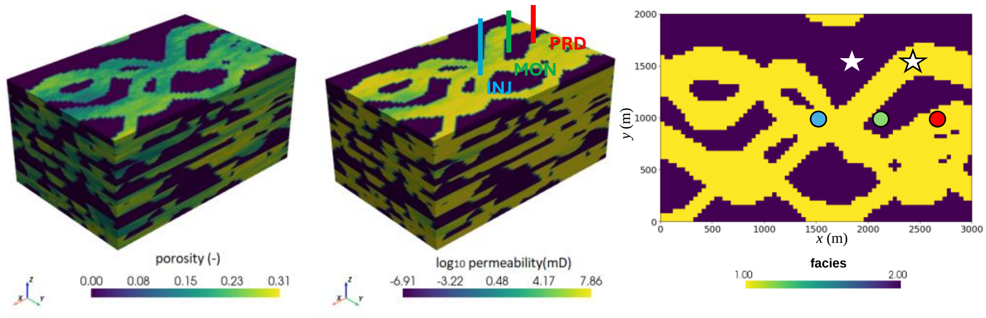

***Figure 1** A fluvial geothermal reservoir model with heterogeneous porosity (left) and permeability (centre) distribution. This particular model is used to generate the reference production data for this study, i.e. is considered to be the "truth case". The two stars on the top surface (right) indicate randomly selected validation points used to assess the reduction of uncertainty in temperature and pressure estimates at coordinates.*

In our simulations we solve the standard governing equations for single-phase mass and energy conservation in a geothermal systems with open-DARTS (Khait and Voskov, 2017), which applies a fully implicit time discretization and a finite-volume spatial discretization using the two-point flux approximation. For parameter estimation, we use the Ensemble Smoother with Multiple Data Assimilation (ESMDA) to invert porosity and permeability within the fluvial channels (Emerick and Reynolds, 2013). It integrates prior equiprobable realizations, forward simulations, and 'observed' data to produce posterior realizations for uncertainty quantification. The method improves upon the iterative ensemble smoother by introducing a series of inflation factors applied to the observation error covariance matrix, thereby controlling the update magnitude at each iteration and reducing the risk of overfitting. Details on simulation and ESMDA parameters can be found in (Song et al., 2025a, b).

Data worth analysis quantifies the amount of information that a data observation scheme is expected to provide. It is usually based on the change of uncertainty in the reservoir performance forecast from prior to the posterior ensemble. The "worth" of each observation in the calibration dataset is ranked against that of others based on its effectiveness in reducing the uncertainty associated with one or more predictions of particular interest (Dausman et al., 2010). To quantify the data worth of different measurements, we introduce a commonly considered information metric, the trace $\mathrm{Tr}_{\mathrm{reduction}}$ of the predictive covariance matrix (Neuman et al., 2012) given by

$$\mathrm{Tr}_{\mathrm{reduction}} = \mathrm{Tr}(C_1) - \mathrm{Tr}(C_2), \tag{1}$$

where Tr indicates the trace (i.e., sum of diagonal entries) of a matrix. C_1 is the prior predictive covariance, while C_2 represents the posterior predictive covariance after data assimilation. A larger $\mathrm{Tr}_{\mathrm{reduction}}$ indicates a greater information gain and thus higher data worth, implying a greater contribution of the observation data to reducing uncertainty and improving understanding of the geothermal system. The use of ensemble-based data assimilation allows direct computation of the prior and posterior variances in reservoir performance prediction from the ensemble statistics. Temperature and pressure at random validation points (Figure 1) access the data worth of observation schemes.

Results

Table 1 summarizes the four observation schemes considered in this study. Case 1 uses 50 years of production temperature and injection bottom-hole pressure, representing the most cost-effective option. In Cases 2 to 4, distributed sensors along the wells are assumed to provide temperature and pressure profiles over the 50-year period.

***Table 1** Case design based on observation types and their associated data worth.*

Case	Observation information				Data worth	
	Production temperature and injection pressure	Distributed temperature and pressure			$\mathrm{Tr}_{\mathrm{reduction}}$ based on temperature $(\mathrm{K})^2$	$\mathrm{Tr}_{\mathrm{reduction}}$ based on pressure $(\mathrm{bar})^2$
		Injection well	Production well	Monitoring borehole		
1	√				353.62	0.94
2	√		√		851.2	1.17
3		√	√		1078.9	1.92
4		√	√	√	1462.3	3.02

Take Case 4 as an example where temperature and pressure profiles along 3 wells (injector, producer, and monitoring borehole) are assimilated. Figure 2 illustrates the prior predictions, posterior predictions, and reference data for the recorded data along the injection and production wells and the monitoring borehole at different times. After ESMDA, the posterior ensemble aligns much better with the “observed” data compared to the prior ensemble. The uncertainty of the predictions is significantly constrained. Figure 3 illustrates the temporal evolution of production temperature and injection pressure for the prior and posterior ensembles, along with the reference model in Case 4. Although the production data were not directly assimilated in Case 4, the posterior production temperature and injection pressure still align more closely with the reference data. These results highlight the capability of ESMDA to refine reservoir performance forecasts, constrain uncertainty, and improve the estimation of reservoir dynamics in channelised fluvial geothermal systems.

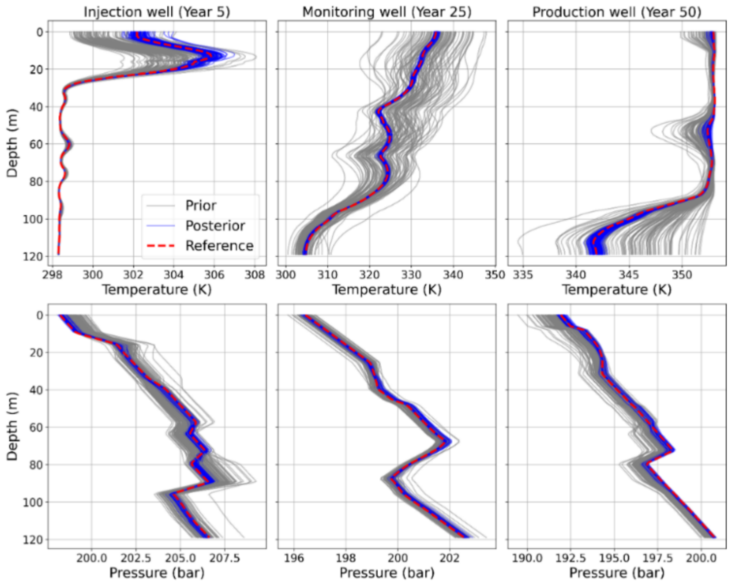

Figure 2 *Prior, posterior, and reference temperature and pressure profiles recorded as a function of depth for the injection well, monitoring borehole, and production well, after 5, 25, and 50 years of production, respectively (from left to right).*

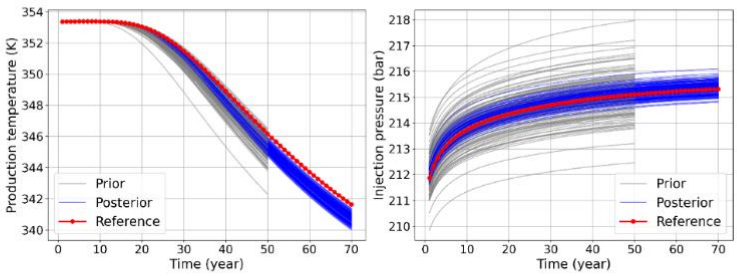

Figure 3 *Comparison of prior, posterior, and reference production temperature (left) and injection pressure (right) for Case 4.*

Figure 4 compares the prior, posterior, and reference temperature and pressure at two validation points for Case 1 and 4. In Case 4, the posterior predictions closely match the reference values and exhibit significantly reduced uncertainty compared to the prior ensemble. In contrast, Case 1, which only assimilates production outlet temperature and injection inlet pressure, shows minimal uncertainty reduction at the same validation points, indicating limited improvement in reservoir characterization and understanding of the reservoir architecture. These results demonstrate that the data collection scheme in Case 1 has substantially lower data worth compared to the more informative well-profile-based strategy in Case 4.

Table 1 summarizes the data worth for four observation schemes. Data worth ranking is Case 4 > Case 3 > Case 2 > Case 1. The comparison between Cases 1 and 2 highlights the importance of incorporating well temperature and pressure profiles. These profiles provide richer spatial and temporal information that better captures reservoir dynamics. This underscores the necessity of deploying distributed sensors along the wells. Furthermore, the comparison between Case 3 and Case 4 shows that including observations from the monitoring well further enhances data worth. The monitoring well provides valuable intermediate information between the injection and production wells and can partially capture the migration of the cold front, contributing to improved uncertainty reduction.

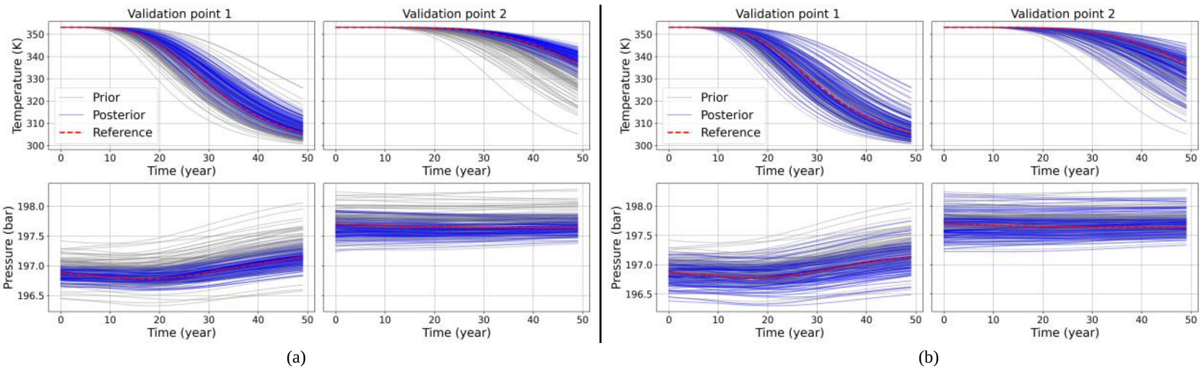

Figure 4 *Prior, posterior, and reference temperature and pressure at two validation points for (a) Case 4 and (b) Case 1.*

Conclusions

The proposed scenario-based data assimilation framework is shown to be effective in reducing both observation mismatch and production uncertainty in a fluvial geothermal system. Building on this, a data worth analysis was performed on a reference reservoir architecture to evaluate the impact of different observation schemes on predictive uncertainty. Results show that the total production temperature and injection pressure alone—though the most cost-effective option—are insufficient to significantly reduce predictive uncertainty. In contrast, distributed well temperature and pressure measurements demonstrate substantially higher data worth, leading to markedly improved uncertainty reduction. Additionally, incorporating a monitoring borehole provides further information gain, by capturing subsurface dynamics between the injection and production wells. These findings underscore the importance of monitoring pressure and temperature profiles in the wells of a geothermal doublet.

Acknowledgements

This research was supported by the TU Delft Excellence Foundation and Energi Simulation.

References

Dausman, A.M., Doherty, J., Langevin, C.D., et al.: Quantifying data worth toward reducing predictive uncertainty, Groundwater, 48(5), (2010), 729-740.

Emerick, A.A.and Reynolds, A.C.: Ensemble smoother with multiple data assimilation, Computers & Geosciences, 55, (2013), 3-15.

Khait, M.and Voskov, D.V.: Operator-based linearization for general purpose reservoir simulation, Journal of Petroleum Science and Engineering, 157, (2017), 990-998.

Neuman, S.P., Xue, L., Ye, M., et al.: Bayesian analysis of data-worth considering model and parameter uncertainties, Advances in Water Resources, 36, (2012), 75-85.

Song, G., Geiger, S., Voskov, D., et al.: Assessing the ensemble smoother with multiple data assimilation for subsurface fluvial geothermal systems, 50th Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, California (2025a).

Song, G., Geiger, S., Voskov, D., et al.: Scenario-based data assimilation framework to improve production estimates for geologically complex geothermal reservoirs, ESS Open Archive, (2025b).

Voskov, D., Abels, H., Barnhoorn, A., et al.: A research and production geothermal project on the TU Delft campus: initial modeling and establishment of a digital twin, 49th Workshop on Geothermal Reservoir Engineering, Stanford University, Stanford, California (2024).