Proceedings World Geothermal Congress 2020+1
Reykjavik, Iceland, April - October 2021

Quantifying Geothermal Drilling Risk Using a Monte Carlo Framework

Noah D. Athens and Jef Caers

Department of Geological Sciences, Stanford University

nathens@stanford.edu

Keywords: Uncertainty Quantification, Numerical Modeling, Drilling Risk

ABSTRACT

Our present ability to explore, develop, and monitor geothermal energy resources is hindered by the complexity of the geologic systems that house these resources. In order to improve decision making about where to drill geothermal wells, we apply a novel approach called Bayesian Evidential Learning to quantify uncertainty of temperature predictions for a prospective well using data from a nearby well. In this method, the relationship between data and prediction variables is learned by generating a training set of data using Monte Carlo simulation of a numerical model, which is then used to sample posterior temperature predictions without any explicit model inversion. We present results for different locations of the prospective temperature well, illustrating the spatial extent that the observed data reduces uncertainty on the prediction. This methodology is both practical and broadly applicable to other problems requiring full uncertainty quantification using spatially complex numerical models.

1. INTRODUCTION

Exploration and development risk remain significant challenges for expanding geothermal resources for renewable energy. Due to the fact that drilling is limited by the financial cost, reducing development risk depends on our ability to accurately integrate information from geological, geophysical, and geochemical investigations into a robust prediction and risk assessment framework. In recent years, many studies have focused on GIS-based aggregation methods to predict locations that are favorable for development, which are particularly useful for regional assessments (Cambazoğlu et al., 2019; Abdel Zaher et al., 2018; Ito et al., 2017; Trumpy et al., 2015). As exploration moves into the development stage, regional assessments are commonly refined by physics-based numerical simulations to identify specific targets or evaluate different production scenarios (Franco and Vaccaro, 2014). However, full uncertainty quantification of complex spatial problems using physics-based numerical models often present computational challenges that may limit how often this kind of analysis is performed in real applications (Chen et al., 2014).

In this paper, we present a hypothetical (but realistic) case study problem to demonstrate how a Monte Carlo based framework for uncertainty quantification can be applied to important questions in geothermal exploration and development. Specifically, we focus on addressing the problem of predicting subsurface temperatures constrained to a previously drilled temperature well in a fault-controlled extensional geothermal system. While we are focusing on temperature prediction, the same approach could also be used to predict other variables of interest, such as the location of high permeability zones, either separately or jointly with temperature. We first develop a prior probability model for model parameters based on data and conceptual understanding of fault-controlled geothermal systems. Next, we use Monte Carlo to generate a training set of data by sampling from the prior model and simulating temperatures using a groundwater and heat transfer numerical model. Finally, we use the Monte Carlo model realizations to establish a statistical relationship between the data variable (a temperature well) and the prediction variable (a prospective temperature well), which we use to generate posterior predictions of temperature. This general approach is termed Bayesian Evidential Learning (BEL), and has been employed in other recent subsurface modeling applications (Hermans et al., 2018; Scheidt et al., 2018).

2. CASE STUDY

In this section, we introduce a case study problem based on the extensively studied history of development in Dixie Valley, Nevada (Iovenitti et al., 2016; Wanner et al., 2014). While three deep temperature wells (> 2 km depth) in the northern part of Dixie Valley reached temperatures in excess of 250 °C, three deep temperature wells in southern Dixie Valley only reached a maximum temperature of ~75 °C (Person et al., 2008). This disparity in temperature in an otherwise similar geologic setting highlights the potential risk of drilling for conventional geothermal resources in a fault-controlled system.

The set-up of the case study problem is as follows: a geothermal company would like to use a previously drilled temperature well to predict temperature in a prospective well at some nearby location to decide if it is worth drilling (Fig. 1a). The approach will be to apply BEL to predict temperature in the prospective well while accounting for both conceptual uncertainty and parameter uncertainty in the numerical model used to simulate temperature. These assumptions will be specified in a prior probability model, and checked through a process of falsification described in the next section.

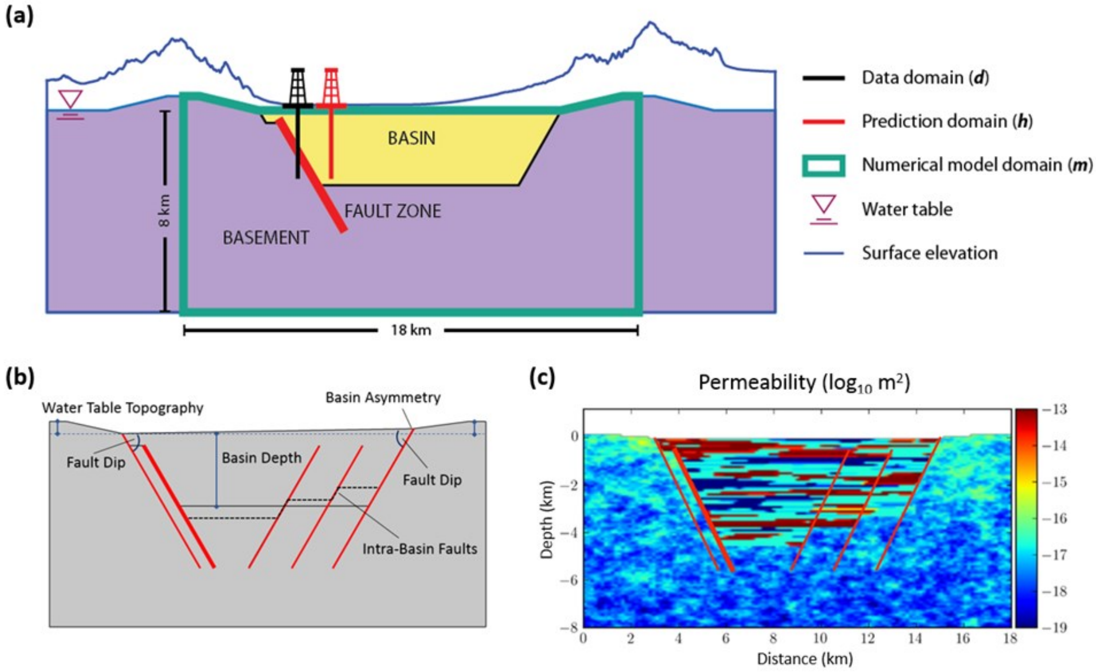

Figure 1: (a) Case study problem set-up. Surface elevation is not to scale. (b) Numerical model domain showing the geometry parameterization specified in Table 1. (c) Example model realization drawn from the prior probability model. Color scale represents permeability.

2.1 Statement of Prior Uncertainty

In the BEL protocol, prior uncertainties are specified in a prior probability model that is used to generate a training set of data and prediction variables using Monte Carlo simulation. Oftentimes the specification of the prior model is one of the most challenging aspects of Bayesian analysis because it requires that all uncertainty is explicitly represented. In subsurface, model-based uncertainty quantification, this means that both the model parameterization and probability distribution for each parameter must be specified a priori.

For our case study problem, the prior model of uncertainty (Table 1) includes 5 parameters related to geometry and structure, 10 parameters related to material properties such as permeability and thermal conductivity, and 2 parameters related to boundary conditions. The parameterization is intended to reflect realistic uncertainty about the hydrologic setting, basin structure, and stratigraphy in an extensional geothermal basin based on published data, previous modeling, and reasonable assumptions. A complete description of the prior model of uncertainty is provided in Athens and Caers (2019).

Table 1: Prior model parameters with prior uncertainty used for hydrothermal simulations

Parameter Type	Parameter Name	Parameter Distribution[1]
Geometry Properties	Basin Depth	U[3, 4] km
	Water Table Topography	U[0.1, 1] km
	Fault Dip	U[55, 70] degrees
	Basin Asymmetry	U[0, 0.1] km
	Intra-Basin Faults	Discrete[0, 1]
Material Properties	Basement Permeability Mean	U[-17.5, -14.5] $\log_{10}$ m^2
	Basement Permeability Correlation Length	U[0.1, 1] km
	Basement Permeability Variance	U[0.05, 0.5] $\log_{10}$ m^2
	Basin Variogram Model	Discrete[Gaussian, Exponential]
	Basin Variogram Correlation Length	U[0.1, 15] km
	Percent Confining	U[10, 40] %
	Percent Reservoir	U[10, 40] %
	Basin Permeability	U[-17, -14] $\log_{10}$ m^2
	Fault Permeability	U[-14, -12] $\log_{10}$ m^2
	Basement Thermal Conductivity	U[2, 3] W m^{-1} K^{-1}
Boundary Conditions	Basal Heat Flux	U[0.075, 0.15] W m^{-2}
	Regional Groundwater	Discrete[0, 1, 2]

[1]U[a, b] refers to a continuous uniform distribution over the specified interval. Discrete[a, b, c] refers to a discrete uniform distribution.

2.2 Numerical Modeling Approach

The numerical modeling approach to simulate temperature draws from previous numerical modeling by Wisian and Blackwell (2004). However, we extend their model in several significant ways to include additional parameter uncertainty. Numerical modeling was performed using COMSOL Multiphysics®, a finite-element modeling package. The governing equations for single-phase, variable-density groundwater flow and heat transfer in porous media, as well as the thermodynamic equations of state for water, are standard and implemented using COMSOL's Subsurface Flow Module and Heat Transfer Module.

The two-dimensional basin-scale model is composed of three primary domains: (1) crystalline basement rock, (2) sedimentary basin, and (3) fault zones. However, fault zones are not explicitly meshed because they are extremely narrow relative to the extent of the modeling domain. Instead, fault zones are represented using COMSOL's fracture flow interface, which couples a 1D solution of the groundwater and heat transfer equations within the fault zone to the surrounding mesh elements.

Boundary conditions for the top of the domain are specified constant temperature of 20 °C and 1 Atmosphere pressure, representing the water table boundary. The sides of the model are assigned inflow and outflow velocities specified in the prior model (Table 1), but do not allow heat transfer. Boundary conditions for the bottom of the domain include a specified basal heat flux (Table 1) and no flow boundary condition. Initial conditions establish a geothermal gradient of 40 °C/km and hydrostatic pressure. Simulations are run for 1 million years, effectively reaching steady state conditions.

2.3 Monte Carlo and Falsification

Monte Carlo simulation was used to generate 750 model realizations (of which 693 converged). For each realization, we extract the data variable and prediction variable (i.e. 1D temperature wells) located at 5 and 6 km in the model domain (Figure 2a). From the realizations, we selected a reference model that represents the true earth in order to illustrate the prediction performance of BEL. We will refer to the temperature profile drawn from the reference model as the observed data ($\boldsymbol{d_{obs}}$).

Following Monte Carlo simulation, the next step is to check whether the specified prior probability model is appropriate, which is determined by whether or not the distribution of simulated data reflects the distribution of actual data. To do this we compared the simulated data to published temperature wells located throughout Dixie Valley (Fig. 2b and c) and verified that the range of temperature is similar. Although presented as a single step, this process of developing the prior probability model is usually iterative, oftentimes requiring increasing the complexity and/or the parameter uncertainty in the prior model multiple times. This method of checking the prior model, termed falsification, has been used in many previous studies (Hermans et al., 2018; Aydin and Caers, 2017).

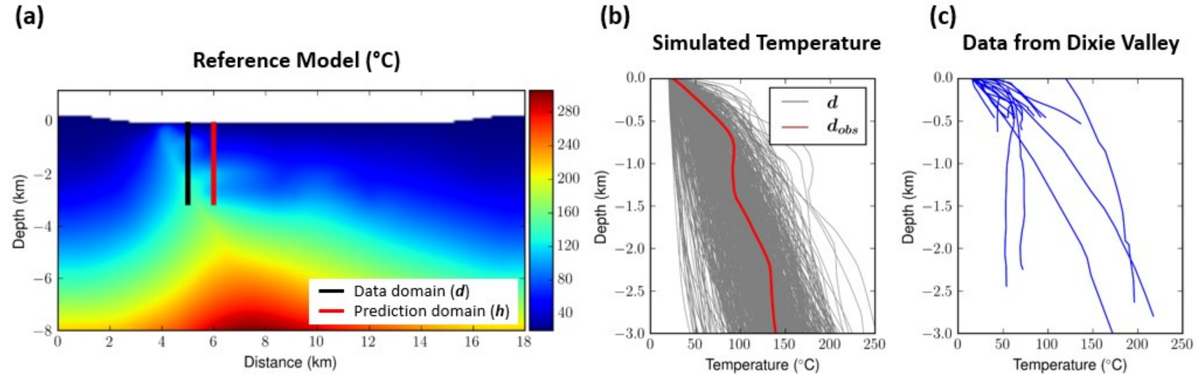

Figure 2: (a) Reference model showing the location of the data domain (black) and prediction domain (red). (b) Prior distribution of the data variable (1D temperature well) where $\boldsymbol{d_{obs}}$ (red) is the data variable from the reference model. (c) Actual published temperature well data from Dixie Valley, Nevada (Person et al., 2008; National Geothermal Data System).

2.4 Inversion by Direct Forecasting

The basic idea of inversion by direct forecasting is to use the training set of data and prediction variables generated by Monte Carlo to learn the multi-variate distribution between those variables $f(\mathbf{h}|\mathbf{d})$. Next, the conditional distribution $f(\mathbf{h}|\mathbf{d}_{obs})$ is calculated from the joint multi-variate distribution, which is used to generate posterior samples of $\boldsymbol{h}$. By directly learning the relationship between data and prediction, there is no need to perform any conventional model inversion.

To learn the multi-variate relationship, we use the method of Satija and Caers (2015). In this approach, data and prediction variables are transformed to a reduced dimension space using Principal Component Analysis (PCA) followed by Canonical Correlation Analysis (CCA) (Fig. 3a). Next, 10,000 samples of the prediction $\boldsymbol{h}$ are drawn from a multi-variate Gaussian distribution at the location of $\boldsymbol{d_{obs}}$. Because PCA and CCA are both bijective transformations, samples of $\boldsymbol{h}$ can be reconstructed to their original dimension (Fig. 3c). This entire process is computationally efficient, requiring only a few seconds to execute.

As shown in Figure 3c, uncertainty in predicted temperature is significantly reduced from the prior uncertainty, indicating that the observed temperature well is informative on the prediction. It can also be seen that the reference temperature profile (red line) drawn from the reference model (i.e. the true earth model) is within the posterior prediction distribution, which is an indication that the method of inversion by direct forecasting is producing reasonable results.

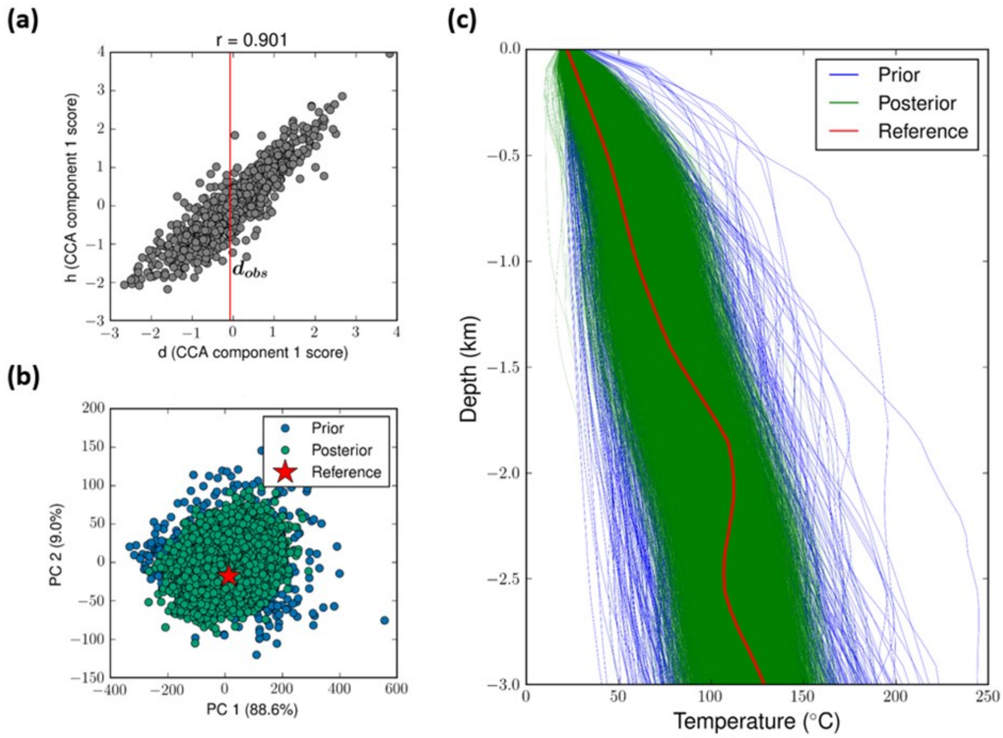

Figure 3: Direct Forecasting summary (a) Scatter plot of the first canonical component of the data and prediction variables. (b) Scatter plot of the prior and posterior distribution in a reduced dimension. (c) Prior and posterior temperature profiles at the location of the prediction domain (Fig. 2a). The posterior temperatures are reconstructed from the reduced dimension PCA space.

3. RESULTS

The results are summarized by showing the bottom hole temperature distributions for 9 well locations, spanning 2 km to either side of the observed temperature well and perpendicular to the strike of the main fault (Figure 4). As expected, the posterior prediction at the location of the observed temperature well has no uncertainty because the observed data is matched in the inversion by direct forecasting. However, as the location of the prediction well is moved away from the observed well, the uncertainty of the prediction increases. In fact, at the location of 3 km, the range of uncertainty of the posterior is approximately equivalent to that of the prior, indicating that the observed temperature data is not informative at this location of the model domain. Nevertheless, at the location of 7 km, the posterior uncertainty is still less than the prior uncertainty. Since both of these prospective wells are 2 km from the data well, this shows that there cannot be a simple rule-of-thumb to calculate the distance over which a temperature well is informative, but rather that it is dependent on the hydrothermal setting and must be quantified. The fact that the reference temperature is significantly outside of the mean of the posterior in most of the proposed well locations is normal, and not indicative of a problem with the posterior prediction.

It is important to note that these results do not generalize to other directions of offset from the main fault. For example, one could imagine that the uncertainty in the along-strike direction of the main fault to be far more constrained than in the perpendicular direction presented here. One way to test this would be to develop a 3D model. Although this would increase the computational complexity, conceptually the same methods used in this study, including inversion by direct forecasting, would apply to 3D models as well.

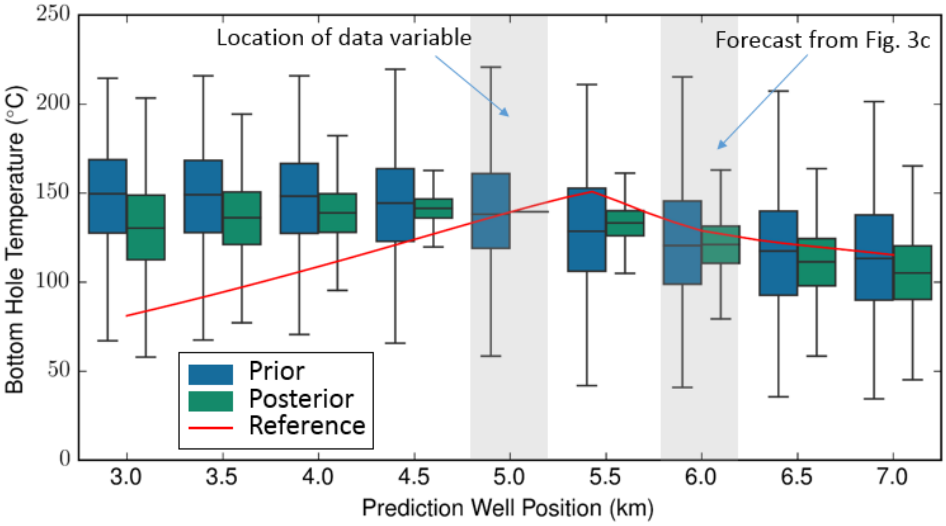

Figure 4: Predicted Bottom Hole Temperature distributions at 9 separate proposed well locations. The extent of the boxes indicate the 25th and 75th quantile, and the extent of the fliers indicate the full data range after removing outliers. For each of the proposed well location predictions, the data variable is located at 5 km as illustrated in Fig. 2a.

4. CONCLUSIONS

In this study, we used a realistic case study problem to show how to apply the BEL protocol to quantify temperature uncertainty in a prospective well location using data. Our primary contribution was to show how to develop a prior model that can be used to generate model realizations using Monte Carlo simulation. Next, we learned the statistical relationship between our data and prediction variables by using PCA and CCA to reduce the dimensionality of the problem. Finally, we fit a multi-variate Gaussian model to sample the posterior temperature predictions, which were reconstructed to their original dimension. By using Monte Carlo and learning the statistical relationship between data and prediction variables, we were able to generate posterior predictions without any traditional model inversion or parameter tuning.

REFERENCES

Athens, N.D., and Caers, J.: A Monte Carlo-based framework for assessing the value of information and development risk in geothermal exploration, *Applied Energy*, **256**, (2019), pp. 11.

Aydin, O., and Caers, J.K.: Quantifying structural uncertainty on fault networks using a marked point process within a Bayesian framework, *Tectonophysics*, (2017), 101-124.

Abdel Zaher, M., Elbarbary, S., El-Shahat, A., Mesbah, H., and Embaby, A.: Geothermal resources in Egypt integrated with GIS-based analysis, *Journal of Volcanology and Geothermal Research*, **365**, (2018), 1-12.

Cambazoğlua, S., Yal, G.P., Eker, A.M., Şen, O., and Akgüna, H.: Geothermal resource assessment of the Gediz Graben utilizing TOPSIS methodology, *Geothermics*, **80**, (2019), 92-102.

Chen, M., Tompson, A.F.B., Mellors, R.J., Ramirez, A.L., Dyer, K.M., Yang, X., and Wagoner, J.L.: An efficient Bayesian inversion of a geothermal prospect using a multivariate adaptive regression spline method, *Applied Energy*, **136**, (2014), 619-627.

Franco, A., and Vaccaro, M.: Numerical simulation of geothermal reservoirs for the sustainable design of energy plants: A review, *Renewable and Sustainable Energy Reviews*, **30**, (2014), 987-1102.

Hermans, T., Nguyen, F., Klepikova, M., Dassargues, A., and Caers, J.: Uncertainty Quantification of Medium-Term Heat Storage From Short-Term Geophysical Experiments Using Bayesian Evidential Learning, *Water Resources Research*, **54**, (2018), 2931-2948.

Iovenitti, J., Ibser, F.H., Clyne, M., Sainsbury, J., Callahan, O.: The Basin and Range Dixie Valley Geothermal Wellfield, Nevada, USA – A test bed for developing an Enhanced Geothermal System exploration favorability methodology, *Geothermics*, **63**, (2016), 195-209.

Ito, G., Frazer, N., Lautze, N., Thomas, D., Hinz, N., Waller, D., Whittier, R., and Wallin, E.: Play fairway analysis of geothermal resources across the state of Hawaii: 2. Resource probability mapping, *Geothermics*, **70**, (2017), 393-405.

Person, M., Banerjee, A., Hofstra, A., Sweetkind, D., and Gao, Y.: Hydrologic models of modern and fossil geothermal systems in the Great Basin: Genetic implications for epithermal Au-Ag and Carlin-type gold deposits, *Geosphere*, **4**, (2008), 888-917.

Satija, A., and Caers, J.: Direct forecasting of subsurface flow response from non-linear dynamic data by linear least-squares in canonical functional principal component space, *Advances in Water Resources*, **77**, (2015), 69-81.

Scheidt, C., Li, L., and Caers, J.: Quantifying uncertainty in subsurface systems, *American Geophysical Union*, (2018), pp. 279.

Trumpy, E., Donato, A., Gianelli, G., Gola, G., Minissale, A., Montanari, D., Santilano, A., and Manzella, A.: Data integration and favourability maps for exploring geothermal systems in Sicily, southern Italy, *Geothermics*, **56**, (2015), 1-16.

Wanner, C., Peiffer, L., Sonnenthal, E., Spycher, N., Iovenitti, J., and Kennedy, B.M.: Reactive transport modeling of the Dixie Valley geothermal area: Insights on flow and geothermometry, *Geothermics*, **51**, (2014), 130-141.

Wisian, K.W., and Blackwell, D.D.: Numerical modeling of Basin and Range geothermal systems, *Geothermics*, **33**, (2004), 713-741.