



HAL
open science

Comparing Well and Geophysical Data for Temperature Monitoring Within a Bayesian Experimental Design Framework

Robin Thibaut, Nicolas Compaire, Nolwenn Lesparre, Maximilian Ramgraber,
Eric Laloy, Thomas Hermans

► **To cite this version:**

Robin Thibaut, Nicolas Compaire, Nolwenn Lesparre, Maximilian Ramgraber, Eric Laloy, et al.. Comparing Well and Geophysical Data for Temperature Monitoring Within a Bayesian Experimental Design Framework. *Water Resources Research*, 2022, 58 (11), 10.1029/2022WR033045 . hal-03863085

HAL Id: hal-03863085

<https://hal.science/hal-03863085>

Submitted on 21 Nov 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Abstract

Temperature logs are an important tool in the geothermal industry. Temperature measurements from boreholes are used for exploration, system design, and monitoring. The number of observations, however, is not always sufficient to fully determine the temperature field or explore the entire parameter space of interest. Drilling in the best locations is still difficult and expensive. It is therefore critical to optimize the number and location of boreholes. Due to its higher spatial resolution and lower cost, four-dimensional (4D) temperature field monitoring via time-lapse Electrical Resistivity Tomography (ERT) has been investigated as a potential alternative. We use Bayesian Evidential Learning (BEL), a Monte Carlo-based training approach, to optimize the design of a 4D temperature field monitoring experiment. We demonstrate how BEL can take into account various data source combinations (temperature logs combined with geophysical data) in the Bayesian optimal experimental design (BOED). To determine the optimal data source combination, we use the Root Mean Squared Error (RMSE) of the predicted target in the low dimensional latent space where BEL is solving the prediction problem. The parameter estimates are accurate enough to use in BOED. Furthermore, the method is not limited to monitoring temperature fields and can be applied to other similar experimental design problems. The method is computationally efficient and requires little training data. For the considered optimal design problem, a training set of only 200 samples and a test set of 50 samples is sufficient.

Plain Language Summary

The design of experiments is a critical step in scientific research to ensure that the data collected is of sufficient quality to answer the research question. The design of an experiment is often optimized to minimize the cost while still fulfilling the desired accuracy in the prediction. In this contribution, we use an approach called Bayesian Evidential Learning to optimize the design of an experiment. The Bayesian philosophy is used to incorporate all available information into the design of the experiment and to quantify the uncertainty in the prediction. In particular, our approach makes it straightforward to use different data sources (sensors), and different types of data (e.g., temperature, pressure, and concentration measurements). We demonstrate the method in the design of a temperature field monitoring experiment to characterize the geothermal energy storage potential of an aquifer, a key step in the development of geothermal energy projects. This framework makes experiment design easier and faster, a prerequisite for robust risk analysis and decision-making. The method is not limited to temperature fields and can be applied to other similar experimental problems, such as the monitoring of a contaminant plume or saltwater intrusion.

1 Introduction

Geothermal systems, including borehole thermal energy storage (BTES) and shallow aquifer thermal energy storage systems (ATES) are becoming more popular as the world looks for ways to reduce greenhouse gas emissions. Such systems use thermal energy extracted from the ground or groundwater to heat or cool buildings, which necessitates some electrical energy input for the heat pump, while storing the excess heat or cold underground. The goal is to re-use this thermal energy during the next season in a cyclic utilization (Saner et al., 2010; Vanhoudt et al., 2011; Bayer et al., 2013; Duijff et al., 2021). The performance of BTES and ATES strongly depends on the subsurface properties. Many variables are involved in geothermal processes, including porosity, hydraulic conductivity, thermal conductivity, and heat capacity. Subsurface temperature fluctuations are strongly influenced by the spatial distribution of these parameters, the boundary conditions, and the aquifer’s hydraulic gradient when modeling the underground response under thermal stress (Ferguson, 2007; Bridger & Allen, 2010; W. Sommer et

68 al., 2013; W. T. Sommer et al., 2014). Previous research demonstrated that time-lapse
69 Electrical Resistivity Tomography (ERT) could monitor spatial temperature changes in
70 the subsurface with a relatively large spatial coverage by utilizing variations in resistiv-
71 ity caused by temperature changes (Hermans et al., 2012, 2014; Arato et al., 2015; Her-
72 mans et al., 2015; Lesparre et al., 2019; Robert et al., 2019). In turn, ERT monitoring
73 experiments can be used to predict the response of the subsurface to thermal exploita-
74 tion (Hermans et al., 2018, 2019).

75 This paper presents a methodology for improving the design of field experiments
76 using well and geophysical data by utilizing Bayesian Evidential Learning (BEL). BEL
77 is a technique that combines Monte Carlo simulations and machine learning that can be
78 used to improve estimation of predictions uncertainty (Hermans et al., 2016, 2018; Michel
79 et al., 2020; Thibaut et al., 2021). It performs the Bayesian inference (usually within a
80 low-dimensional latent space) using a direct relationship between predictor (data) and
81 target (prediction) learned from a training set sampled from the prior distribution. Its
82 efficiency has been demonstrated with extensive synthetic validation, but also against
83 rejection sampling (Scheidt et al., 2015), MCMC algorithms (Michel et al., 2020, 2022),
84 field data (Hermans et al., 2019), and experimental design (Thibaut et al., 2021). Pre-
85 vious research has shown that BEL can estimate the posterior distribution of paramet-
86 ers in a variety of contexts, including geothermal systems (Hermans et al., 2018, 2019;
87 Athens & Caers, 2019), contaminant transport (Satija & Caers, 2015; Scheidt et al., 2015),
88 geophysical inversion (Hermans et al., 2016; Michel et al., 2020). In addition, the BEL
89 framework has been successfully applied to a range of subsurface field cases, such as ground-
90 water, shallow and deep geothermal and oil/gas predictions (J. Park & Caers, 2020; Prad-
91 han & Mukerji, 2020; Tadjer & Bratvold, 2021).

92 To forecast the behavioral patterns of such processes, mathematical, physics-based
93 models are frequently used. These models are typically described by partial differential
94 equation systems, but are often too complex to solve analytically and must be solved nu-
95 merically. These simulations produce a prediction defined on a grid of locations in space
96 and/or time, which is conditioned on the uncertainty of the model parameters. Typi-
97 cally, model outputs are compared to field data via forward modeling in order to cali-
98 brate the model unknown parameters, which is known as inverse modeling (Tarantola,
99 2005, 2006). In prediction-focused approaches, such as BEL, the prediction is directly
100 estimated from the data (Hermans, 2017). In both strategies, data must be collected through
101 an experimental process. Because this process is typically costly and/or time-consuming,
102 it is critical to optimize resource utilization and information gain by controlling the ex-
103 perimental conditions for data acquisition (Attia et al., 2018). This is referred to as op-
104 timal experimental design (OED), which is commonly viewed as an optimization prob-
105 lem (Ryan et al., 2016; Kleinegesse & Gutmann, 2020; Haan et al., 2021; Lykkegaard &
106 Dodwell, 2022).

107 By combining prior knowledge of the model’s unknown parameters with likelihood-
108 the data’s contribution to those parameters—Bayesian statistics produce the posterior dis-
109 tribution, from which it is possible to draw conclusions about the model’s unknown pa-
110 rameters. Applying the Bayesian paradigm to OED at the definition of the objective func-
111 tion leads to a so-called Bayesian OED (BOED) approach. There are several approaches
112 to solving (B)OED problems, all of which involve maximization or minimization of a util-
113 ity function (Lindley, 1956; Chaloner & Verdinelli, 1995; Ryan et al., 2016). However,
114 the computational burden is significant because BOED requires solving the inverse prob-
115 lem for each possible dataset, as the posterior distribution must first be estimated to com-
116 pute the Bayesian utility function. Surrogate modeling and Bayesian Model Averaging
117 (BMA) in combination with preposterior estimation are the two main simplifications that
118 have been suggested in the field of Earth Sciences to make posterior computation tractable.
119 Thibaut et al. (2021) provide an overview of these two simplifications. Iterative approaches
120 requiring multiple forward problem runs, such as Markov chain Monte Carlo (MCMC)

121 methods, are commonly used to solve the problem (Laloy & Vrugt, 2012; Vrugt, 2016;
 122 Ryan et al., 2016). Zhang et al. (2015), for example, use a MCMC algorithm to solve the
 123 BOED of sampling well locations and source parameters identification of groundwater
 124 contaminants, but defined a surrogate for the contaminant transport equation to reduce
 125 the computational burden of MCMC. Tarakanov and Elsheikh (2020) developed a BOED
 126 methodology aimed at subsurface flow problems relying on a polynomial chaos expansion
 127 surrogate model for the utility function embedded in a MCMC algorithm.

128 The main benefit of BEL is that the inferred predictor-target relationship can be
 129 applied to any data set that is consistent the prior distribution. This is advantageous
 130 for BOED because the prediction problem can be solved quickly and without expensive
 131 computation for any newly proposed data set. Thibaut et al. (2021) demonstrated, for
 132 instance, how BEL can be used to predict the wellhead protection area (WHPA) sur-
 133 rounding pumping wells. The WHPA (target) was predicted using a small number of trac-
 134 ing experiments (predictors). To determine the optimal data source design, the predic-
 135 tive power of the posterior distribution was evaluated in the original high-dimensional
 136 space.

137 In this contribution, we propose a methodology for solving a BOED problem in-
 138 volving two different data types, using the BEL framework in a low-dimensional latent
 139 space, which alleviates the “curse of dimensionality” (Bellman, 1961) and reduces com-
 140 putational and memory demands. This study is based on the monitoring of a heat injection-
 141 storage-pumping experiment monitored by time-lapse Electrical Resistivity Tomogra-
 142 phy (ERT) and well data mimicking the field experiment of Lesparre et al. (2019). In
 143 addition to having the potential to improve subsurface information, the fusion of var-
 144 ious data types can also be used to plan surveys and choose the optimal set of instru-
 145 ments (JafarGandomi & Binley, 2013). There is a tradeoff between non-invasive data ac-
 146 quisition methods such as ERT and invasive data acquisition methods such as drilling,
 147 with the latter being generally more expensive. It is critical to determine whether drilling
 148 is required or if a simple geophysical survey will suffice, depending on the specific prob-
 149 lem at hand. It can also be useful to know whether it is worthwhile to conduct a long
 150 geophysical survey and mobilize a large amount of equipment if we can use an existing
 151 borehole in the area.

152 This contribution is unique in five ways when compared to previous applications
 153 of BEL to experimental design:

- 154 1. We use two different type of predictors, geophysical and borehole data, which vary
 155 in time, to predict a four-dimensional target, the temperature field magnitude in
 156 the aquifer over time. It will first be demonstrated how to predict the target us-
 157 ing each data set separately.
- 158 2. Then, it will be demonstrated how to use them together. Because the predictors
 159 are of different types, the dimensionality reduction step in the pre-processing sec-
 160 tion is applied to each instance separately, and then concatenated before being fed
 161 to the learning algorithm.
- 162 3. The Transport Map method is used to sample in the low dimensional space.
- 163 4. The data utility function is calculated in the low dimensional latent space, with-
 164 out the need to back-transform the data to its original space.
- 165 5. We identify the optimal choice between two ERT protocols with our proposed method-
 166 ology.

167 Overall, BEL is well-suited to solve the considered BOED problem at a low cost (con-
 168 sidering both time and computational demand).

2 Methodology

2.1 BEL

The BEL framework has been extensively explained in contributions such as Hermans et al. (2018) or Scheidt et al. (2018), and we will only summarize the fundamental principles in this contribution. The objective of BEL is to determine the target H 's posterior probability distribution $p(H|D_{obs})$, conditioned on an observed predictor D_{obs} . Both the target and the predictor are real, multidimensional random variables. First, training sets of H and D are generated by sampling the prior distribution through forward modeling. Then, the target and the predictor are separately subjected to dimensionality reduction through Principal Component Analysis (PCA) and projected to the new principal component (PC) space. Next, Canonical Correlation Analysis (CCA) analyses the PCs to find their underlying correlations while transforming them to new, maximally correlated canonical variates (CVs). Given that the CVs' relationships describe the behavior of each target dimension for each predictor dimension, an observed data point can be used to infer the posterior distribution of each unknown target dimension. To do this, we first project the observation onto the data CV axes using the same transformations we derived for the data samples. Then, for each target dimension, we use the corresponding bivariate distribution and the observed CV to derive the posterior distribution of that dimension. Thibaut et al. (2021), for example, demonstrated how this can be easily accomplished using multivariate Gaussian inference, provided that the CVs' bivariate distributions are both Gaussian and linear. Kernel Density Estimation (KDE) is another method for approximating the bivariate distribution for each CCA dimension without requiring such assumptions to be verified (e.g., Hermans et al. (2019); Michel et al. (2020)). KDE, on the other hand, has two parameters that must be adjusted: the kernel type, which defines the shape of the distribution at each coordinate, and the kernel bandwidth, which describes the size of the kernel at each position. Transport maps (Spantini et al., 2018) form another method that allows to calculate the posterior distribution and is used in this work (§2.3.3).

2.2 Experimental setup

We present a method for estimating the posterior distribution of an unknown four-dimensional temperature field during a heat injection-storage-pumping experiment. The experiment has been described in Lesparre et al. (2019). In short, hot water was injected at 3 m³/h into an aquifer at a temperature of 42 °C for six hours, followed by another injection at 14.5 °C for 20 minutes. Then, it was stored in the aquifer for 92 hours, and then pumped back out for 16 hours and 15 minutes at a flow rate of 3 m³/h. The goal was to track the evolution of temperature distribution in the aquifer over time. The experiment was monitored using time-lapse ERT and temperature data in wells. The time-varying predictor (measured voltage for ERT, direct temperature for wells) and the target (temperature distribution in the aquifer) are high-dimensional, and their relationship is non-linear. By performing geothermal field experiments such as injection and hot water pumping tests, we can gain a better understanding of the aquifer's behavior (Palmer et al., 1992; Macfarlane et al., 2002; Vandenbohede et al., 2009, 2011; V. Wagner et al., 2014; Wildemeersch et al., 2014; B.-H. Park et al., 2015; Klepikova et al., 2016). Combined with these tests, geophysical and thermal monitoring can track heat transfer in the aquifer. The location of injection wells is the focal point of the hydrogeological models used in this study. These models' grids are 60 meters in length in the known direction of natural aquifer flow and 40 meters in length in the perpendicular direction to that direction (Figure 1A). The grid layers begin at the surface of the aquifer's saturated zone at a depth of 3 meters and conclude at the surface of the impermeable basement at 10 meters. The space step along the Z axis (depth) is 0.5 meters. The space step along the X and Y axes ranges from 2.5 centimeters at the injection point to 2.5 meters at the model's

220 edges, with a 0.25 centimeter refinement in a 3 meter radius around the injection points
 221 in the hydraulic flow direction (Y axis).

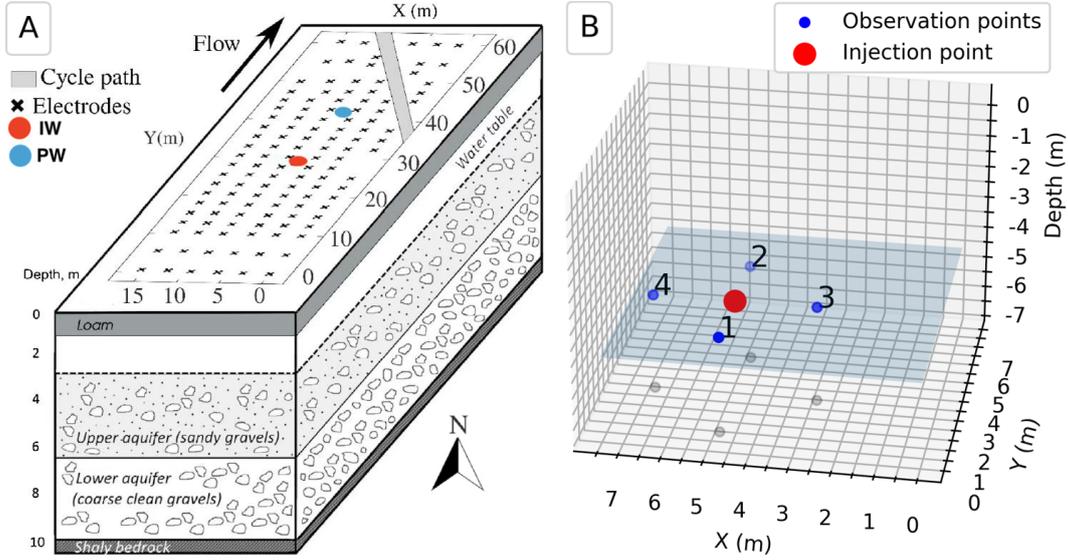


Figure 1. A. Model design (modified from Lesparre et al. (2019)). IW: Injecting well. PW: Pumping well. B. Positions of observation and injection wells.

222 The mean and variance of hydraulic conductivity (K), the anisotropy factor, and
 223 its direction in the horizontal plane and the range of the spherical variogram are used
 224 to build the models of the hydraulic conductivity distribution in the aquifer. In addition,
 225 the homogeneous porosity and the natural gradient are set as uncertain parameters.
 226 All other parameters including water and matrix thermal properties are set as constant.
 227 Bulk thermal properties are calculated based on porosity using the arithmetic average.
 228 The prior is the sum of all the definitions of the parameters under consideration.
 229 We detail the ranges of the variables in Table 1.

Uncertain parameter	Range
Log K mean	$U[-4, -1]$, K in m/s
Log K variance	$U[0.05, 2]$, K in m/s
Effective porosity	$U[0.05, 0.3]$
Variogram main range	$U[1, 10]$ m
Anisotropy ratio	$U[0, 0.5]$
Orientation	$U[0, \pi]$ rad
Natural gradient	$U[0.05, 0.167]$ %

Table 1. Parameters of the prior model. $U[a, b]$ refers to the continuous uniform distribution bounded by the values a and b .

Each parameter is generated randomly and independently according to a uniform distribution. The sequential Gaussian simulation algorithm (Goovaerts, 1997) is then used to generate the hydraulic conductivity fields. In the direction of Y, the hydraulic flow at the grid boundaries is zero, and it respects the natural gradient in the X-direction. The direction of the flow due to the natural gradient is indicated on Figure 1A. For this investigation, 250 hydrogeological models were built and used to generate the temperature fields resulting from the injection-storage-pumping experiment. The HydroGeoSphere code was used to simulate the temperature field in the hydraulic conductivity grids during the injection-storage-pumping experiment (Brunner & Simmons, 2012). There are 106-time observations in each simulation. The temperature grids obtained are reduced to sub-grids around the injection well of size $16 \times 16 \times 14$ of elementary volumes $0.5 \text{ m} \times 0.5 \text{ m} \times 0.5 \text{ m}$ to reduce the amount of data (Figure 1B). The temperature is averaged when the subgrid elemental volumes contain several elemental volumes from the initial grid.

The following equations were used to transform the temperature T into conductivity (Hermans et al., 2014):

$$\frac{\sigma_{f,T}}{\sigma_{f,25}} = m_{f,25}(T - 25) + 1, \text{ and} \quad (1)$$

$$\Delta T = \frac{1}{m_{f,25}} \left[\frac{\sigma_{b2,T}}{\sigma_{b1}} \frac{\sigma_{f1}}{\sigma_{f,25}} - 1 \right] + 25 - T_{init}, \quad (2)$$

where $m_{f,25}$ is the fractional change of the fluid conductivity per degree Celsius around the reference temperature of 25°C , σ_{b1} is the background conductivity, $\sigma_{b2,T}$ is the background conductivity at temperature T which we aim to determine at each time-step, σ_{f1} is the fluid conductivity at the initial state, $\sigma_{f,25}$ is the fluid conductivity at 25°C , and T_{init} is the initial temperature.

The synthetic study replicates a real ERT monitoring campaign performed on a well documented site (Dassargues, 1997; Derouane & Dassargues, 1998; Brouyère, 2001; Hermans et al., 2015; Klepikova et al., 2016). Therefore, the initial temperature was measured directly in the field, and the background resistivity σ_{b1} was determined using their inversion using real field data. The temperature and fluid conductivity trend from on-site water samples was calculated to be $m_{f,25}=0.0194$ and $\sigma_{f,25}$ was estimated to be 0.0791S/m (Hermans et al., 2015; Lesparre et al., 2019). $\sigma_{f1} = 0.0614\text{S/m}$ was estimated from Equation 2 with an initial temperature of 13.44°C .

The forward modeling process went as follows:

1. Temperature field simulations on the discretized grid (HydroGeoSphere, Brunner and Simmons (2012)).
2. Computation of the conductivity with the petrophysical law for each simulation and time step (Equation 2).
3. Computation of the ERT profiles for each simulation and time step (EIDORS, Polydorides and Lionheart (2002)).

The ERT simulations are computed with the following experimental setup: 6 parallel profiles of 21 electrodes in the X direction, with a 2.5 meter spacing between the 17 central electrodes and a 5-meter spacing between the four electrodes at the profile's edge. Two protocols were used: multiple gradient (MG) and dipole-dipole (DD). The DD array was chosen due to its widespread use in ERT investigations, which is due in part to the low electromagnetic coupling between the circuits. The MG array was another appealing option due to its sensitivity distribution and high S/N ratio, as well as its robustness and multi-channel compatibility (Dahlin & Zhou, 2006). For each of the 106 temperature simulation time observations, an ERT simulation is generated for each model. At each time step, 1100 quadrupoles are used for the DD protocol and 848 quadrupoles are used for

273 the MG protocol. The target for each model is thus made of $16 \times 16 \times 14$ temperature grids
 274 for the 106 observation times, and the predictor is made of the resistances measured by
 275 1100 or 848 quadrupoles for the 106 observation times.

276 **2.3 Heat prediction**

277 **2.3.1 Pre-processing**

278 **2.3.1.1 Target** The target H is a three-dimensional temperature field subdivided
 279 into $n_{rows} \times n_{cols} \times n_{lay}$ over n_{step} observation time steps. It is critical to reduce dimen-
 280 sionality because some small-scale variations of the target do not need to be perfectly
 281 reconstructed as they are already beyond the predictor's resolution. Before performing
 282 dimension reduction, the raw target is scaled to unit variance, because the dimension re-
 283 duction step is sensitive to the scale of the data. The dimension reduction itself is done
 284 by linear PCA. The principal components are new variables produced by combining the
 285 initial variables in a linear way.

286 **2.3.1.2 Predictor** The geophysical predictor D_g is made up of resistance values
 287 measured by n_{quad} quadrupoles over n_{step} observation time steps, and the borehole pre-
 288 dictor D_b is made of temperature curves measured at observation well's locations. Be-
 289 fore performing dimension reduction, both raw predictors are scaled to unit variance. The
 290 dimension reduction is performed by linear PCA, and the number of components to keep
 291 is automatically set by setting the amount of variance that needs to be explained to 99%.
 292 When working with geophysical data, filtering higher dimensions allows to reduce the
 293 effect of noise on the prediction (Hermans et al. (2016); Michel et al. (2020)).

294 After dimensionality reduction, the PCs of both predictor and target are scaled to
 295 unit variance, because covariance matrices are sensitive to the scale of the data.

296 **2.3.2 Training**

After dimensionality reduction, the next step is to find the relationship between
 the predictor and the target in the reduced space. The CCA algorithm projects the data
 onto a new set of axes which maximize the correlation between the two data sets. The
 transformed variables are called Canonical Variates (CVs). Similarly to PCA, each pair
 of transformed variables is orthogonal (uncorrelated) with the other pairs. The CVs are
 the transformed variables representing the mutual information between the two data sets.
 Let δ be the number of PCs necessary to explain the required amount of variance in the
 predictor, and $n_{training}$ be the number of pairs of predictors and targets used for train-
 ing. In our case, δ is the sum of the number of components of the geophysical data and
 the borehole temperature curves, i.e., $\delta = \delta_g + \delta_b$. This is simply done by concatenat-
 ing the geophysical and borehole temperature PCs into a single matrix. Following pre-
 processing, the model is trained with the CCA algorithm to establish a multivariate re-
 lationship between D and H in PC space. The number of CCA components is set to δ ,
 the maximum number that can be used (Meloun & Militký, 2012). The number of com-
 ponents to keep depends on the dimensionality reduction of the predictor. To allow back-
 transformation, more components must be used for the predictor than for the target be-
 fore learning the relationship between the two. However, to avoid overfitting and noise
 propagation, it is recommended that both have a similar number of components. It does
 not have to be strictly the same. Let the superscript c denote the canonical space. The
 canonical variates (CVs) pairs are stored in the $(n_{training} \times \delta)$ matrices

$$D_\delta^c = d_{i,1}^c, d_{i,2}^c, \dots, d_{i,\delta}^c | i = 1, \dots, n_{training} \quad (3)$$

$$H_\delta^c = h_{i,1}^c, h_{i,2}^c, \dots, h_{i,\delta}^c | i = 1, \dots, n_{training} \quad (4)$$

297 With the pairs of canonical variates $(d_{:,1}^c, h_{:,1}^c) \sim \pi_1$ to $(d_{:,\delta}^c, h_{:,\delta}^c) \sim \pi_\delta$ established,
 298 we may infer the posterior in the canonical subspaces by independently conditioning each

299 of the resulting bivariate joint distributions $\pi_j, j = 1, \dots, \delta$ on a new observation pro-
 300 jected into the canonical predictor space $d_{new,j}^c$. In this study, we consider a new approach
 301 based on triangular transport, which offers a good tradeoff between computational ef-
 302 ficiency and accuracy.

303 2.3.3 Conditioning with transport methods

In essence, transport methods seek a monotone, invertible transport map \mathbf{S} that transforms samples from a target distribution π_j into samples from a simpler, user-specified reference distribution η , typically a standard multivariate normal distribution $N(0, I)$, where I is the identity matrix. This map allows us to sample conditionals of the target distribution, and thus implements the conditioning operation we are principally interested in. Transport methods are a nuanced topic, and the interested reader is referred to (Villani, 2009; El Moselhy & Marzouk, 2012; Spantini et al., 2018) for a more detailed discussion of their theoretical properties. In this study, we focus only on the parts necessary for the conditioning operation. Following Spantini et al. (2022) and Baptista et al. (2022), the map \mathbf{S} is triangular, meaning it has as many parameterized map components as there are dimensions in π_j (two in this case, so $\mathbf{S} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$). Each map component depends on one more dimension of the target’s probability density function (pdf) than its predecessor. In our setting, the map is structured as:

$$\mathbf{S}(d_{:,j}^c, h_{:,j}^c) = \begin{bmatrix} S_1(d_{:,j}^c) \\ S_2(d_{:,j}^c, h_{:,j}^c) \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \mathbf{z} \quad (5)$$

304 where \mathbf{z} are samples distributed according to the reference pdf η . These map component
 305 functions are made up of user-specified combinations of polynomials or radial basis func-
 306 tions, and must be monotone in their last argument. This means that $\partial_{d_{:,j}^c} S_1(d_{:,j}^c) >$
 307 0 and $\partial_{h_{:,j}^c} S_2(d_{:,j}^c, h_{:,j}^c) > 0$. This will ensure that the entire map \mathbf{S} remains monotone
 308 and thus invertible. This monotonicity requirement can be ensured with diligent param-
 309 eterization of the map components (e.g., Equation 2 in Baptista et al. (2022)).

If we are only interested in conditioning, we only need to define, optimize, and evaluate the second map component $S_2 : \mathbb{R}^2 \rightarrow \mathbb{R}$ (Spantini et al., 2022). We optimize this map component S_2 by minimizing the Kullback-Leibler divergence between the target pdf π_j , presumed to be known only through samples, and the map’s approximation to the target pdf, which is obtained by sending the standard Gaussian reference η through the inverse map \mathbf{S}^{-1} . From this, we can derive the following objective function \mathcal{J} (Spantini et al., 2018):

$$\mathcal{J}(S_2) = \sum_{i=1}^{n_{training}} \left(\frac{1}{2} (S_2(d_{i,j}^c, h_{i,j}^c))^2 - \log \frac{\partial S_2(d_{i,j}^c, h_{i,j}^c)}{\partial h_{i,j}^c} \right). \quad (6)$$

310 The forward evaluation of S_2 is a simple evaluation of its constituent basis functions. We
 311 can also invert S_2 efficiently through an appropriate one-dimensional root finding algo-
 312 rithm. Note that while the forward map components S_1 and S_2 can be evaluated inde-
 313 pendently, their inverses must be evaluated in sequence:

$$\mathbf{S}^{-1}(\mathbf{z}) = \begin{bmatrix} S_1^{-1}(z_1) \\ S_2^{-1}(d_{:,j}^c, z_2) \end{bmatrix} = \begin{bmatrix} d_{:,j}^c \\ h_{:,j}^c \end{bmatrix} \quad (7)$$

314 where the inverse of the second map component S_2^{-1} depends on the outcome $d_{:,j}^c$
 315 of the first inversion S_1^{-1} . This dependence makes triangular transport highly useful for
 316 Bayesian inference. In fact, the inverse map \mathbf{S}^{-1} factorizes the target pdf π_j according
 317 to the ordering of the variables in Equation 5 as $\pi_j(d_{:,j}^c, h_{:,j}^c) = \pi_j(d_{:,j}^c) \pi_j(h_{:,j}^c | d_{:,j}^c)$,
 318 where the second term on the right-hand side ($\pi_j(h_{:,j}^c | d_{:,j}^c)$) corresponds to the poste-
 319 rior and is sampled by the map component inverse S_2^{-1} (e.g., Section 7.1 of Spantini et
 320 al. (2018)).

321 In other words, this means that we can sample conditionals of π_j by simply replac-
 322 ing the argument $d_{:,j}^c$ during the map component inversion S_2^{-1} with any values on which
 323 we want to condition. For instance, supplying duplicates of the observation $d_{obs,j}^c \mathbf{1}^\top$, where
 324 $\mathbf{1}$ is a column vector of 1s, we can sample the desired conditional $\pi_j(h_{i,j}^c | d_{obs,j}^c)$:

$$\mathbf{t}_{:,j}^{c,\text{cond.}} = S_2^{-1}(z_2; d_{obs,j}^c \mathbf{1}^\top) \sim \pi_j(h_{i,j}^c | d_{obs,j}^c). \quad (8)$$

325 Note that the required reference samples z_2 can be either drawn from a standard
 326 Gaussian distribution, or (better) obtained from the forward map $z_2 = S_2(d_{:,j}^c, h_{:,j}^c)$
 327 (see Spantini et al. (2022)). With the conditioned samples $\mathbf{t}_{:,j}^{c,\text{cond.}}$ for each pair of co-
 328 variates $(d_{:,j}^c, h_{:,j}^c)$, we can then back-transform the posterior samples into the original
 329 target space by ascending from CCA, PCA, and undoing any transformations.

330 2.4 Experimental design

331 BEL can be used to assess the amount of information delivered by various data sources.
 332 The actual data can have any value within the prior data space, and data sources can
 333 be placed anywhere across the grid. To identify highly informative data sets based on
 334 their location, one data-utility function must be maximized or minimized (Thibaut et
 335 al., 2021). To restrict computation time and keep some realism, we will only consider
 336 four well positions around the injection well, which is sufficient to demonstrate the ap-
 337 proach. With the geophysical data, a total of 31 different combinations of data sets are
 338 possible. These unique combinations are combined with a total of 50 unknown ground
 339 truths (test set). Each sample of the test set is sampled with 500 posterior samples. This
 340 number is arbitrarily chosen to be high enough to ensure that the posterior distribution
 341 is reasonably well-sampled while remaining low enough to keep the computation time
 342 manageable. We end up with an array of shape $(31, 50, n_{samples}, \delta) = (\text{number of com-}$
 343 $\text{binations, test set size, sample set size, number of canonical components})$.

344 One of the challenges of our case study is the high dimensionality of the problem.
 345 This is why we use the principal components to perform BOED, which is a novel approach
 346 in this context. The true target is transformed to the PC space and its RMSE with the
 347 predicted targets' PCs is computed. The advantage of working in a lower dimension is
 348 that we need to predict a smaller number of dimensions, which is computationally faster.
 349 Because the back-transformation from the PCs to the observation targets is a linear op-
 350 eration, it is simple to demonstrate that the prediction error is minimized if the distance
 351 between the PCs of the predicted and observed targets is also minimized. The minimiza-
 352 tion of the target prediction error is therefore equivalent to the minimization of the dis-
 353 tance between the PCs of the predicted and the observed targets. The PCs are weighted
 354 by their explained variance ratio during distance computation because they account for
 355 the different importance of the different components.

356 We validate our approach with k-fold cross validation. We repeat our computations
 357 across 5-folds and average the metric results. Averaging across folds is not strictly nec-
 358 essary but will be used here to increase the robustness of the BOED.

359 The methodology is summarized below:

360 **Input:** training set (D_{train}, H_{train}) , test set (D_{test}, H_{test}) , data utility function UF
 361 **Output:** Averaged cross-validation rankings
 362 **for all** Fold f in 5-fold cross-validation **do**
 363 $H_{train,f} \leftarrow$ Target training data for fold f
 364 $H_{train,f} \leftarrow PCA_h.fit_transform(H_{train,f})$
 365 $H_{test,f} \leftarrow$ Target test data for fold f
 366 $H_{test,f} \leftarrow PCA_h.transform(H_{test,f})$
 367 **for all** Possible Combinations **do**

```

368    $O \leftarrow \text{Combination}$ 
369    $D_{train,f,O} \leftarrow \text{Predictor training data for fold } f \text{ and combination } O$ 
370    $D_{train,f,O} \leftarrow PCA_d.fit\_transform(D_{train,f,O})$ 
371    $D_{test,f,O} \leftarrow \text{Predictor test data for fold } f \text{ and combination } O$ 
372    $D_{test,f,O} \leftarrow PCA_d.transform(D_{test,f,O})$ 
373    $TrainedModel \leftarrow CCA.fit(D_{train,f,O}, H_{train,f})$   $\triangleright$  Training step
374   for all Ground Truths in  $(D_{test,f,O}, H_{test,f})$  do
375      $D_{true} \leftarrow \text{True predictor}$ 
376      $D_{true} \leftarrow PCA_d.transform(D_{true})$ 
377      $H_{true} \leftarrow \text{True target}$ 
378      $H_{true} \leftarrow PCA_h.transform(H_{true})$ 
379      $H_{posterior} \leftarrow TrainedModel.predict(D_{true})$   $\triangleright$  Predicting in PC space
380      $Utility \leftarrow UF(H_{posterior}, H_{true})$ 
381      $Results.add(Utility)$ 
382   end for
383    $Ranking \leftarrow add(Results)$ 
384    $Results.clear()$ 
385 end for
386 end for
387  $Final \leftarrow mean(Ranking)$ 

```

388 The method was developed in Python on a standard laptop with 16 GB of RAM
389 and a 2.30 GHz 8-core Intel Core i9 processor. Profiling has been performed on the code
390 for temperature prediction with the most complex combination (ERT combined with the
391 four wells). It took 38 seconds to initialize the model, load the dataset, pre-process the
392 data, and save the files. The prediction part of the experiment took 27 seconds, bring-
393 ing the total time to 65 seconds, including the time required to save the results to a file.
394 Throughout, the process used no more than 2.08 GB of RAM. Given that the entire im-
395 plementation was created for research purposes, the implementation can be enhanced
396 for speed and memory management (e.g., parallelization, low-level programming language
397 implementation, etc.). Finally, the computational costs differ depending on the case study
398 and the size and number of data sets used.

399 3 Application

400 3.1 Optimal protocol determination

401 Before determining the optimal sensor combination, we can determine the optimal
402 protocol between MG and DD following our methodology. Common approaches to op-
403 timizing electrode arrays for ERT either seek the best measurement configuration on a
404 given set of electrodes (e.g., Wilkinson et al. (2015)) or select electrode locations (e.g.,
405 F. M. Wagner et al. (2015)). Both strategies are based on the resolution matrix. Uhlemann
406 et al. (2018) propose a new approach that combines these two strategies by introducing
407 an additional weight that penalizes the addition of electrode locations to the optimized
408 set. Qiang et al. (2022) take a different approach. Instead of focusing on image resolu-
409 tion, they propose that the ERT survey be optimized for a specific target of interest by
410 maximizing the information gained from a target area using Bayesian experimental de-
411 sign.

412 In contrast to these methods, we seek to determine the optimal protocol for a given
413 set of electrodes, and not the electrode locations or quadrupole configurations themselves.
414 It could, however, be extended for that purpose. Furthermore, previous approaches use
415 the resolution matrix, whereas our approach uses the predictive power of the posterior
416 distribution in a low-dimensional latent space. Qiang et al. (2022) also makes use of the
417 posterior distribution's predictive power, but in a physical, high-dimensional space. Lastly,
418 these approaches are limited to a single target of interest, whereas our machine learn-

419 ing approach can generalize to unseen targets if the prior distribution is representative
 420 of the true distribution.

421 We consider four cases:

- 422 1. The MG array (848 quadrupole measures), to which PCA is applied to explain
 423 99% of the variance.
- 424 2. The DD array (1100 quadrupole measures), to which PCA is applied to explain
 425 99% of the variance.
- 426 3. Combination 1: concatenation of the MG and DD arrays to which PCA is applied
 427 to explain 99% of the variance.
- 428 4. Combination 2: union of the previously calculated principal components of the MG
 429 and DD configurations. Thus, it is equivalent to joining the principal components
 430 from items 1 and 2.

431 Combinations 1 & 2 are not among the available options for continuing the study,
 432 but they are included here for completeness. The computation is performed on the 50
 433 test cases over 5 different folds for each of the four cases, and the averaged results are
 434 shown in Figure 2. The MG array, despite having fewer quadrupoles than the DD ar-
 435 ray, is the best single protocol. In fact, of the four cases, the DD array has the lowest
 436 predictive power. Therefore, the MG array is selected as the optimal protocol for the study.
 437 Interestingly, Combination 1 has a lower predictive power than the MG array, although
 438 it contains more quadrupoles. The MG and DD configurations are not equivalent and
 439 produce different outcomes. As a result, when using PCA, joining the two configurations
 440 in this manner resulted in information loss. Combination 2, however, significantly im-
 441 proves the predictive power over the individual cases. It is explained by the fact that the
 442 two configurations are complementary and that the union of the previously calculated
 443 principal components of the two configurations captures the most meaningful informa-
 444 tion for the problem. The posterior distribution is better constrained as a result of the
 445 additional information, improving predictive ability.

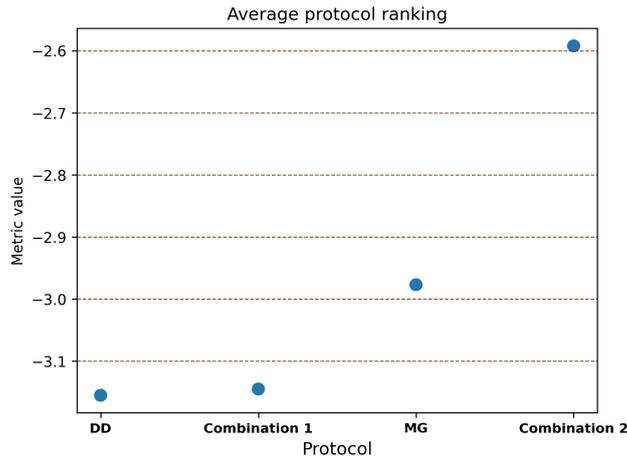


Figure 2. Average score of the different protocols over 5 folds. DD: PCA(Dipole-Dipole), MG: PCA(Multiple-Gradient). Combination 1: PCA(DD+MG), Combination 2: PCA(DD)+PCA(MG).

446

3.2 Target prediction

447

448

449

450

451

452

453

454

455

456

457

458

459

460

461

462

463

464

465

This section shows how to predict a single four-dimensional temperature field, using three distinct predictors: (i) the geophysical data, made of resistance measurements from 848 quadrupoles over 106 observation time steps, which are combined to form the predictor, (ii) a single temperature profile over 106 observation times, and (iii) the combination of the geophysical data with all 4 borehole temperature curves. We show that BEL can accurately estimate the target posterior distribution with varying uncertainty levels according to the type of predictor used. The dataset in this section has a total size of $n = 250$. The training set is then reduced to $n_{80\%} = 200$ models, with the remaining models being used to validate BEL's ability to predict the target, and later on for OED. Previous BEL applications have demonstrated that making accurate predictions with a dataset of this size is possible (Hermans et al., 2016, 2018, 2019; Athens & Caers, 2019; Michel et al., 2020; J. Park & Caers, 2020; Yin et al., 2020; Thibaut et al., 2021). While a small training set size is inevitable due to the time-consuming nature of the simulations, it is sufficient because the prediction is a temperature distribution that varies smoothly in both time and space and results from advection, diffusion, and dispersion processes. Such target is much simpler than the underlying K model. BEL is a Bayesian method that incorporates uncertainty. Therefore, a large training set size is not required and the method is more robust against overfitting. Furthermore, the use of cross-validation ensures the robustness of the results.

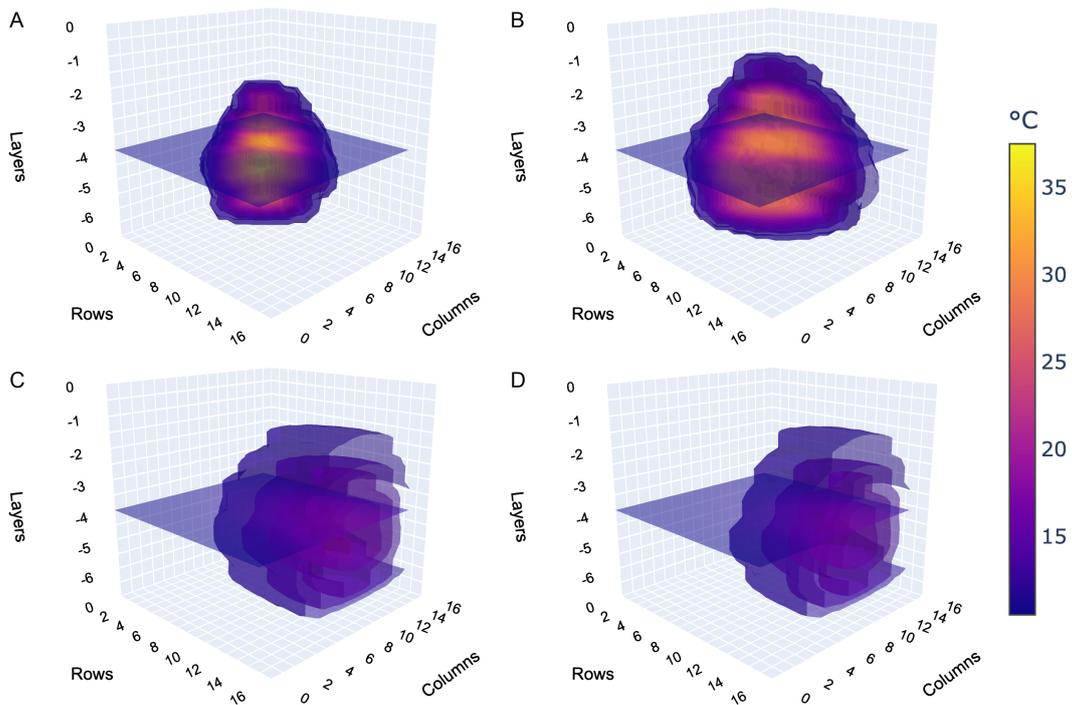


Figure 3. Snapshots of the temperature field at time steps 4 (A), 14 (B), 61 (C) and 74 (D) for one example. The injection well discharge is at (column, row, layer) = (9, 6, -5). The reference plane at the wells level is highlighted.

466

3.2.1 Pre-processing

467

468

The three predictors undergo the pre-processing described in §2.3.1. When geophysical data is combined with borehole temperature curves, the following steps are taken:

- 469 1. PCA is performed to explain 99% of the variance in the geophysical data to ob-
 470 tain the PCs $D_{G,99}$
 471 2. PCA is performed to explain 99% of the variance in each of the n_b boreholes tem-
 472 perature curves to obtain the PCs $D_{B,99}^{(i)}$ for each borehole i . These PCs are then
 473 concatenated into a single array $D_{B,99}$.
 474 3. The geophysical PCs are concatenated with the borehole temperature curves PCs
 475 to obtain the array $D_{G,B,99}$

476 We have to perform PCA separately for each predictor, because if we simply con-
 477 catenated the predictor arrays and then applied PCA, the resulting predictor would be
 478 more representative of the geophysical data than the borehole temperature curves. Fur-
 479 thermore, the PCs are scaled to unit variance before concatenation because the PCs mag-
 480 nitude can vary greatly from predictor to predictor. The resulting principal components
 481 for all cases can directly be compared in Figure 4. The left column shows the predictor
 482 PCs, and the right column the target PCs for all cases. For each PC dimension (X-axis),
 483 the PC value for each of the 200 training instances is plotted (Y-axis). The example test
 484 instance is plotted on top of the training instances. The required number of PCs is cho-
 485 sen to account for 99 percent of the variance in the predictor. In parallel, PCA is ap-
 486 plied to the target. In line with our methodology, the number of components to keep is
 487 set to the maximum number of PCs required by the predictor, and the corresponding
 488 amount of variance explained in the target is summarized in Table 2.

Parameter \ Combination	G	1	1, 2	1, 2, 3	1, 2, 3, 4	G, 1, 2, 3, 4
δ	10	3	7	10	13	23
Explained variance (target)	77%	57%	71%	77%	80%	88%

Table 2. Effect of the number of PCs (δ) on the target PCA explained variance. **G** stands for geophysical data. **1, 2, 3, 4** stand for the borehole temperature curves. Case (i) is **G**, case (ii) is **1** and case (iii) is **G, 1, 2, 3, 4**.

489 When a single temperature profile is used as a predictor (Case (ii)), the target vari-
 490 ance is not explained as well as when the entire set of predictor data is used (Case (iii)),
 491 hence the smaller number of target PCs required. When the predictor is a combination
 492 of geophysical data and all temperature profiles, some additional variance is captured.
 493 The remaining target PCs for each case are shown in Figure 4B, D and C. These addi-
 494 tional target components are not used in training. They are, however, saved for subse-
 495 quent use in OED.

496 3.2.2 Training and prediction

497 The CCA mapping allows the fusion of the low-dimensional representation of the
 498 predictor with the low-dimensional representation of the target in order to make predic-
 499 tions of the target distribution. It is run on all cases to find the canonical variates that
 500 define the relation between the predictor and the target, using a maximum of 23 PCs
 501 on the predictor. The first three canonical variate pairs for each case are shown in Fig-
 502 ure 5. Each row corresponds to a single case, and each column to a single canonical vari-
 503 ate pair. The variable ρ is the correlation between the canonical variates, and can be in-
 504 terpreted as the amount of mutual information shared between the target and predic-
 505 tor. For all cases, the predictor explains a large part of the target: ρ_1 (first pair) = 0.999,
 506 0.959 and 1. for case (i), (ii) and (iii) respectively. Hence, CCA is well-suited for our pur-
 507 poses. The canonical variates for case (ii) provide a poorer explanation for the target (ρ_2

508 (second pair) = 0.462). This is because the temperature curves used as predictors are
 509 one-dimensional and only convey information about a small portion of our model over
 510 time. The strong correlations in the canonical variates pairs of cases (i) and (iii) is fur-
 511 ther indication of the suitability of CCA for this study. Transport map inference is run
 512 for each pair for each case, using the test predictor (on the X-axis (d^c) of Figure 5). The
 513 500 samples of $p(h^c|d_*^c)$ in Figure 5 represent our predictions of the unknown target dis-
 514 tribution in the canonical space. They can then be sequentially back-transformed to the
 515 principal component space (see posterior samples in Figure 4) and to the original space
 516 (see Figures 6 and 7 for 1D and 2D representations, respectively).

517 Figure 6 shows the temperature curves of the 500 samples at the location of the
 518 observation well number 2 for each case. This point was arbitrarily chosen to illustrate
 519 the one-dimensional temperature curve at one observation point. Across cases, the level
 520 of uncertainty, expressed by the spread of the temperature curves, is higher during the
 521 injection and storage phases than it is during the pumping phase, which is a positive de-
 522 velopment because this portion of the curve is the part that is typically inferred for ATEs
 523 systems (energy recovery). The magnitude of the spread of the predicted samples in the
 524 canonical space (Figure 5) is sequentially transmitted through the principal component
 525 space (Figure 4) and the original space (Figure 6), since the data flows through linear
 526 transformations.

527 On Figure 7A, the true (test) temperature cross-section at layer 9 (at a depth of
 528 4.5m) and time $t=105.75$ h (pumping phase) is shown. Figure 7B shows one randomly
 529 drawn example from the 500 sampled temperature profiles when using the ERT predic-
 530 tor alone. The results are not only visually close to the truth, but the absolute temper-
 531 ature difference ranges from -0.8 to 0.4 degrees Celsius, which is the magnitude of ac-
 532 curacy we expect when using resistance data (Hermans et al., 2015). Figures 7C and 7D
 533 show the same results when a single borehole temperature profile and the full combina-
 534 tion of ERT and temperature profiles are used, respectively. As expected, the single bore-
 535 hole image is the least accurate, while the combination of all temperature profiles and
 536 ERT is the most accurate.

537 Because the underlying internal behavior of the aquifer is nonlinear, given the num-
 538 ber of parameters involved, the underlying physics of the model and the nature of the
 539 predictor (indirect geophysical data), such a level of uncertainty is expected (Hermans
 540 et al., 2015, 2016, 2018). Considering the level of complexity of the underlying physics
 541 and the relatively small training set, BEL is capable of inferring different possible out-
 542 comes with reasonable uncertainty.

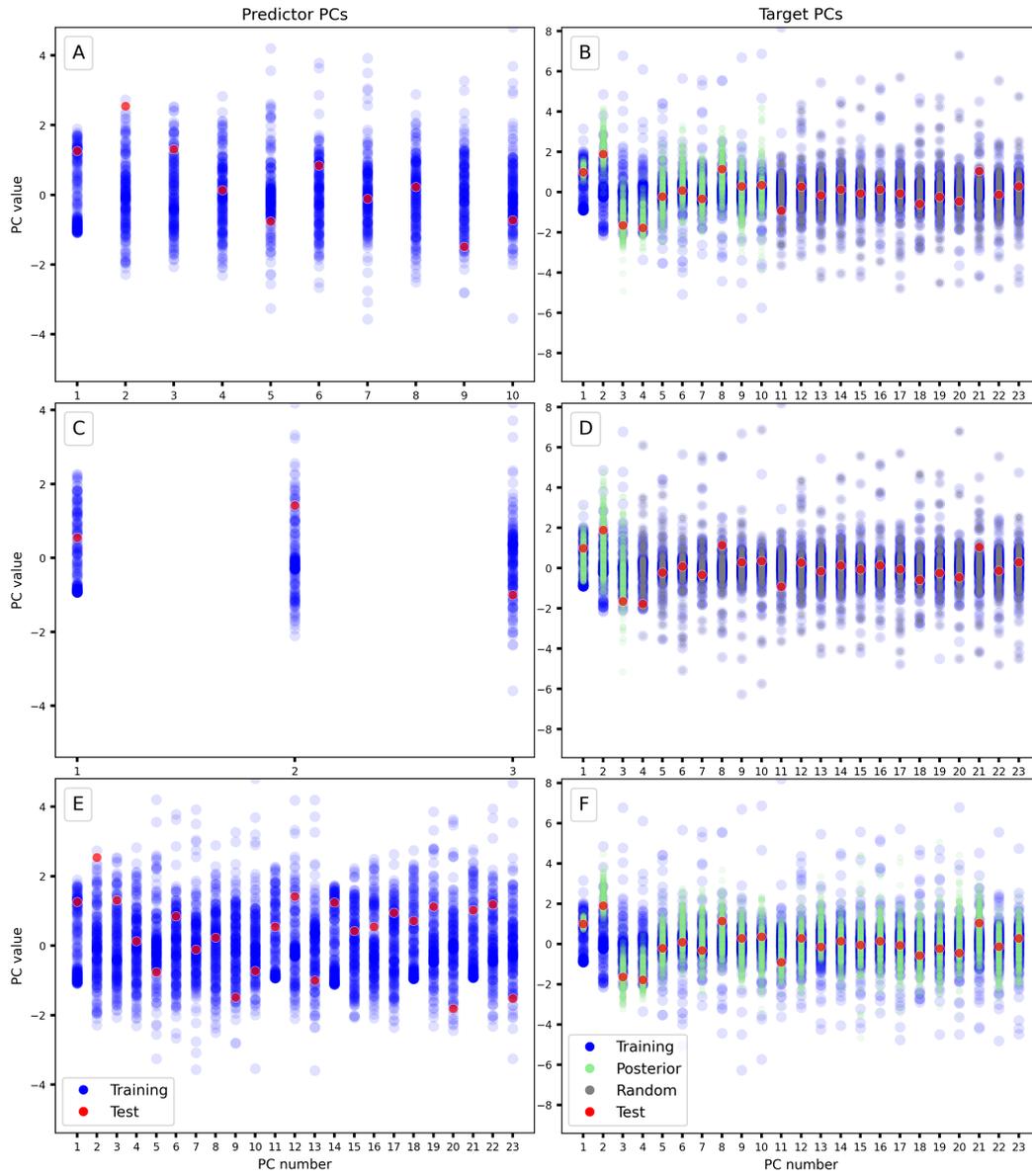


Figure 4. Principal Component Scores. The ‘Random’ PC samples are drawn at random from the target PC training set. They will be used in the OED section 4. A. Case (i). Predictor: ERT data, B. Case (i). Target. C. Case (ii). Predictor: Temperature profile from borehole 1. D. Case (ii). Target. E. Case (iii). Predictor: Full combination (ERT data + four boreholes temperature profiles). F. Case (iii). Target.

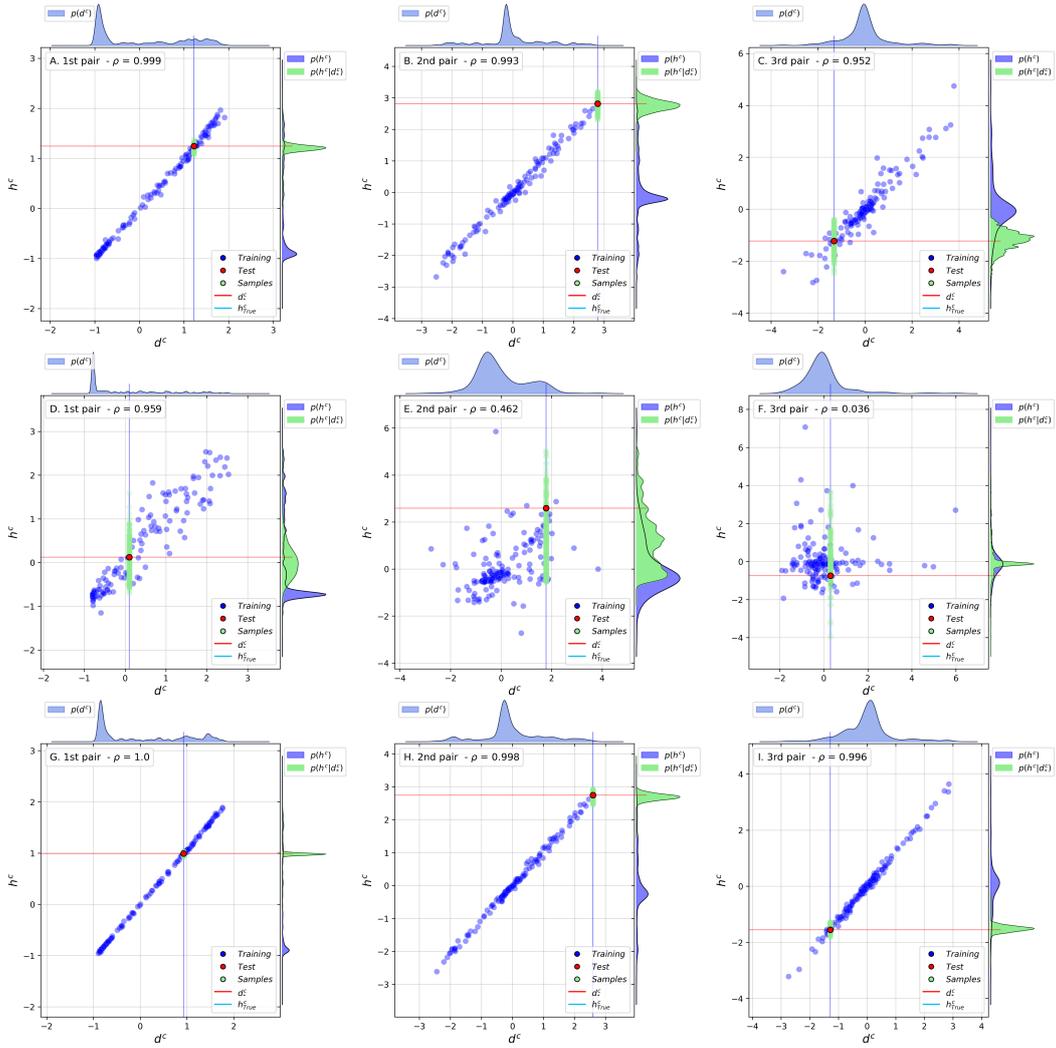


Figure 5. Canonical Variate Pairs (1 to 3). The first row (A, B, C), case (i): uses the geophysical predictor, the second row (D, E, F), case (ii): uses the borehole predictor, and the third row (G, H, I), case (iii): uses both predictors. The true point coordinates (Test) are highlighted by the two lines in each dimension.

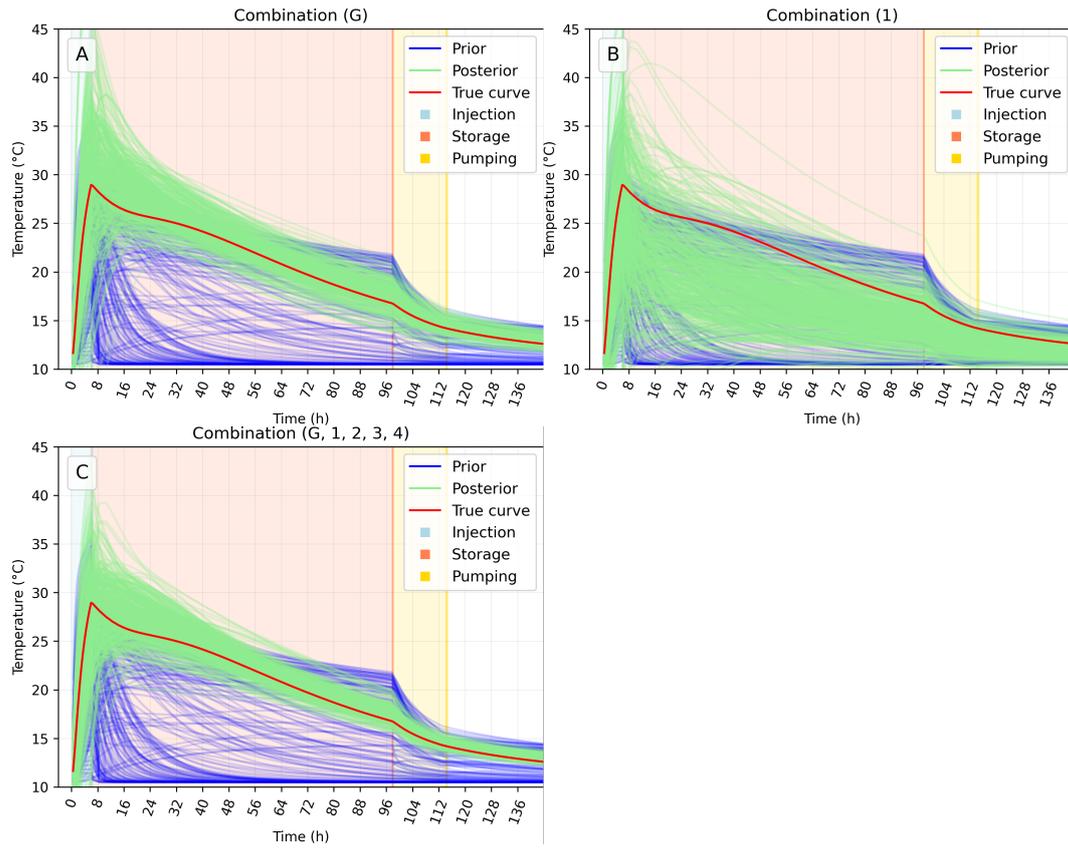


Figure 6. Temperature curves across all time steps, at the observation well 2. A. Using the ERT as predictor. B. Using the temperature curves at well 1 as predictor. C. Using the ERT and all the temperature curves at wells 1, 2, 3 and 4 as predictors.

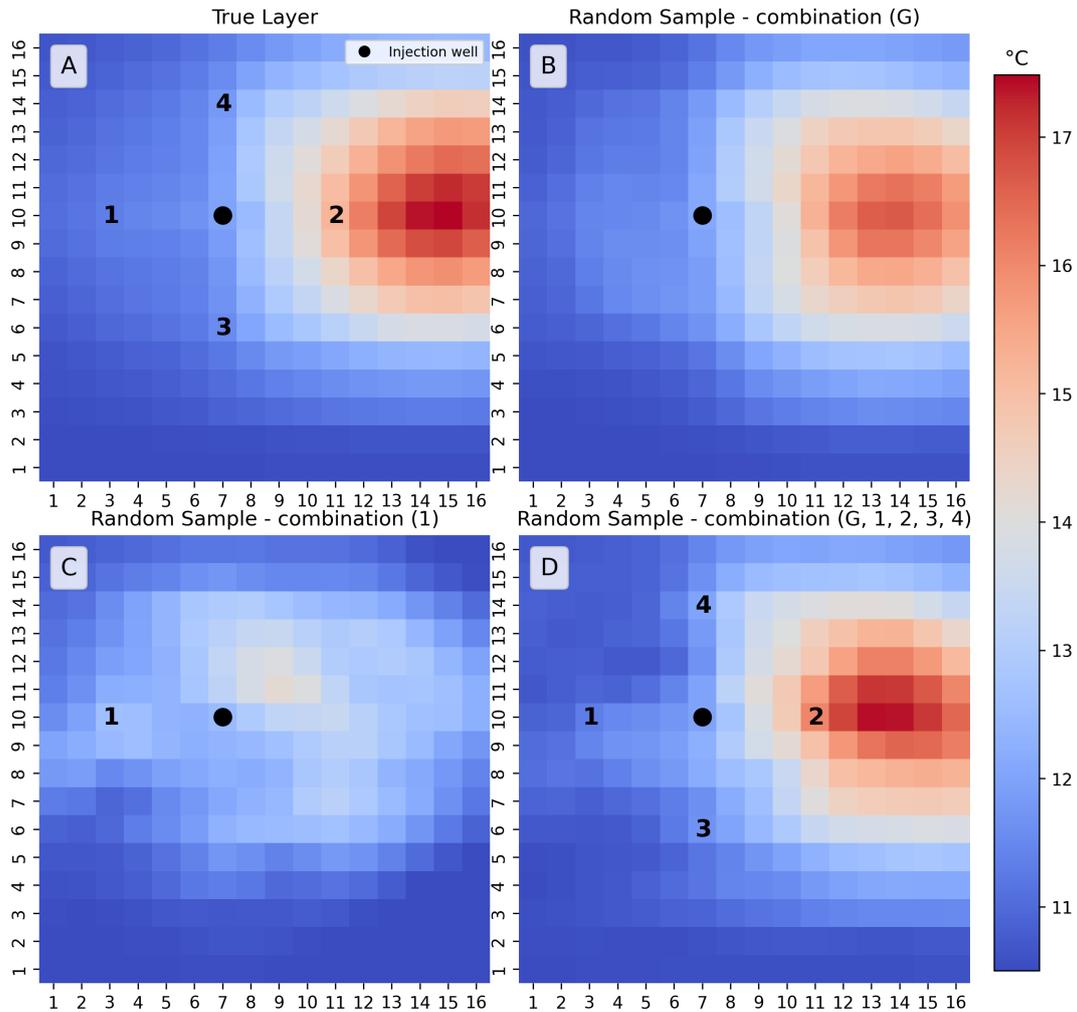


Figure 7. Cross-section of the temperature field at time step 74 (105.75 hours - pumping phase) and layer 9 heat injection level. A. Ground truth. B. Using the ERT as predictor. C. Using the temperature curves at well 1 as predictor. D. Using the ERT and all the temperature curves at wells 1, 2, 3 and 4 as predictors.

543 4 Results

544 To compare the different methods fairly, the same number of components must be
 545 used in each case. Random components from the training set are chosen for BOED, as
 546 shown in Figure 4. The extra random components, which were not used for training, could
 547 thus not be predicted, but they must be included to avoid bias in the results. Selecting
 548 them at random for each PC dimension ensures that the score of the combination will
 549 be representative of the actual prediction performance. The expected outcome is a ranking
 550 of combinations ranging from a small number of combinations to all considered data
 551 sources, with the most unfavourable case being the use of a single observation well and
 552 the most favourable case being the use of geophysical data and all available wells, as can
 553 be visually evaluated across cases on Figure 6.

554 To derive a robust BOED, we average the results over the test set of ground truth
 555 models. However, the results depend on the chosen ground truth models set, so a combination
 556 that is ranked as the best in one case may not be ranked as such for another

557 set, which is why we use a 5-fold cross-validation to validate our findings. We use k-fold
558 cross-validation to produce predictions over 5 different training and test sets. We aver-
559 age the rankings across folds to obtain the final ranking. Figure 8A depicts the well lo-
560 cations for the default case. Figure 8B depicts the average results of all 31 combinations
561 over the 5 folds for the default case. The geophysical data is labeled as ‘G’ and the well
562 data are labeled by their well ID (1, 2, 3, 4). For example, ‘G12’ is the combination of
563 geophysical data with wells 1 and 2. The results are consistent with what was expected,
564 and using more data sources yields the best results, which is logical given the symme-
565 try of our observation wells network. Our findings show that the ERT data alone pro-
566 vides the most information to the model, with a clear difference in metric value between
567 the wells alone and the ERT data. To highlight that there is no overlap between the two,
568 a darker background colour indicates when the ERT data is used, and a lighter background
569 colour indicates when the wells alone are used. Our results corroborate previous find-
570 ings that ERT is a valuable tool for monitoring the development of thermally affected
571 zones in aquifers (Lesparre et al., 2019; Hermans et al., 2016).

572 In terms of observation wells, wells 1 and 4 are consistently ranked as the least fa-
573 vorable cases, and well 2 is ranked as the most favorable case. This is due to the direc-
574 tion of the heat plume over time. As shown on Figures 3, 6 and 7, the pumping causes
575 the plume to move downstream, allowing well 2 to record more variation from the plume
576 over time. The aquifer’s natural gradient has an X component (see Figure 1A), which
577 causes the plume to slowly drift towards well 3, which is why well 3 is ranked in the mid-
578 dle. However, because the plume is moving away from them, wells 1 and 4 are more likely
579 to record redundant data from it (e.g., a flat temperature curve), which makes them less
580 relevant for model training.

581 Similar observations were made in the case of solute transport in Thibaut et al. (2021):
582 downstream wells provided the most useful data. In their settings, their tracer curves
583 provided more information on their underlying hydraulic conductivity field, lowering their
584 prediction’s uncertainty. They did not investigate the information gained from any geo-
585 physical data, which in our case is sufficient to understand the evolution of the heat plume,
586 but it is important to note that the overall information gain from the combination of all
587 wells and ERT is the highest. Since ERT is an indirect data source, adding at least one
588 borehole allows to provide direct information on the temperature and thus to reduce the
589 uncertainty.

590 To corroborate our findings, we reproduce the experiment using four different well
591 locations, but with the same data (the previous k-fold shuffle seed is reused). The al-
592 ternative well positions are shown in Figure 8C and the results are shown in Figure 8D.
593 The outcomes are in line with expectations. The worst-case scenario is well number 3.
594 It is located upstream and is less affected by the heat plume. Because of its proximity
595 to the heat injection, well number 1 is the best case. The combination of the various wells
596 produces more insightful results than before. Wells 1 and 2 cover both upstream and down-
597 stream areas, and are far apart. As a group, they collect more data from the heat plume
598 than the other group of two observation wells. Following the same logic, the combina-
599 tion of wells 1, 2, and 4 (‘124’) is the best case. Because alternative well locations are
600 more dispersed and provide less redundant information, the increase in information pro-
601 vided by geophysical data is not as strong as in the previous case. In the case of alter-
602 native well locations, the full well combination ‘1234’ is not the best case (Figure 8D).
603 As described in section 3, feeding to the training algorithm an array of concatenated data
604 from various sources, some of which provide redundant information, can reduce the model’s
605 prediction power.

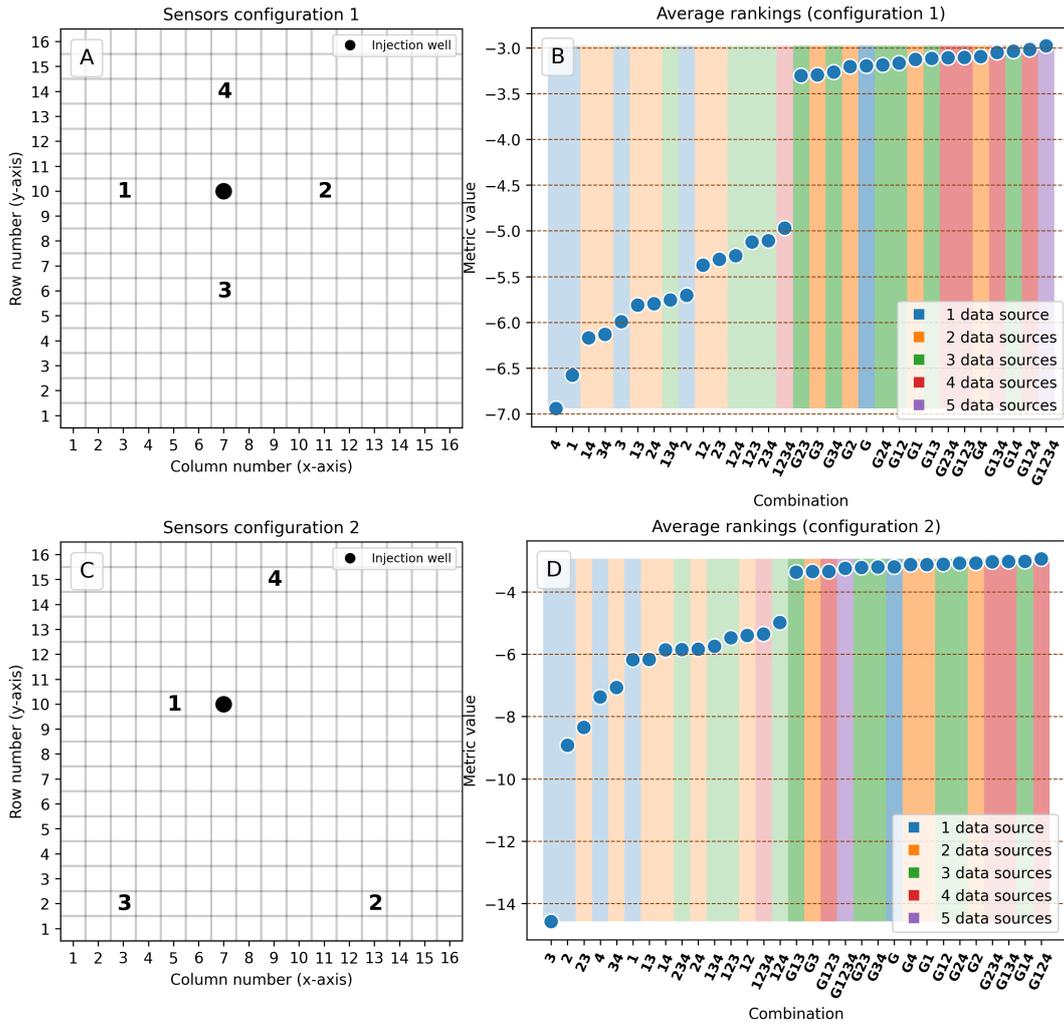


Figure 8. Sensors position for the two cases and ranking of the different combinations of data sources. To visualize a higher score for the best combination, the metric (RMSE) values opposite are displayed. The use of ERT data (**G**) is indicated by a darker background shade, whereas the use of wells alone is indicated by a lighter shade. A. Default well locations. B. Average ranking of 5 folds for the default well locations. C. Alternative well locations. D. Average ranking of 5 folds for the alternative well locations.

5 Discussion

A fundamental question is how we will put our framework into action in order to make the best decisions. Testing for all well positions and combinations is possible in our synthetic study, but time-consuming, because each combination requires a training and testing phase. As a reminder, the number of possible combinations is $\frac{n!}{w!(n-w)!}$, where n is the number of possible positions and w is the number of wells. For a number of 3 wells and 256 possible positions (number of cells for one plane in the grid), the number of possible combinations is 2,763,520. This is assuming uniform grid cells and no hydrogeological property scaling issues. Indeed, petrophysical laws may not be viable at different scales (see Singha et al. (2015) for a discussion on this topic). However, regardless of which combination is selected, the computations will be based on the same 250 heat flow and transport simulations. This ensures that the most time-consuming com-

ponent will not be performed more than once. As a result, running the experimental design for a representative subset of these combinations is feasible. Constraints such as at least 5 m between wells, at least one well upstream and one downstream, and so on can be used to reduce the number of possible combinations.

One ideal situation in which to apply our framework is when potential well locations are predetermined and limited by factors such as land occupation. The situation where the wells are already positioned, and we are looking for the single next-best well position is an alternative ideal situation. This topic, however, is beyond the scope of this contribution. This would entail sequentially updating our prior knowledge from the existing well(s), as each new measurement would necessitate rerunning the forward model with updated parameters. This would be a challenging task given our computationally expensive forward model. Future research will investigate the ability of the framework to sequential BOED, and solve the problem of allocating new sensors placement adaptively such as, e.g., Haan et al. (2021), Lykkegaard and Dodwell (2022), Wang et al. (2022).

In addition to temperature monitoring and borehole sampling optimization, the method as it is presented here can be used to solve a wide range of data sources fusion and location problems with any utility function. Haan et al. (2021) mention that solving such allocation problems helps answer questions such as: Which additional sensor type would yield the highest gain? What is the most effective data source combination? How many and where should sensors be activated given logistical or energy constraints?

In the field of Earth Sciences, Eidsvik et al. (2015) integrates these concepts into a decision-theoretic framework in which the utility function is defined as Value of Information (VOI), which is the expected utility of the information gained from the experiment. It encompasses the notions that the information gained must be (i) informative, (ii) result in material (practical) decisions, and ultimately (iii), lead to economic decisions. These notions are sequentially conditional, such that, in order to make an economic decision, the information must first lead to a material decision, which depends on the information gained. The idea of VOI was applied and developed in works such as Liu et al. (2012) to the remediation of groundwater sites, Trainor-Guitton et al. (2013) to monitor the detection of CO₂ leaks, Nenna and Knight (2014) to evaluate the benefits of acquiring geophysical data as part of a groundwater management strategy, and a geophysical perspective on VOI is presented by Trainor-Guitton (2014). In addition, a new concept called efficacy of information (EOI), which is similar to VOI but without financial rewards or costs, was recently introduced by Caers et al. (2022). Our approach does not take into account the financial or logistical costs of the experiment, but instead focuses on the information gain from the experiment; however, decision-theoretic metrics can easily be incorporated into our framework into a VOI or EOI framework, because information gain is the first stage of the VOI framework.

Several other dynamic processes can be monitored using time-lapse electrical techniques (Singha et al., 2015; Hermans et al., 2018), and the methodology could be applied to cases such as the monitoring of a contaminant plume (Power et al., 2014; Robinson et al., 2020; Tso et al., 2020; Nazifi et al., 2022) or saltwater intrusion (SWI) (Johnson et al., 2015; Palacios et al., 2020). For example, to assess the efficiency of a remediation strategy for a contaminated aquifer, the position of a well can be optimized to minimize the cost of drilling, while still fulfilling the desired accuracy in the prediction of the contaminant concentration. In SWI-prone areas, it could also be used to minimize the uncertainty in the most vulnerable areas in the prediction of the saltwater front by optimizing the number and positions of measurements. Other geophysical monitoring applications could also benefit from the methodology. Wilkinson et al. (2015) review some examples of geophysical monitoring applications. Newer developments include applications such as CO₂ sequestration (Auken et al., 2014; Lu et al., 2015; Pezard et al., 2016; Schmidt-Hattenberger et al., 2016), soil properties for agriculture applications (Blanchy

670 et al., 2020), or even erosion (Masi et al., 2020) and ground-loosening processes (Kim
671 et al., 2022).

672 We also point out that the approach can be used to combine datasets of any type
673 and any dimension in order to stochastically solve geophysical inverse problems. Although
674 different geophysical methods are complementary, JafarGandomi and Binley (2013) ar-
675 gue that inconsistency in their acquisition geometries (i.e., 1D, 2D, or 3D) makes the data
676 fusion process difficult. In their inversion strategy, they re-casted the subsurface model
677 to a set of locally layered 1D earth-models to overcome the incompatibility of the acqui-
678 sition geometries and also to reduce the computational cost of the 2D and 3D inversions.
679 Our fusion method easily combines datasets of any dimension because it first reduces their
680 dimensionality before concatenating them and feeding them to the learning algorithm.

681 **6 Conclusion**

682 In this paper, we present a method for optimizing 4D temperature field monitor-
683 ing experiments using a Bayesian approach. The proposed methodology uses a combi-
684 nation of ODE and Bayesian inference to identify informative observation well locations.
685 We apply our method to a synthetic case study involving the prediction of a four-dimensional
686 temperature field from data collected by electrical resistivity tomography (ERT) and four
687 observation wells recording the temperature over time. These predictors of different nature
688 are combined in the principal component space to form a new predictor. Using our
689 method, the optimal ERT electrode configuration is determined prior to optimization,
690 and we found that, in our case, the multiple-gradient array outperforms the dipole-dipole
691 array in terms of information gain. Following the training step, targets are sampled from
692 the inferred posterior distribution in a low-dimensional latent space using transport map
693 methods, which are a powerful tool in our Bayesian inference framework for sampling
694 an unknown target given a known predictor. Applying a simple metric (RMSE) to the
695 principal components of the predicted and true targets allows to determine the locations
696 of observation wells that minimize uncertainty. This method can be used to optimize the
697 design of 4D temperature field monitoring experiments and to reduce the cost of data
698 collection by choosing a threshold between precision and number and nature of the data.

699 Our findings indicate that the placement of observation wells must take into ac-
700 count the direction of the heat plume, and that an observation network that includes
701 wells both downstream and upstream of the original injection well is optimal. A com-
702 bination of various observation wells and geophysical data always yields the best results.
703 We also noticed that increasing the number of observation wells does not always improve
704 prediction accuracy if the wells are not placed in the proper locations and thus provide
705 redundant or no information. Careful consideration of the information provided by the
706 different data types and how they should be combined is necessary prior to optimiza-
707 tion. Our method provides a straightforward procedure for examining these trade-offs,
708 and is thus a useful technique for determining the optimal sensor combination.

709 One of our framework’s limitations is that it requires prior knowledge of the un-
710 known model parameters as well as knowledge of the physics of the problem to allow for-
711 ward modeling, which can be computationally expensive. However, with a small num-
712 ber of examples (200 for training and 50 for testing), the method is able to predict the
713 posterior distribution of the temperature field with a reasonable accuracy and to find
714 the best combination of data sources among 4 prescribed well positions and their com-
715 binations, with the caveat that the training set should be as representative as possible
716 of the real data. As a more general drawback, OED can not guarantee assumptions such
717 as experiment effectiveness, model form validity, and criterion reflecting experiment ob-
718 jectives. To satisfy the aforementioned assumptions, OED must carefully consider the
719 experiment (Smucker et al., 2018).

The proposed method is not limited to the monitoring of temperature fields and can be applied to the prediction of any type of high-dimensional target from a fusion of multiple predictors in a wide range of contexts, such as geophysical, environmental, and engineering applications.

7 Open Research

The software developments related to this paper are distributed in the software SKBEL (v2.0.0) published on GitHub under the BSD 3-Clause License (Thibaut & Ramgraber, 2021). Additionally, the dataset is available on Kaggle (Lesparre et al., 2022) under the terms of the CC BY-SA 4.0 License. It includes the data used in the paper (Temperature field monitoring with ERT data) and Jupyter notebooks to reproduce the results.

References

- Arato, A., Boaga, J., Comina, C., Seta, M. D., Sipio, E. D., Galgaro, A., . . . Mandrone, G. (2015, 8). Geophysical monitoring for shallow geothermal applications – two Italian case histories. *First Break*, *33*. Retrieved from <https://www.earthdoc.org/content/journals/0.3997/1365-2397.33.8.82010> doi: 10.3997/1365-2397.33.8.82010
- Athens, N. D., & Caers, J. K. (2019, 12). A monte carlo-based framework for assessing the value of information and development risk in geothermal exploration. *Applied Energy*, *256*, 113932. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S0306261919316198> doi: 10.1016/j.apenergy.2019.113932
- Attia, A., Alexanderian, A., & Saibaba, A. K. (2018, 9). Goal-oriented optimal design of experiments for large-scale Bayesian linear inverse problems. *Inverse Problems*, *34*, 095009. doi: 10.1088/1361-6420/aad210
- Auken, E., Doetsch, J., Fiandaca, G., Christiansen, A. V., Gazoty, A., Cahill, A. G., & Jakobsen, R. (2014, 2). Imaging subsurface migration of dissolved CO₂ in a shallow aquifer using 3-d time-lapse electrical resistivity tomography. *Journal of Applied Geophysics*, *101*, 31-41. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S0926985113002681> doi: 10.1016/j.jappgeo.2013.11.011
- Baptista, R., Marzouk, Y., & Zahm, O. (2022). On the representation and learning of monotone triangular transport maps. *arXiv*. Retrieved from <https://arxiv.org/abs/2009.10303v2> doi: 10.48550/ARXIV.2009.10303
- Bayer, P., Rybach, L., Blum, P., & Brauchler, R. (2013, 10). Review on life cycle environmental effects of geothermal power generation. *Renewable and Sustainable Energy Reviews*, *26*, 446-463. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S1364032113003420> doi: 10.1016/j.rser.2013.05.039
- Bellman, R. (1961). *Adaptive control processes: A guided tour. (a random corporation research study)*. Princeton, N. J.: Princeton University Press, XVI, 255 p. (1961).
- Blanchy, G., Watts, C. W., Richards, J., Bussell, J., Huntentburg, K., Sparkes, D. L., . . . Binley, A. (2020, 1). Time-lapse geophysical assessment of agricultural practices on soil moisture dynamics. *Vadose Zone Journal*, *19*. Retrieved from <https://onlinelibrary.wiley.com/doi/10.1002/vzj2.20080> doi: 10.1002/vzj2.20080
- Bridger, D. W., & Allen, D. M. (2010, 1). Heat transport simulations in a heterogeneous aquifer used for aquifer thermal energy storage (ATES). *Canadian Geotechnical Journal*, *47*, 96-115. Retrieved from <http://www.nrcresearchpress.com/doi/10.1139/T09-078> doi: 10.1139/T09-078
- Brouyère, S. (2001). *Etude et modélisation du transport et du piégeage des solutés en milieu souterrain variablement saturé (study and modelling of transport and*

- 771 *retardation of solutes in variably saturated media*). (PhD thesis ed.). Retrieved
 772 from <https://hdl.handle.net/2268/40804>
- 773 Brunner, P., & Simmons, C. T. (2012, 3). Hydrogeosphere: A fully integrated, phys-
 774 ically based hydrological model. *Ground Water*, *50*, 170-176. doi: 10.1111/j
 775 .1745-6584.2011.00882.x
- 776 Caers, J., Scheidt, C., Yin, Z., Wang, L., Mukerji, T., & House, K. (2022, 6). Ef-
 777 ficacy of information in mineral exploration drilling. *Natural Resources Re-*
 778 *search*, *31*, 1157-1173. doi: 10.1007/s11053-022-10030-1
- 779 Chaloner, K., & Verdinelli, I. (1995, 8). Bayesian experimental design: A review.
 780 *Statistical Science*, *10*. doi: 10.1214/ss/1177009939
- 781 Dahlin, T., & Zhou, B. (2006). Multiple-gradient array measurements for multichan-
 782 nel 2d resistivity imaging. *Near Surface Geophysics*, *4*, 113-123. doi: 10.3997/
 783 1873-0604.2005037
- 784 Dassargues, A. (1997, 3). Modeling baseflow from an alluvial aquifer using
 785 hydraulic-conductivity data obtained from a derived relation with apparent
 786 electrical resistivity. *Hydrogeology Journal*, *5*, 97-108. Retrieved from [http://](http://link.springer.com/10.1007/s100400050125)
 787 link.springer.com/10.1007/s100400050125 doi: 10.1007/s100400050125
- 788 Derouane, J., & Dassargues, A. (1998, 11). Delineation of groundwater protection
 789 zones based on tracer tests and transport modeling in alluvial sediments. *Envi-*
 790 *ronmental Geology*, *36*, 27-36. Retrieved from [http://link.springer.com/10](http://link.springer.com/10.1007/s002540050317)
 791 [.1007/s002540050317](http://link.springer.com/10.1007/s002540050317) doi: 10.1007/s002540050317
- 792 Duijff, R., Bloemendal, M., & Bakker, M. (2021, 12). Interaction effects between
 793 aquifer thermal energy storage systems. *Groundwater*, gwat.13163. Retrieved
 794 from <https://onlinelibrary.wiley.com/doi/10.1111/gwat.13163> doi: 10
 795 .1111/gwat.13163
- 796 Eidsvik, J., Mukerji, T., & Bhattacharjya, D. (2015). *Value of information in*
 797 *the earth sciences*. Cambridge University Press. Retrieved from [https://](https://www.cambridge.org/core/product/identifier/9781139628785/type/book)
 798 www.cambridge.org/core/product/identifier/9781139628785/type/book
 799 doi: 10.1017/CBO9781139628785
- 800 El Moselhy, T. A., & Marzouk, Y. M. (2012, October). Bayesian inference with
 801 optimal maps. *Journal of Computational Physics*, *231*(23), 7815–7850. Re-
 802 trieved 2022-03-01, from [https://linkinghub.elsevier.com/retrieve/pii/](https://linkinghub.elsevier.com/retrieve/pii/S0021999112003956)
 803 [S0021999112003956](https://linkinghub.elsevier.com/retrieve/pii/S0021999112003956) doi: 10.1016/j.jcp.2012.07.022
- 804 Ferguson, G. (2007, 7). Heterogeneity and thermal modeling of ground water.
 805 *Ground Water*, *45*, 485-490. Retrieved from [https://onlinelibrary](https://onlinelibrary.wiley.com/doi/10.1111/j.1745-6584.2007.00323.x)
 806 [.wiley.com/doi/10.1111/j.1745-6584.2007.00323.x](https://onlinelibrary.wiley.com/doi/10.1111/j.1745-6584.2007.00323.x) doi: 10.1111/
 807 [j.1745-6584.2007.00323.x](https://onlinelibrary.wiley.com/doi/10.1111/j.1745-6584.2007.00323.x)
- 808 Goovaerts, P. (1997). *Geostatistics for natural resources evaluation*. New York: Ox-
 809 ford University Press.
- 810 Haan, S., Ramos, F., & Müller, R. D. (2021, 1). Multiobjective bayesian optimiza-
 811 tion and joint inversion for active sensor fusion. *Geophysics*, *86*, ID1-ID17. Re-
 812 trieved from <https://library.seg.org/doi/10.1190/geo2019-0460.1> doi:
 813 10.1190/geo2019-0460.1
- 814 Hermans, T. (2017, 1). Prediction-focused approaches: An opportunity for hydrology.
 815 *Groundwater*, *55*, 683-687. doi: 10.1111/gwat.12548
- 816 Hermans, T., Lesparre, N., De Schepper, G., & Robert, T. (2019, August). Bayesian
 817 evidential learning: a field validation using push-pull tests. *Hydrogeology Jour-*
 818 *nal*, *27*(5), 1661–1672. Retrieved 2020-11-21, from [http://link.springer](http://link.springer.com/10.1007/s10040-019-01962-9)
 819 [.com/10.1007/s10040-019-01962-9](http://link.springer.com/10.1007/s10040-019-01962-9) doi: 10.1007/s10040-019-01962-9
- 820 Hermans, T., Nguyen, F., Klepikova, M., Dassargues, A., & Caers, J. (2018, 1).
 821 Uncertainty quantification of medium-term heat storage from short-term geo-
 822 physical experiments using bayesian evidential learning. *Water Resources*
 823 *Research*, *54*, 2931-2948. doi: 10.1002/2017WR022135
- 824 Hermans, T., Nguyen, F., Robert, T., & Revil, A. (2014, 8). Geophysical meth-
 825 ods for monitoring temperature changes in shallow low enthalpy geothermal

- 826 systems. *Energies*, 7, 5083-5118. Retrieved from [http://www.mdpi.com/](http://www.mdpi.com/1996-1073/7/8/5083)
 827 1996-1073/7/8/5083 doi: 10.3390/en7085083
- 828 Hermans, T., Oware, E., & Caers, J. (2016, 9). Direct prediction of spatially and
 829 temporally varying physical properties from time-lapse electrical resistance
 830 data. *Water Resources Research*, 52, 7262-7283. Retrieved from [http://](http://doi.wiley.com/10.1002/2016WR019126)
 831 doi.wiley.com/10.1002/2016WR019126 doi: 10.1002/2016WR019126
- 832 Hermans, T., Vandenbohede, A., Lebbe, L., & Nguyen, F. (2012, 1). A
 833 shallow geothermal experiment in a sandy aquifer monitored using elec-
 834 tric resistivity tomography. *GEOPHYSICS*, 77, B11-B21. Retrieved
 835 from <https://library.seg.org/doi/10.1190/geo2011-0199.1> doi:
 836 10.1190/geo2011-0199.1
- 837 Hermans, T., Wildemeersch, S., Jamin, P., Orban, P., Brouyère, S., Dassargues,
 838 A., & Nguyen, F. (2015). Quantitative temperature monitoring of a heat
 839 tracing experiment using cross-borehole ert. *Geothermics*, 53, 14-26. doi:
 840 10.1016/j.geothermics.2014.03.013
- 841 JafarGandomi, A., & Binley, A. (2013, 9). A bayesian trans-dimensional approach
 842 for the fusion of multiple geophysical datasets. *Journal of Applied Geophysics*,
 843 96, 38-54. doi: 10.1016/j.jappgeo.2013.06.004
- 844 Johnson, T., Versteeg, R., Thomle, J., Hammond, G., Chen, X., & Zachara, J.
 845 (2015). Four-dimensional electrical conductivity monitoring of stage-driven
 846 river water intrusion: Accounting for water table effects using a transient mesh
 847 boundary and conditional inversion constraints. *Water Resources Research*,
 848 51, 6177-6196. doi: 10.1002/2014WR016129
- 849 Kim, B., Joung, I., Cho, A., Shin, D., Han, Y., & Nam, M. (2022). Monitor-
 850 ing the perturbation zone near a foundation excavation with electrical re-
 851 sistivity tomography: Comparison between time-lapse 3d and 2d inver-
 852 sions in single-profile study. *Journal of Applied Geophysics*, 205. doi:
 853 10.1016/j.jappgeo.2022.104772
- 854 Kleinegesse, S., & Gutmann, M. U. (2020, 13–18 Jul). Bayesian experimental
 855 design for implicit models by mutual information neural estimation. In
 856 H. D. III & A. Singh (Eds.), *Proceedings of the 37th international confer-*
 857 *ence on machine learning* (Vol. 119, pp. 5316–5326). PMLR. Retrieved from
 858 <https://proceedings.mlr.press/v119/kleinegesse20a.html>
- 859 Klepikova, M. V., Borgne, T. L., Bour, O., Dentz, M., Hochreutener, R., &
 860 Lavenant, N. (2016, 7). Heat as a tracer for understanding transport
 861 processes in fractured media: Theory and field assessment from multiscale
 862 thermal push-pull tracer tests. *Water Resources Research*, 52, 5442-5457.
 863 Retrieved from <http://doi.wiley.com/10.1002/2016WR018789> doi:
 864 10.1002/2016WR018789
- 865 Laloy, E., & Vrugt, J. A. (2012, jan). High-dimensional posterior exploration of
 866 hydrologic models using multiple-try DREAM (ZS) and high-performance
 867 computing. *Water Resources Research*, 48(1). Retrieved from [http://](http://doi.wiley.com/10.1029/2011WR010608)
 868 doi.wiley.com/10.1029/2011WR010608 doi: 10.1029/2011WR010608
- 869 Lesparre, N., Compaire, N., Hermans, T., & Thibaut, R. (2022). *4D Temperature*
 870 *Monitoring* [Dataset]. Kaggle. Retrieved from [https://www.kaggle.com/dsv/](https://www.kaggle.com/dsv/3819983)
 871 [3819983](https://www.kaggle.com/dsv/3819983) doi: 10.34740/KAGGLE/DSV/3819983
- 872 Lesparre, N., Robert, T., Nguyen, F., Boyle, A., & Hermans, T. (2019, 1). 4d
 873 electrical resistivity tomography (ert) for aquifer thermal energy stor-
 874 age monitoring. *Geothermics*, 77, 368-382. Retrieved from [https://](https://linkinghub.elsevier.com/retrieve/pii/S03755650517303772)
 875 linkinghub.elsevier.com/retrieve/pii/S03755650517303772 doi:
 876 10.1016/j.geothermics.2018.10.011
- 877 Lindley, D. V. (1956, 12). On a measure of the information provided by an exper-
 878 iment. *The Annals of Mathematical Statistics*, 27, 986-1005. doi: 10.1214/
 879 aoms/1177728069
- 880 Liu, X., Lee, J., Kitanidis, P. K., Parker, J., & Kim, U. (2012, 4). Value of

- 881 information as a context-specific measure of uncertainty in groundwa-
 882 ter remediation. *Water Resources Management*, 26, 1513-1535. doi:
 883 10.1007/s11269-011-9970-3
- 884 Lu, C., Zhang, C., Hunag, H., & Johnson, T. C. (2015, 2). Monitoring co-
 885 subsurface sequestration into deep saline aquifer and associated salt in-
 886 trusion using coupled multiphase flow modeling and time-lapse electrical re-
 887 sistivity tomography. *Greenhouse Gases: Science and Technology*, 5, 34-49.
 888 Retrieved from <https://onlinelibrary.wiley.com/doi/10.1002/ghg.1437>
 889 doi: 10.1002/ghg.1437
- 890 Lykkegaard, M. B., & Dodwell, T. J. (2022, 6). Where to drill next? a dual-
 891 weighted approach to adaptive optimal design of groundwater surveys. *Ad-
 892 vances in Water Resources*, 164, 104219. doi: 10.1016/j.advwatres.2022.
 893 .104219
- 894 Macfarlane, A., Förster, A., Merriam, D., Schrötter, J., & Healey, J. (2002, 12).
 895 Monitoring artificially stimulated fluid movement in the cretaceous dakota
 896 aquifer, western kansas. *Hydrogeology Journal*, 10, 662-673. Retrieved
 897 from <http://link.springer.com/10.1007/s10040-002-0223-7> doi:
 898 10.1007/s10040-002-0223-7
- 899 Masi, M., Ferdos, F., Losito, G., & Solari, L. (2020). Monitoring of internal erosion
 900 processes by time-lapse electrical resistivity tomography. *Journal of Hydrology*,
 901 589. doi: 10.1016/j.jhydrol.2020.125340
- 902 Meloun, M., & Militký, J. (Eds.). (2012). *Statistical data analysis: a practical guide*
 903 (Reprinted ed.). New Delhi: WPI, Woodhead Publ. India Pvt. Ltd. (OCLC:
 904 696087077)
- 905 Michel, H., Hermans, T., & Nguyen, F. (2022, 9). Iterative prior resampling and
 906 rejection sampling to improve 1d geophysical imaging based on bayesian
 907 evidential learning (bell1d). *Geophysical Journal International*. doi:
 908 10.1093/gji/ggac372
- 909 Michel, H., Nguyen, F., Kremer, T., Elen, A., & Hermans, T. (2020, 5). 1d ge-
 910 ological imaging of the subsurface from geophysical data with bayesian ev-
 911 idential learning. *Computers & Geosciences*, 138, 104456. Retrieved from
 912 <https://linkinghub.elsevier.com/retrieve/pii/S0098300419306028>
 913 doi: 10.1016/j.cageo.2020.104456
- 914 Nazifi, H. M., Gülen, L., Gürbüz, E., & Pekşen, E. (2022). Time-lapse elec-
 915 trical resistivity tomography (ert) monitoring of used engine oil contam-
 916 ination in laboratory setting. *Journal of Applied Geophysics*, 197. doi:
 917 10.1016/j.jappgeo.2022.104531
- 918 Nenna, V., & Knight, R. (2014, 1). Demonstration of a value of information met-
 919 ric to assess the use of geophysical data for a groundwater application. *GEO-
 920 PHYSICS*, 79, E51-E60. doi: 10.1190/geo2012-0474.1
- 921 Palacios, A., Ledo, J. J., Linde, N., Luquot, L., Bellmunt, F., Folch, A., ... Carrera,
 922 J. (2020). Time-lapse cross-hole electrical resistivity tomography (chert) for
 923 monitoring seawater intrusion dynamics in a mediterranean aquifer. *Hydrology
 924 and Earth System Sciences*, 24, 2121-2139. doi: 10.5194/hess-24-2121-2020
- 925 Palmer, C. D., Blowes, D. W., Frind, E. O., & Molson, J. W. (1992, 10). Thermal
 926 energy storage in an unconfined aquifer: 1. field injection experiment. *Water
 927 Resources Research*, 28, 2845-2856. Retrieved from [http://doi.wiley.com/10.
 928 .1029/92WR01471](http://doi.wiley.com/10.1029/92WR01471) doi: 10.1029/92WR01471
- 929 Park, B.-H., Bae, G.-O., & Lee, K.-K. (2015, 11). Importance of thermal dis-
 930 persivity in designing groundwater heat pump ({GWHP}) system: {Field}
 931 and numerical study. *Renewable Energy*, 83, 270-279. Retrieved from
 932 <https://linkinghub.elsevier.com/retrieve/pii/S0960148115500316X>
 933 doi: 10.1016/j.renene.2015.04.036
- 934 Park, J., & Caers, J. (2020, October). Direct forecasting of global and spatial model
 935 parameters from dynamic data. *Computers & Geosciences*, 143, 104567. Re-

- 936 trieved 2021-01-30, from [https://linkinghub.elsevier.com/retrieve/pii/](https://linkinghub.elsevier.com/retrieve/pii/S0098300420305562)
937 S0098300420305562 doi: 10.1016/j.cageo.2020.104567
- 938 Pezard, P. A., Denchik, N., Lofi, J., Perroud, H., Henry, G., Neyens, D., ... Lev-
939 annier, A. (2016, 5). Time-lapse downhole electrical resistivity monitoring of
940 subsurface co₂ storage at the maguelone shallow experimental site (languedoc,
941 france). *International Journal of Greenhouse Gas Control*, *48*, 142-
942 154. Retrieved from [https://linkinghub.elsevier.com/retrieve/pii/](https://linkinghub.elsevier.com/retrieve/pii/S1750583615301523)
943 S1750583615301523 doi: 10.1016/j.ijggc.2015.12.005
- 944 Polydorides, N., & Lionheart, W. R. B. (2002, 12). A matlab toolkit for three-
945 dimensional electrical impedance tomography: a contribution to the electrical
946 impedance and diffuse optical reconstruction software project. *Measurement*
947 *Science and Technology*, *13*, 1871-1883. doi: 10.1088/0957-0233/13/12/310
- 948 Power, C., Gerhard, J., Karaoulis, M., Tsourlos, P., & Giannopoulos, A. (2014).
949 Evaluating four-dimensional time-lapse electrical resistivity tomography for
950 monitoring dnapi source zone remediation. *Journal of Contaminant Hydrology*,
951 *162-163*, 27-46. doi: 10.1016/j.jconhyd.2014.04.004
- 952 Pradhan, A., & Mukerji, T. (2020, 6). Seismic bayesian evidential learning: estima-
953 tion and uncertainty quantification of sub-resolution reservoir properties. *Com-
954 putational Geosciences*, *24*, 1121-1140. Retrieved from [http://link.springer](http://link.springer.com/10.1007/s10596-019-09929-1)
955 .com/10.1007/s10596-019-09929-1 doi: 10.1007/s10596-019-09929-1
- 956 Qiang, S., Shi, X., Kang, X., & Revil, A. (2022, 7). Optimized arrays for electri-
957 cal resistivity tomography survey using bayesian experimental design. *GEO-
958 PHYSICS*, *87*, E189-E203. doi: 10.1190/geo2021-0408.1
- 959 Robert, Paulus, Bolly, Lin, K. S., & Hermans. (2019, 9). Heat as a proxy to im-
960 age dynamic processes with 4d electrical resistivity tomography. *Geosciences*,
961 *9*, 414. Retrieved from <https://www.mdpi.com/2076-3263/9/10/414> doi: 10
962 .3390/geosciences9100414
- 963 Robinson, J., Johnson, T., & Rockhold, M. (2020). Feasibility assessment of long-
964 term electrical resistivity monitoring of a nitrate plume. *Groundwater*, *58*,
965 224-237. doi: 10.1111/gwat.12899
- 966 Ryan, E. G., Drovandi, C. C., McGree, J. M., & Pettitt, A. N. (2016, 4). A review
967 of modern computational algorithms for bayesian optimal design. *International*
968 *Statistical Review*, *84*, 128-154. doi: 10.1111/insr.12107
- 969 Saner, D., Juraske, R., Kübert, M., Blum, P., Hellweg, S., & Bayer, P. (2010,
970 9). Is it only co₂ that matters? a life cycle perspective on shallow geother-
971 mal systems. *Renewable and Sustainable Energy Reviews*, *14*, 1798-1813.
972 Retrieved from [https://linkinghub.elsevier.com/retrieve/pii/](https://linkinghub.elsevier.com/retrieve/pii/S1364032110001061)
973 S1364032110001061 doi: 10.1016/j.rser.2010.04.002
- 974 Satija, A., & Caers, J. (2015, March). Direct forecasting of subsurface flow response
975 from non-linear dynamic data by linear least-squares in canonical functional
976 principal component space. *Advances in Water Resources*, *77*, 69-81. Re-
977 trieved 2021-01-30, from [https://linkinghub.elsevier.com/retrieve/pii/](https://linkinghub.elsevier.com/retrieve/pii/S0309170815000044)
978 S0309170815000044 doi: 10.1016/j.advwatres.2015.01.002
- 979 Scheidt, C., Li, L., & Caers, J. (2018). *Quantifying Uncertainty in Subsur-
980 face Systems*. Hoboken, NJ, USA: John Wiley & Sons, Inc. Retrieved
981 2020-09-29, from <http://doi.wiley.com/10.1002/9781119325888> doi:
982 10.1002/9781119325888
- 983 Scheidt, C., Renard, P., & Caers, J. (2015, February). Prediction-Focused Sub-
984 surface Modeling: Investigating the Need for Accuracy in Flow-Based Inverse
985 Modeling. *Mathematical Geosciences*, *47*(2), 173-191. Retrieved 2021-01-
986 30, from <http://link.springer.com/10.1007/s11004-014-9521-6> doi:
987 10.1007/s11004-014-9521-6
- 988 Schmidt-Hattenberger, C., Bergmann, P., Labitzke, T., Wagner, F., & Rippe, D.
989 (2016). Permanent crosshole electrical resistivity tomography (ert) as an es-
990 tablished method for the long-term monitoring of the ketzin pilot

- 991 site. *International Journal of Greenhouse Gas Control*, 52, 432-448. doi:
 992 10.1016/j.ijggc.2016.07.024
- 993 Singha, K., Day-Lewis, F. D., Johnson, T., & Slater, L. D. (2015, 3). Advances in
 994 interpretation of subsurface processes with time-lapse electrical imaging. *Hy-*
 995 *drological Processes*, 29, 1549-1576. doi: 10.1002/hyp.10280
- 996 Smucker, B., Krzywinski, M., & Altman, N. (2018, 8). Optimal experimental design.
 997 *Nature Methods*, 15, 559-560. doi: 10.1038/s41592-018-0083-2
- 998 Sommer, W., Valstar, J., van Gaans, P., Grotenhuis, T., & Rijnaarts, H. (2013,
 999 12). The impact of aquifer heterogeneity on the performance of aquifer
 1000 thermal energy storage. *Water Resources Research*, 49, 8128-8138. Re-
 1001 trieved from <http://doi.wiley.com/10.1002/2013WR013677> doi:
 1002 10.1002/2013WR013677
- 1003 Sommer, W. T., Doornenbal, P. J., Drijver, B. C., van Gaans, P. F. M., Leusbrock,
 1004 I., Grotenhuis, J. T. C., & Rijnaarts, H. H. M. (2014, 2). Thermal perfor-
 1005 mance and heat transport in aquifer thermal energy storage. *Hydrogeology*
 1006 *Journal*, 22, 263-279. Retrieved from <http://link.springer.com/10.1007/s10040-013-1066-0> doi: 10.1007/s10040-013-1066-0
- 1007 Spantini, A., Baptista, R., & Marzouk, Y. (2022). Coupling techniques for nonlin-
 1008 ear ensemble filtering. *SIAM Review*, in press. Retrieved from [https://arxiv](https://arxiv.org/abs/1907.00389v2)
 1009 [.org/abs/1907.00389v2](https://arxiv.org/abs/1907.00389v2)
- 1010 Spantini, A., Bigoni, D., & Marzouk, Y. (2018). Inference via low-dimensional cou-
 1011 plings. *arXiv*. Retrieved from <https://arxiv.org/abs/1703.06131> doi: 10
 1012 .48550/arxiv.1703.06131
- 1013 Tadjer, A., & Bratvold, R. B. (2021). Managing uncertainty in geological co2 storage
 1014 using bayesian evidential learning. *Energies*, 14. doi: 10.3390/en14061557
- 1015 Tarakanov, A., & Elsheikh, A. H. (2020, October). Optimal bayesian experi-
 1016 mental design for subsurface flow problems. *Computer Methods in Ap-*
 1017 *plied Mechanics and Engineering*, 370, 113208. Retrieved 2021-07-14, from
 1018 <https://linkinghub.elsevier.com/retrieve/pii/S0045782520303935>
 1019 doi: 10.1016/j.cma.2020.113208
- 1020 Tarantola, A. (2005). *Inverse problem theory and methods for model paramete-*
 1021 *ter estimation*. Society for Industrial and Applied Mathematics. Retrieved
 1022 from <http://epubs.siam.org/doi/book/10.1137/1.9780898717921> doi:
 1023 10.1137/1.9780898717921
- 1024 Tarantola, A. (2006, 8). Popper, bayes and the inverse problem. *Nature Physics*, 2,
 1025 492-494. doi: 10.1038/nphys375
- 1026 Thibaut, R., Laloy, E., & Hermans, T. (2021, 12). A new framework for ex-
 1027 perimental design using bayesian evidential learning: The case of well-
 1028 head protection area. *Journal of Hydrology*, 603, 126903. Retrieved from
 1029 <https://linkinghub.elsevier.com/retrieve/pii/S0022169421009537>
 1030 doi: 10.1016/j.jhydrol.2021.126903
- 1031 Thibaut, R., & Ramgraber, M. (2021, September). *SKBEL - Bayesian Evi-*
 1032 *dential Learning framework built on top of scikit-learn* [Software]. Zen-
 1033 odo. Retrieved from <https://doi.org/10.5281/zenodo.6205242> doi:
 1034 10.5281/zenodo.6205242
- 1035 Trainor-Guitton, W. J. (2014, 3). A geophysical perspective of value of information:
 1036 examples of spatial decisions for groundwater sustainability. *Environment Sys-*
 1037 *tems and Decisions*, 34, 124-133. doi: 10.1007/s10669-013-9487-9
- 1038 Trainor-Guitton, W. J., Ramirez, A., Yang, X., Mansoor, K., Sun, Y., & Car-
 1039 roll, S. (2013, 10). Value of information methodology for assessing the
 1040 ability of electrical resistivity to detect co2/brine leakage into a shallow
 1041 aquifer. *International Journal of Greenhouse Gas Control*, 18, 101-113. doi:
 1042 10.1016/j.ijggc.2013.06.018
- 1043 Tso, C.-H., Johnson, T., Song, X., Chen, X., Kuras, O., Wilkinson, P., . . . Binley,
 1044 A. (2020). Integrated hydrogeophysical modelling and data assimilation for
 1045

- 1046 geolectrical leak detection. *Journal of Contaminant Hydrology*, 234. doi:
1047 10.1016/j.jconhyd.2020.103679
- 1048 Uhlemann, S., Wilkinson, P. B., Maurer, H., Wagner, F. M., Johnson, T. C., &
1049 Chambers, J. E. (2018, 7). Optimized survey design for electrical resistiv-
1050 ity tomography: combined optimization of measurement configuration and
1051 electrode placement. *Geophysical Journal International*, 214, 108-121. doi:
1052 10.1093/gji/ggy128
- 1053 Vandenbohede, A., Hermans, T., Nguyen, F., & Lebbe, L. (2011, 10). Shal-
1054 low heat injection and storage experiment: Heat transport simulation and
1055 sensitivity analysis. *Journal of Hydrology*, 409, 262-272. Retrieved from
1056 <https://linkinghub.elsevier.com/retrieve/pii/S0022169411005701>
1057 doi: 10.1016/j.jhydrol.2011.08.024
- 1058 Vandenbohede, A., Louwyck, A., & Lebbe, L. (2009, 1). Conservative solute versus
1059 heat transport in porous media during push-pull tests. *Transport in Porous*
1060 *Media*, 76, 265-287. Retrieved from [http://link.springer.com/10.1007/](http://link.springer.com/10.1007/s11242-008-9246-4)
1061 [s11242-008-9246-4](http://link.springer.com/10.1007/s11242-008-9246-4) doi: 10.1007/s11242-008-9246-4
- 1062 Vanhoudt, D., Desmedt, J., Bael, J. V., Robeyn, N., & Hoes, H. (2011, 12). An
1063 aquifer thermal storage system in a belgian hospital: Long-term experimen-
1064 tal evaluation of energy and cost savings. *Energy and Buildings*, 43, 3657-
1065 3665. Retrieved from [https://linkinghub.elsevier.com/retrieve/pii/](https://linkinghub.elsevier.com/retrieve/pii/S0378778811004427)
1066 [S0378778811004427](https://linkinghub.elsevier.com/retrieve/pii/S0378778811004427) doi: 10.1016/j.enbuild.2011.09.040
- 1067 Villani, C. (2009). *Optimal Transport* (Vol. 338). Berlin, Heidelberg: Springer Berlin
1068 Heidelberg. Retrieved 2022-03-01, from [http://link.springer.com/10.1007/](http://link.springer.com/10.1007/978-3-540-71050-9)
1069 [978-3-540-71050-9](http://link.springer.com/10.1007/978-3-540-71050-9) doi: 10.1007/978-3-540-71050-9
- 1070 Vrugt, J. A. (2016, January). Markov chain monte carlo simulation using the
1071 DREAM software package: Theory, concepts, and matlab implementation.
1072 *Environmental Modelling & Software*, 75, 273-316. Retrieved 2021-02-05, from
1073 <https://linkinghub.elsevier.com/retrieve/pii/S1364815215300396>
1074 doi: 10.1016/j.envsoft.2015.08.013
- 1075 Wagner, F. M., Günther, T., Schmidt-Hattenberger, C., & Maurer, H. (2015,
1076 3). Constructive optimization of electrode locations for target-focused
1077 resistivity monitoring. *GEOPHYSICS*, 80, E29-E40. Retrieved from
1078 <https://library.seg.org/doi/10.1190/geo2014-0214.1> doi: 10.1190/
1079 geo2014-0214.1
- 1080 Wagner, V., Li, T., Bayer, P., Leven, C., Dietrich, P., & Blum, P. (2014, 2). Ther-
1081 mal tracer testing in a sedimentary aquifer: field experiment (lauswiesen,
1082 germany) and numerical simulation. *Hydrogeology Journal*, 22, 175-187. Re-
1083 trieved from <http://link.springer.com/10.1007/s10040-013-1059-z> doi:
1084 10.1007/s10040-013-1059-z
- 1085 Wang, Y., Zechner, M., Mern, J. M., Kochenderfer, M. J., & Caers, J. K. (2022, 8).
1086 A sequential decision-making framework with uncertainty quantification for
1087 groundwater management. *Advances in Water Resources*, 166, 104266. doi:
1088 10.1016/j.advwatres.2022.104266
- 1089 Wildemeersch, S., Jamin, P., Orban, P., Hermans, T., Klepikova, M., Nguyen,
1090 F., ... Dassargues, A. (2014, 11). Coupling heat and chemical tracer
1091 experiments for estimating heat transfer parameters in shallow alluvial
1092 aquifers. *Journal of Contaminant Hydrology*, 169, 90-99. Retrieved from
1093 <https://linkinghub.elsevier.com/retrieve/pii/S0169772214001247>
1094 doi: 10.1016/j.jconhyd.2014.08.001
- 1095 Wilkinson, P. B., Uhlemann, S., Meldrum, P. I., Chambers, J. E., Carrière, S.,
1096 Oxby, L. S., & Loke, M. (2015, 10). Adaptive time-lapse optimized survey
1097 design for electrical resistivity tomography monitoring. *Geophysical Journal*
1098 *International*, 203, 755-766. Retrieved from [https://academic.oup.com/gji/](https://academic.oup.com/gji/article-lookup/doi/10.1093/gji/ggv329)
1099 [article-lookup/doi/10.1093/gji/ggv329](https://academic.oup.com/gji/article-lookup/doi/10.1093/gji/ggv329) doi: 10.1093/gji/ggv329
- 1100 Yin, Z., Strebelle, S., & Caers, J. (2020, 2). Automated monte carlo-based quan-

1101 tification and updating of geological uncertainty with borehole data (au-
1102 tobel v1.0). *Geoscientific Model Development*, 13, 651-672. Retrieved
1103 from <https://gmd.copernicus.org/articles/13/651/2020/> doi:
1104 10.5194/gmd-13-651-2020
1105 Zhang, J., Zeng, L., Chen, C., Chen, D., & Wu, L. (2015, 1). Efficient bayesian
1106 experimental design for contaminant source identification. *Water Resources*
1107 *Research*, 51, 576-598. Retrieved from [http://doi.wiley.com/10.1002/](http://doi.wiley.com/10.1002/2014WR015740)
1108 2014WR015740 doi: 10.1002/2014WR015740