

Stochastic inversion of gravity, magnetic, tracer, lithology, and fault data for geologically realistic structural models: Patua Geothermal Field case study

Ahinoam Pollack^{a,*}, Trenton T. Cladouhos^b, Michael W. Swyer^b, Drew Siler^c, Tapan Mukerji^a, Roland N. Horne^a

^a Stanford University, 367 Panama St, Stanford, CA 94305, USA

^b Cyrr Energy Inc., 4010 Stone Way, Suite 400, Seattle, WA 98115, USA

^c Geology, Minerals, Energy, and Geophysics Science Center, U.S. Geological Survey, Moffett Field, CA, USA

ARTICLE INFO

Keywords:

Stochastic inversion
Uncertainty quantification
Joint inversion
Structural models
MCMC
Genetic algorithms

ABSTRACT

Financial risk due to geological uncertainty is a major barrier for geothermal development. Production from a geothermal well depends on the unknown location of subsurface geological structures, such as faults that contain hydrothermal fluids. Traditionally, geoscientists collect many different datasets, interpret the datasets manually, and create a single model estimating faults' locations. This method, however, does not provide information about the uncertainty regarding the location of faults and often does not fully respect all observed datasets. Previous researchers investigated the use of stochastic inversion schemes for addressing geological uncertainty, but often at the expense of geologic realism. In this paper, we present algorithms and open-source code to stochastically invert five typical datasets for creating geologically realistic structural models. Using a case study with real data from the Patua Geothermal Field, we show that these inversion algorithms are successful in finding an ensemble of structural models that are geologically realistic and match the observed data sufficiently. Geoscientists can use this ensemble of models to optimize reservoir management decisions given structural uncertainty.

1. Introduction

Globally, 25% percent of greenhouse gas emissions result from electricity generation that is powered by burning fossil fuels (Pachauri et al., 2014). To mitigate climate change due to these emissions, we must increase the portion of electricity generated by low-carbon resources, such as geothermal energy.

One of the major barriers for geothermal development is financial risk due to geological uncertainty (Gehringer and Loksha, 2012; Salmon et al., 2012; Speer et al., 2014). Drilling a geothermal well costs several millions of dollars, yet there is always uncertainty regarding whether the well will produce hydrothermal fluids or not. A well will only produce fluids if it intersects a permeable fault carrying hydrothermal fluids. The locations of the faults in the subsurface, however, are uncertain. Geoscientists need to reduce the uncertainty regarding the subsurface fault structures to alleviate the risk of drilling a "dry well."

Improvements in fault characterization methods can also open access to significant geothermal resources. While 3 GW of geothermal energy

are currently produced in the United States, equivalent to 0.2% of the total U.S. energy consumption (U.S. Energy Information Administration, 2016), the U.S. Geological Survey assessed that an additional 30 GW of geothermal resources exists in hidden geothermal systems in the U.S. (Williams et al., 2008). In hidden geothermal resources, hydrothermal fluids upwelling along faults do not reach the surface to form an easy-to-spot hot spring, fumaroles, or hydrothermal mineral deposits, but may rather be covered by the accumulation of sediments or lava bodies. Even in locations with surface manifestation of geothermal activity and faulting, the exact location of faults and their geometry below the surface can still be difficult to pinpoint. Improving the techniques for delineating fault structures will help both to find the largely hidden geothermal resource in the U.S. and to minimize geothermal exploration risk, incentivizing further geothermal development.

Quantifying geological uncertainty is a vital component of improving exploration techniques. Uncertainty quantification can help determine the financial risk associated with different reservoir development decisions, such as where to drill a new well or stimulate an

* Corresponding author.

E-mail address: ahinoamp@gmail.com (A. Pollack).

existing one. Estimating and visualizing geological uncertainty can not only help identify locations with higher likelihood of success but can also assist in communicating with investors and insurance companies.

Many of the current methods for characterizing fault structures are deterministic and do not properly quantify geological uncertainty. These methods create only a single geologic model without uncertainty information. Current stochastic inversion methods are often not geologically realistic or only invert for a few data types. In this paper, we address these shortcomings by describing a method to create an ensemble of realistic structural models that match five different observed data types: gravity, magnetic, granite top well markers, wellbore fault markers, and tracer data. The workflow is applied to case study of the Patua Geothermal Field in Fernley, Nevada in the United States. This work shows both how to reduce geological uncertainty by jointly leveraging multiple data types and how to create an ensemble of geologic models that conveys the posterior uncertainty regarding the faulting structures in the subsurface.

In the sections below, we discuss background information and prior work, the inversion methodology, and the results of the stochastic inversion.

2. Background and prior work

Geoscientists characterize fault structures and create computational structural geologic models by collecting different datasets and interpreting them in combination with background knowledge about local geology. In the following subsections, we describe the typical datasets used to characterize fault structures and review previous literature on inversion methodology.

2.1. Typical datasets for characterizing fault structures and their associated uncertainty

Typical datasets that are collected for structural inversion in a geothermal setting can be broken into four main types:

- Geological: such as surface geology, stratigraphic horizon, dips, and surface fault traces with subsurface projection.
- Geophysical: such as gravity, magnetic, magnetotelluric, and seismic.
- Wellbore: such as mud logs with lithology and lost circulation, temperature-pressure logs, and geophysical logs.
- Flow-related data: such as flow tests with pressure transient analysis, tracer data, and pressure-spinner-logs.

Each of these data types provides different information regarding the location and geometry of faults in the subsurface. Yet, none of the data sources provide complete certainty regarding the faults' locations, and therefore geoscientists often need to leverage their individual professional experience when interpreting the data sources.

Geological data can provide information regarding the surface trace, dip and thickness of faults partially exposed on the surface. Geological data can be projected to depth to predict the geometry of faults below the surface, but these interpretations lose certainty with increasing depth. Geophysical data are used to infer fault locations indirectly from contrasts in measurements of physical properties or fields, such as the gravity potential field. Gravity varies everywhere on earth based on the local density of rocks. Faults displace rocks with different densities, creating areas with sharp lateral differences in rock density. Therefore, geophysicists can detect faults by observing sharp contrasts in gravity measurements. This interpretation, however, does not have a unique solution, and commonly geophysicists will interpret the contrasts differently.

Drilling a wellbore into the ground is an expensive way to detect faults, usually costing several million dollars. After the well is drilled, various sensors and measurement tools are lowered into the well to log

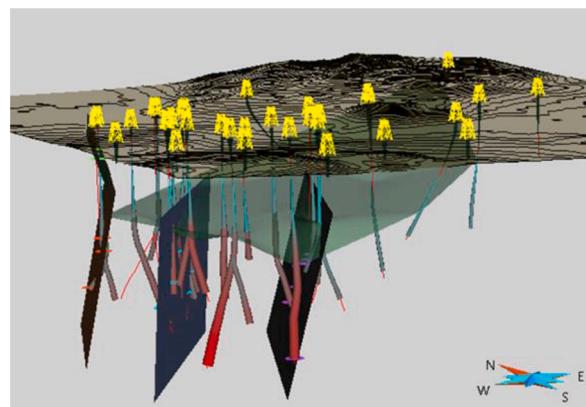


Fig. 1. Three-dimensional representation of the Patua structural model from Cladouhos et al. (2017), where the dark surfaces are faults and the green surface is the top of the granite. The wellbores in the model show the portion of well markers used to match the four fault surfaces. The color of the wellbore, ranging from blue to red, indicates the relative static temperatures of the wellbore from cold to hot.

the properties of the earth at depth. For example, if the well intersects a fault with flowing hot water, a temperature log will show an isothermal section extending between an inflow and an outflow point. A reservoir engineer can then either mark the inflow and outflow points as locations where the well intersects two separate faults, or as an extensively fractured fault zone. Other logs, such as image logs and mud logs, can also detect faults via sinusoidal traces seen on microresistivity measurements and zones where drilling fluid is lost to a permeable structure, respectively. Yet, well logs do not indicate the geometry and extension of faults away from the wellbore. Wellbore data only give information regarding points where faults intersect the wellbore.

In mature geothermal fields with several wells, tracer data can be used to characterize hydraulically permeable faults. In geothermal reservoirs, the matrix permeability and porosity are often very low, and therefore flow connectivity between wells is often attributed to faults. Geoscientists can inject a chemical tracer into one well and monitor the tracer concentration in other wells over time. If a tracer is recorded in a monitoring well, then this hydraulic connectivity may indicate a fault network extending between the injection and monitoring wells. Yet, this does not mean the wells are directly connected by a single fault that forms a straight line between the two wells. The tracer data does not indicate fault geometry, only a topological indicator of well connectivity.

Each of the data types separately leaves room for many possible interpretations of the subsurface structure. These data types complement each other and need to be combined to attain the most information about the location of faults in the subsurface. Yet, even when the data sets are combined, there is still great uncertainty regarding the subsurface structural model. The task of the geoscientist is therefore twofold: to combine all these data types together and leverage their complementary information as well as to estimate the uncertainty regarding the different possible elements in the subsurface.

2.2. Review of methods used to generate structural models

Typically, geoscientists create structural models that fit collected datasets by convening a group of domain experts to interpret the relevant data sets individually. After the geologists, geophysicists, and reservoir engineers interpret the datasets, these domain experts combine the various interpretations into one consistent interpretation regarding the location of faults. This is an extremely challenging and iterative process because interpretations of the individual datasets often do not match each other initially. The whole analysis and integration procedure can often take months because it is done manually.

For example, a previous numerical reservoir model of Patua Geothermal Field, the case study area of this paper, was constructed using a deterministic method and used for resource forecasting. Cladouhos et al. (2017) evaluated the regional tectonic regime, tracer and well tests, reservoir operational data, temperature profiles, and wellbore data, especially zones of lost circulation in the mud logs. Based on the connectivity indicated in the tracer data, three primary faults and two secondary faults were delineated. The geometry of the planar faults was assessed by using a least squares algorithm to fit points of lost circulation in the connected wellbores. This structural model, visualized in Fig. 1, does a good job of explaining the reservoir performance, tracer data and some of the wellbore data, but did not attempt to fit other important data types, such as the gravity, magnetic, and granite markers. In addition, the created model does not include a balanced geological section showing the displacement of the granite top due to faulting, nor does it extend beyond the developed reservoir area.

Jolie et al. (2015) described an advanced version of a typical deterministic workflow for creating a complex three-dimensional structural model of fault controlled geothermal systems. In this workflow, geoscientists draw several cross-sections of the subsurface using mapped surface faults and surface geological boundaries, interpretation of two-dimensional seismic reflection profiles, gravity data, core data and drill cuttings. These cross sections in conjunction with the surface traces are then interpolated in three-dimensional space using geologic modeling software. As an example, they created a detailed structural model of Brady's geothermal field.

While the structural model created by Jolie et al. (2015) is highly complex, it does not contain information regarding the uncertainty associated with each element in the subsurface model and does not provide a quantitative measure of the model fit to the observed data. Subsequent work on the same Brady's model by Siler et al. (2016) shows that it is possible to construct some measure of uncertainty within a deterministic framework by using an uncertainty metric based on the cell distance from the input data. Interpretations are more certain closer to input data than they are in areas further away, where interpretations are extrapolated. This approach, however, only considers one of the elements that contribute to subsurface uncertainty: lack of data coverage. It does not address other major sources of uncertainty, such as the nonunique nature of interpretations and prior expectations regarding structural continuity. To provide this uncertainty information, it is necessary to create multiple models that match the observed data, a task that is often too time consuming to perform in a partly manual deterministic process.

To better address structural uncertainty and match observed data,

researchers have been developing computational algorithms for creating structural models. These methods computationally generate geomodel realizations and systematically perturb these models to find realizations with low mismatch between simulated and observed data. We will review below some examples of such methods, highlighting innovation along four elements:

- 1 **Task:** what data types were used, to determine which target rock properties or structural elements?
- 2 **Model parameterization:** how does the algorithm generate geomodel realizations?
- 3 **Score or likelihood:** what is the loss function or scoring function?
- 4 **Search method:** what is the search or optimization method?

Witter et al. (2016) performed an inversion with the task of finding the subsurface density distribution that would match observed gravity and geologic data at Brady's Geothermal Area. They used a model parameterization that kept the lithological unit boundaries and structure constant and varied the density of each layer with heterogeneous density properties. The proprietary iterative search method used in the program has a basic scoring of root mean squared misfit of the simulated versus the observed gravity data. The model parameterization used in their inversion scheme, as well as an inversion project by Vogt et al. (2012a), use spatially correlated heterogeneous rock density, which is more realistic geologically than a constant density model. Yet, a major shortfall of the work is the use of fixed geologic boundaries. The fault structures are the most important elements of a structural model for a geothermal field, and these were not varied. Similarly, but using a different search method, Jardani and Revil (2009) used an Adaptive Metropolis algorithm for the task of finding the permeability of 10 geological units at the Cerro Prieto Geothermal field in Baja California by jointly inverting temperature and self-potential data. Yet, they also consider the stratigraphic geometry as deterministic.

Several methods have been developed to vary geometry in the model parameterization. Fullagar et al. (2008) wrote an algorithm for the task of combining core hole lithology data with geophysical data to find stratigraphic surfaces. Their model parameterization method allows both rock properties and the surfaces to vary to match the geophysical data. The surface movements up and down are constrained by "control points" from drill hole data. The search method in Fullagar et al. (2008) yields a single resulting model, but there are other researchers who have used a combination of such a model parameterization together with a search method, Markov Chain Monte Carlo (MCMC), that is stochastic (Bosch et al., 2006; Chen et al., 2012). Compared to Witter et al., these algorithms are an improvement as they allow a variable geometry of the stratigraphic surfaces. However, these methods still cannot identify the

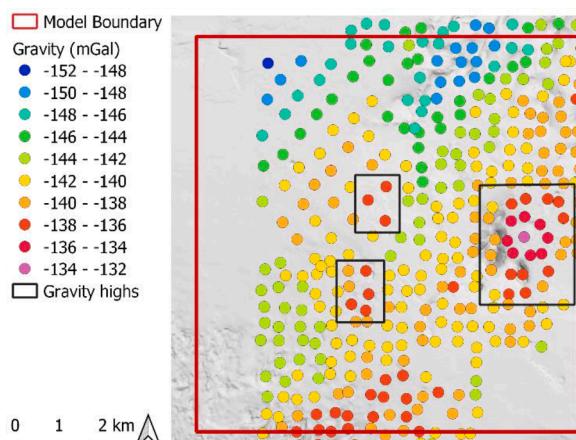


Fig. 2. Points show the measured gravity readings at stations after Complete Bouguer Anomaly Correction to an elevation of 0 meters relative to sea level using a density of 2400 kg/m^3 . The red rectangle marks the geological model boundary. Three black rectangles mark three oval shaped gravity highs.

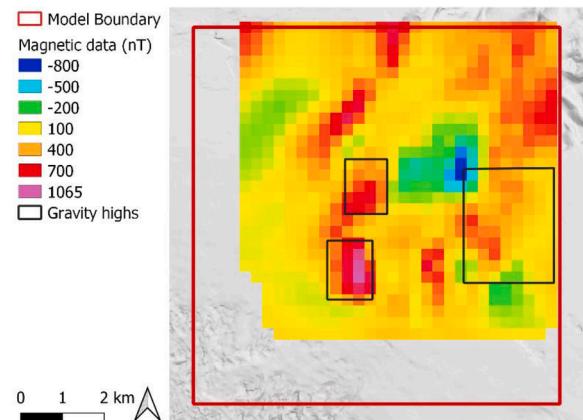


Fig. 3. Grid of reduced to pole magnetic data collected at Patua. The red rectangle marks the geological model boundary. Three black rectangles mark three oval shaped gravity highs that partially overlap with magnetic highs.

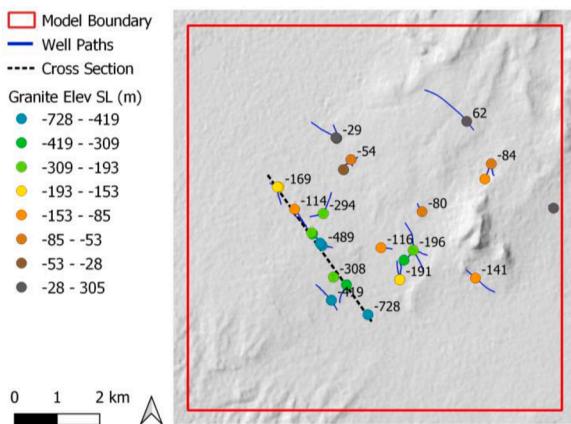


Fig. 4. Projection to the surface of wellbore markers of the top of the granitic basement. Both the point colors and the labels near the points indicate the position of the top of the granitic basement relative to sea level. The red rectangle marks the model boundary. The dotted black line indicates the location of the cross section shown in Fig. 5.

location of faults, except by analyzing the inversion results and looking for lineaments where the lithological surfaces have large sharp offsets.

Object-based model parameterization methods have been developed to invert for structural elements directly. A simple object-based model parameterization was developed by Mellors et al. (2014), in which the geometric aspects and location of a single fault controlling a geothermal system were allowed to vary as part of a stochastic inversion. Aydin and Caers (2017) performed a structural inversion to delineate faults conditioned on partially observed interpretation of seismic data. Their model parameterization is more complex than that by Mellors et al. (2014) and involved many faults and a marked Strauss point process for randomly positioning faults. To add more geologic realism to the fault objects, fault placement was restricted by age-relations rules between different fault groups. Object-based parameterization, however, still does not ensure a geologically balanced section, which may be important for geothermal applications that need to match datasets such as gravity, magnetics, and markers of the top of stratigraphic units, which depend on the movement of stratigraphy due to slip on faults.

Recently, there has been more research on the use of “potential field” interpolation approaches (De La Varga et al., 2019), not to be confused with potential-field geophysics. These methods can include faulting

events and surface discontinuities at fault location within the interpolation scheme and can directly condition to fault and stratigraphy dips and points. While potential field methods also do not ensure a balanced section, such methods can incorporate geologic rules. For example, Wellmann et al. (2018) used this method in conjunction with a likelihood term that rewards models for respecting certain geologic expectations regarding overlap.

Many papers have been written with a focus on enhancing the search method. For example, Farrell et al. (1996) used genetic algorithms to match gravity and magnetic data. M. Chen et al. (2014) used a multi-variate adaptive regression spline to create a surrogate model for computationally costly flow simulations that are needed as part of an inversion process. De Pasquale et al. (2019a, 2019b) used an empirical-Bayes-within-Gibbs algorithm, and have a sequential framework of first perturbing the geometry and only then perturbing the rock properties. Cui et al. (2011) used an adaptive delayed-acceptance Metropolis-Hastings algorithm and (Vogt et al., 2012b) used an Ensemble Kalman Filter.

2.3. Challenges with current methods

There are several issues with current methods for performing a computational inversion:

- Uncertainty quantification: some methods are deterministic and do not provide stochastic realizations for uncertainty quantification.
- Geologic realism: the parametrization of geothermal models often does not directly invert for faults, the central structures of geothermal reservoirs, or the structural models are too simple or not geologically reasonable.
- Number of input data: many of the algorithms often invert for only one or two observed data types. While theoretically, many methods can simply include additional data types, we have found that with five data types, additional adjustments to the optimization algorithm are necessary to address different misfit scales and different convergence rates.

We address these shortcomings of current methods by using the methodology described below.

3. Inversion methodology

In this section, we review our inversion methodology separated into

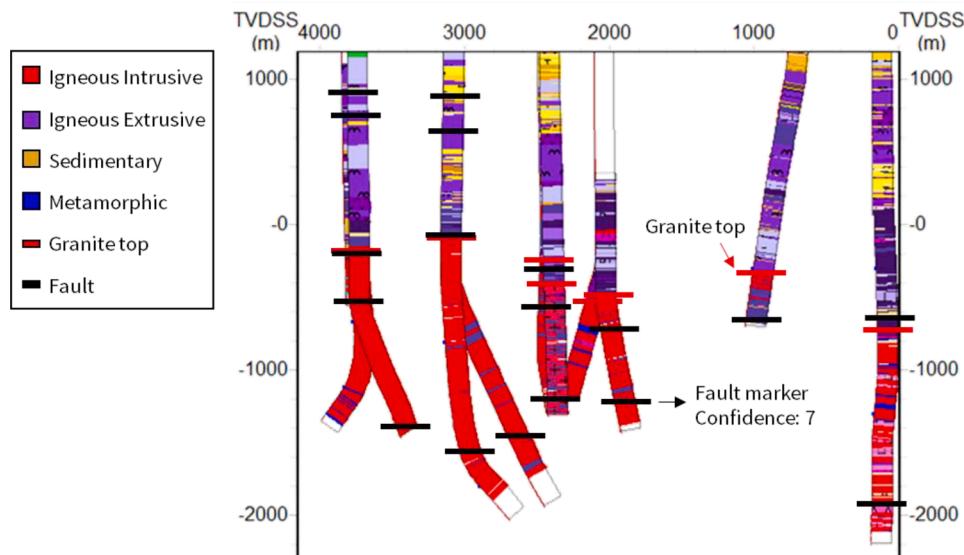


Fig. 5. A cross section through the Patua Geothermal Field reservoir along the line indicated by the black dotted line in Fig. 4 with projected mud logs and well paths for wells within 200 meters of the cross section. The mud logs show the rock formations in the wells. The red and black lines are markers of the position of the top of the granite basement in the wells and locations where we believe faults intersect the wellbore based on analysis of well logs. The elevation is marked by the acronym TVDSS (true vertical depth relative to sea level).

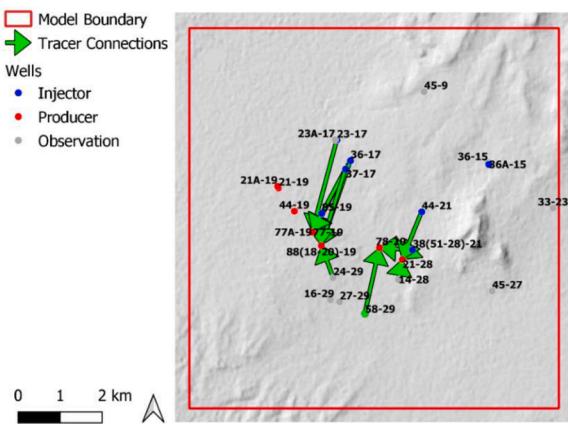


Fig. 6. Map of wellbores at Patua and connectivity found between wellbores via tracer tests. Hydraulically connected wellbores are shown with a green arrow connecting them, where the source of the arrow begins at the well that injected the tracer.

four components: inversion task, geomodel parameterization, mismatch calculation, and the search and optimization methods.

3.1. Task

The task of the inversion is to characterize the subsurface fault structures and the uncertainty associated with the fault locations at the Patua Geothermal Field. The goal is to have many diverse models that fit five real collected data sets: gravity, magnetic, granite top well markers, wellbore fault markers, and tracer data. The generated geomodel realizations have X, Y, Z dimensions of 8850 m, 9000, 3900 m, respectively, with a discretization of 150 m. Though, these values are flexible and can be easily changed in the code. The uncertainty quantification is achieved by creating and displaying the multiplicity of models that match the data. The resulting ensemble of matching models will then be used in future research on optimizing reservoir management decisions at the Patua Geothermal Field. The five input datasets are described in further detail in the subsections following.

3.1.1. Gravity data

Gravity data, shown in Fig. 2, were measured at 622 stations throughout Patua and were corrected to the Complete Bouguer Anomaly values assuming a rock density of 2400 kg/m³.

3.1.2. Magnetic data

Magnetic data were collected over the Patua field using an ultralight-borne magnetic survey. The data were leveled and reduced according to the International Geomagnetic Reference Field (IGRF), and this residual magnetic intensity was then reduced to pole (PRJ Inc. and Wave Geophysics, 2005). The gridded magnetic data are shown in Fig. 3.

3.1.3. Granite top data

32 of the wells at Patua (including side-tracks of wells) have mud logs that show a clear transition between volcanic rocks and granodiorite or granite basement rocks. Fig. 4 shows the points along the wellbore, projected to the surface, where the granite top was marked in the wellbore. The numbers next to the points indicate the position of the top of the granite surface relative to sea level.

3.1.4. Wellbore fault markers data

Mud logs, temperature logs, mud-loss logs, resistivity, and sonic logs were used to determine locations where faults intersect the wellbores. We marked fault picks with a confidence measure on a scale between 0 and 10 to indicate whether the pick was certain or questionable, respectively. These markers form the observed wellbore fault markers data. Fig. 5 is a cross section through the reservoir, showing several well trajectories and associated mud log data. This figure shows an example of both the granite top markers discussed in the previous sections as well as fault picks discussed in this section.

3.1.5. Tracer data

Tracer tests performed at Patua show hydraulic connectivity between various wells in the field. Fig. 6 shows the well connectivity determined by analyzing tracer returns above a certain threshold. Structural geologic models proposed during the inversion process will be rated based on their ability to recreate this observed connectivity between wells, either by a single fault or by a fault network.

Table 1
Ranges of values for log of magnetic susceptibility (SI units).

	Min	Max	Sources
Sedimentary	-6	-3	(Clark and Emerson, 1991; Schon, 2011)
Extrusive mafic	-4	-1.5	(Clark and Emerson, 1991; Schon, 2011)
Extrusive felsic	-6	-2.5	(Clark and Emerson, 1991; Schon, 2011)
Granitic	-4	-1.5	(Clark and Emerson, 1991; Schon, 2011)
Intrusion	-4	-1.5	Assumed same as mafic

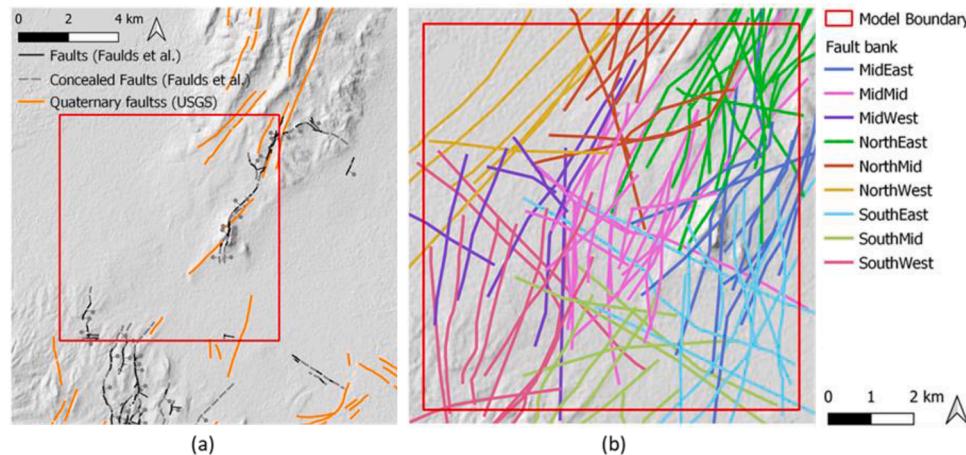


Fig. 7. (a) Locally mapped faults. Black and gray faults were mapped by Faulds et al. (2011). Orange faults were mapped by the U.S. Geological Survey (2006). The structural model boundary is shown in red. (b) A fault-traces ‘bank’ from which faults are randomly drawn during the optimization process. The fault selection process considers the zone of the faults, indicated by the various colors, in order to ensure an even distribution of faults in space.

Table 2

Range of property values for the geological and structural events.

Event 1: Stratigraphy				
Layer	Density (kg/m ³)	Log10(Magnetic susceptibility) (SI)	Thickness (m)	
Sedimentary	$U(2100, 2360)$	$U(-6, -3)$	$U(100, 350)$	
Volcanic mafic	$U(2200, 2420)$	$U(-4, -1.5)$	$U(250, 950)$	
Volcanic felsic	$U(2300, 2650)$	$U(-6, -2.5)$	$U(250, 950)$	
Granitic basement	$U(2500, 2850)$	$U(-4, -1.5)$	Not applicable	

Event 2: Tilting: Tilt: $U[0, 3.5]$ Tilt direction: $U[0, 360]$

Event 3-5: Intrusions				
Intrusion number	Intrusion density (kg/m ³)	Magnetic susceptibility (SI)	Intrusion radii (m)	Center of intrusion (m)
1	$N(2500, 2900)$	$N(-4, -1.5)$	X-Axis: $U(200, 800)$ Y-Axis: $U(300, 1000)$ Z-Axis: $U(1000, 2500)$	X: $U(6950, 7950)$ Y: $U(3980, 4980)$ Z: $U(-3400, -1400)$
2	$N(2500, 2900)$	$N(-4, -1.5)$	X-Axis: $U(150, 600)$ Y-Axis: $U(150, 600)$ Z-Axis: $U(1000, 2500)$	X: $U(3150, 4150)$ Y: $U(2880, 3880)$ Z: $U(-3400, -1400)$
3	$N(2500, 2900)$	$N(-4, -1.5)$	X-Axis: $U(150, 600)$ Y-Axis: $U(150, 600)$ Z-Axis: $U(1000, 2500)$	X: $U(3650, 4650)$ Y: $U(5080, 6080)$ Z: $U(-3400, -1400)$

Event 6-(6 + n _{faults}): Faults
--

Dip (°)	Maximum slip (m)	Deformation ellipse radii (m)	Center of deformation (m)
$U[45, 90]$ Dip direction: rand(East/ West)	$ratio_{slip} : N(0.05, 0.2)$ Slip:ratio _{slip} ·Length _{fault}	Length: from fault bank Depth: $U[0.25, 0.75] · length_{fault}$ Outward normal: $U[0.25, 0.75] · len_{fault}$	X, Y, Z: from fault bank

3.2. Model parameterization and prior uncertainty

Before rigorous matching of observed data, we reviewed the regional geologic data to understand the possible subsurface scenarios, in a step termed "quantifying prior uncertainty" (Mosegaard and Tarantola, 1995; Scheidt et al., 2018). Specifically, we investigated the likely fault dips and azimuths, and stratigraphic layers, rock density properties, and magnetic susceptibility properties. We set up a method for generating the geomodel realization based on a set of control parameters. The stochastic structural models generated according to the ranges of parameter values defined in this step are called "prior models." These models, which have not been tested for matching observed data, reflect the uncertainty regarding the subsurface prior to integrating the observed data.

3.2.1. Expected structural elements

Patua is in Western Nevada in a transition zone between the Walker Lane dextral (right lateral strike-slip) faulting zone and the Basin and Range Province normal faulting area. Specifically, Patua is uniquely located between three subdomains: Pyramid Lake, Carson, and Humboldt (Faulds et al., 2010; Faulds and Henry, 2008). Pyramid Lake is a dominantly northwest striking right lateral faulting zone. Carson is an east striking left-lateral zone (Faulds and Henry, 2008). Humboldt structural zone is a dominantly north-northeast striking normal faulting zone (Faulds et al., 2010). The zones are distinct in both faulting orientation, which changes from eastern to north-western directions, as well as faulting type, which varies from strike slip faulting in both Pyramid Lake and Carson (which are part of Walker Lane) to normal faulting in the Humboldt structural zone. Based on regional tectonic patterns, Patua may be expected to have more north-northeast normal faulting structures in the north of the field (characteristic of the Basin and Range province) and possibly some east-west or northwest strike-slip structures (characteristic of the Walker Lane) in the south of the field.

Faulds et al. (2011) and the U.S. Geological Survey (2006) mapped faults at Patua, shown by black and orange lineaments, respectively, in Fig. 7(a). These faults follow expected regional trends with north-northeast normal faults to the north and south of the geological model boundary, and east-west faults southeast as well as southwest of the geological model boundary. No faults are mapped in the center and west of the geological model boundary, either because the faults do not reach the surface, have been covered by sediments, or do not exist.

Based on this review of expected structural elements, we made a "fault bank" (Fig. 7(b)), from which to sample faults during the inversion process. We will explain the concept of the fault bank and details of the algorithm for generating geomodel realizations in Section 3.2.6.

3.2.2. Expected lithological units

Following a review of the surface and subsurface geology from mud logs, we have determined the model region has a basic geologic stratigraphy with the following layers:

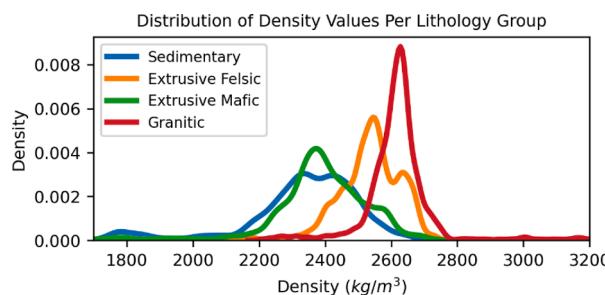


Fig. 8. Probability density function of the density of various lithology groups.

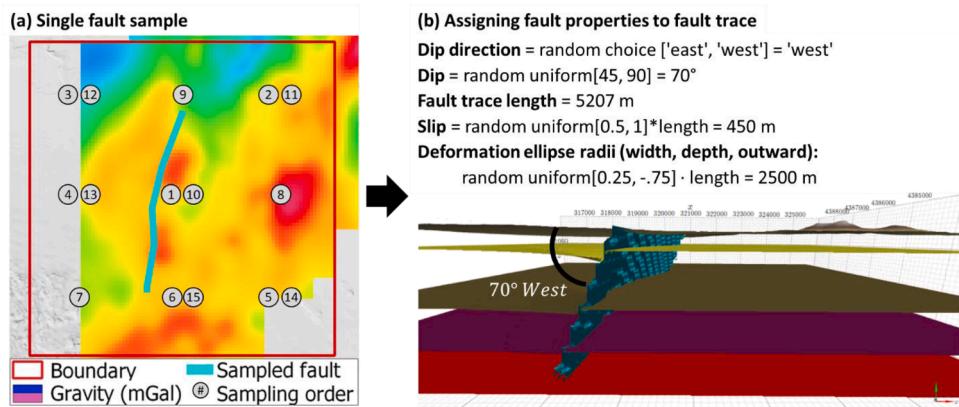


Fig. 9. The process of generating faults in the geomodel realization. (a) A fault is sampled from the fault bank from the first zone in the randomly generated zone order. (b) the fault properties are selected by sampling from the distribution of fault properties.

- 1 Sedimentary cover,
- 2 Mafic volcanic rocks,
- 3 Felsic volcanic rocks, and
- 4 Granitic basement rock.

The thicknesses of each of these layers prior to deformation is uncertain. The range of thicknesses is determined based on observations from mud logs and defined in Table 2.

3.2.3. Density of lithologic units

To match the gravity observations, it is necessary to define the bounds of density values of the different rock types. The density data collected at Patua consists of two types of density logs: ZDEN (formation bulk density) and ZDNC (borehole size/mud weight corrected density). Fig. 8 shows the distribution of rock density by lithology type. The density measurements show that in general the granitic basement has higher density than the volcanic rocks, which have higher density than the sedimentary rocks, as expected. The uncertainty regarding the density values is used during the stochastic generation of prior models by sampling density values for the lithology that are partially informed by the distributions found in this section and partially by initial analysis of inversion results.

3.2.4. Magnetic susceptibility of lithological units

To match the magnetic data, it is necessary to define a range of uncertainty for rock magnetic susceptibility values. There are no collected magnetic susceptibility data for the area. We therefore use the following estimated magnetic susceptibility data collected from different sources.

3.2.5. Model parameterization using PyNoddy

One of the most crucial steps for creating prior models is generating

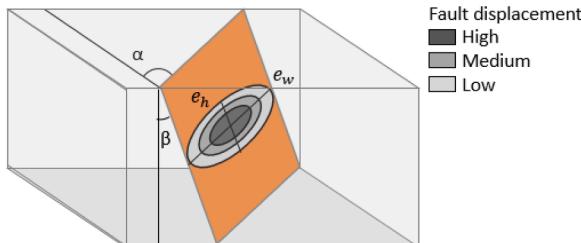


Fig. 10. Example of the parameterization of a planar fault with an elliptic fault displacement distribution, where fault displacement diminishes outwards according to a shape of an ellipsoid. The dark gray color in the center of the deformation ellipsoid on the fault indicates the area with the highest displacement, and the outer light gray ellipsoid indicates an area with reduced displacement.

geologically realistic subsurface models in a reasonable computational time. For this work, we used the kinematic structural modeling software PyNoddy (Wellmann et al., 2016), which is a python wrapper for the software Noddy (Jessell and Valenta, 1996). PyNoddy can be used to simulate faulting, folding, strain, shearing, rotation, dykes, and other geological events. The Noddy algorithm works by starting with a definition of flat stratigraphy, discretized into 25 to 250 m cubes, and applying a series of geologic deformation events to this discretized model. This process ensures that the final geomodel is a balanced section, meaning that the final stratigraphic and faulting configuration can be explained by a series of geological events from a predeformed state. In addition, PyNoddy can also forward-simulate the gravity and magnetic signatures of the geomodel based on the provided density, magnetic susceptibility and geophysical survey parameters.

We used PyNoddy for generating structural realizations because kinematic simulations strike a balance between geologic realism and computational speed. Physical process simulators, such as discrete element simulations that model fault growth and interaction of faults given stress boundary conditions (Finch and Gawthorpe, 2017), would create the most physically realistic structural models. Yet, these process-based simulations take significant computational time and are very difficult to condition to observed data. On the other hand, traditional stochastic methods in earth sciences can easily match observed data and are quick, but often lack geologic realism. For example, Sequential Gaussian Simulation creates smooth realizations of the subsurface and cannot generate fault discontinuities. PyNoddy is in the middle between these two extremes, being relatively quick (few seconds per simulation) and with the potential to be geologically realistic. PyNoddy can generate many geological events, including faults, the

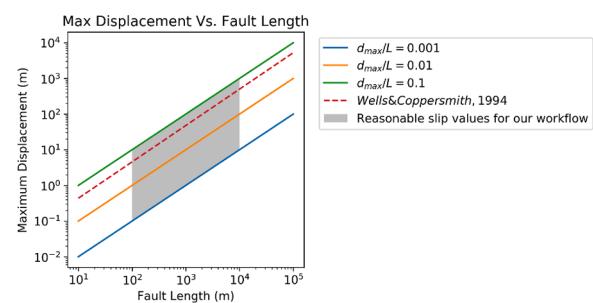


Fig. 11. Relationship between fault length and maximum fault displacement. The solid lines show the relationship between fault length and fault displacement following the equation: $d_{max} = cL^n$, when $n = 1$, and $c = \frac{d_{max}}{L}$, and given different values of $\frac{d_{max}}{L}$. The red dashed line showed the relationship observed in a study by Wells and Coppersmith (1994). The gray zone indicates points with acceptable values of slip given the fault length for our study.

cornerstones of fault-dominated geothermal systems. Crucially, PyNoddy simulates the displacement of stratigraphic layers along faults, given an assigned slip value. In our study, we condition to data such as gravity and magnetics, that show signatures that relate to the displacement of stratigraphy due to faulting and therefore PyNoddy's kinematic simulation captures this important physical process.

3.2.6. Algorithm for stochastic generation of structural models

Our overall parameterization of the geologic model has five base events and additional faulting events, as follows:

- Event 1: stratigraphy
- Event 2: stratigraphy tilting
- Events 3-5: three basaltic intrusion events
- Events 6+: a randomly chosen number of faulting events (n_{faults})

For generating a geological model in Noddy, each defined geologic event necessitates several input parameters. For example, each strati-

graphic layer needs to be defined with a density, magnetic susceptibility, and thickness. Each intrusion has a radius in three directions, a location, as well as a density and magnetic susceptibility. These input parameters are sampled from Table 2, which aggregates the ranges of parameters' uncertainty defined in previous sections. The overall model creation algorithm is:

```

nfaults = random(min = 10, max = 20)
Select nfaults from the fault bank
Randomly select each fault's dip and slip (relative to length)
Sample stratigraphy parameters from Table 2
Sample tilt event parameters from Table 2
Sample intrusion parameters from Table 2
Create Noddy input file
Run Noddy using input to simulate geology and geophysical response

```

We added the tilting and intrusion events based on analysis of the granite top and gravity, respectively. The data regarding the top of the granite show that the top of the granite is higher towards the northeast

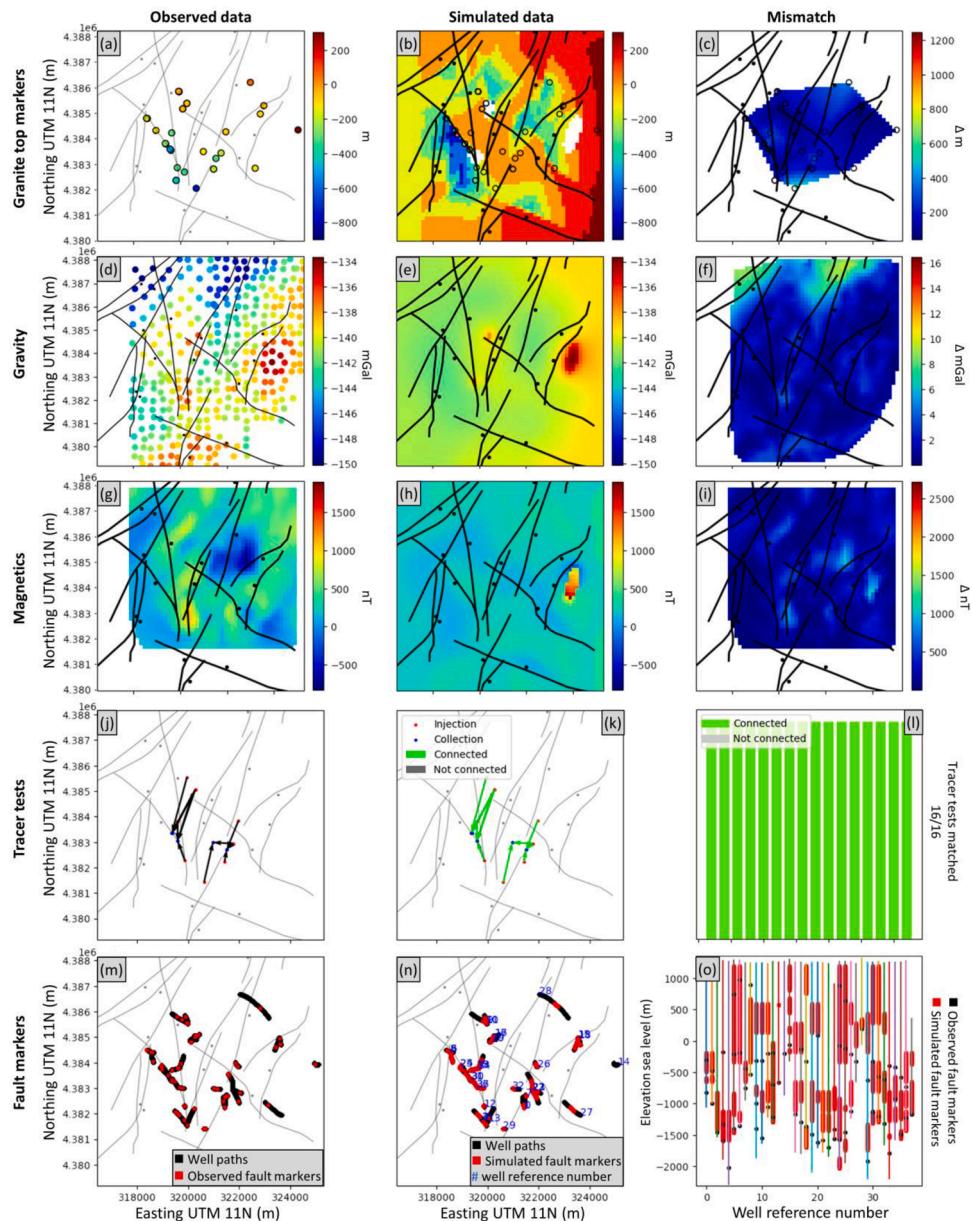


Fig. 12. Example of a single optimization job. The first column shows the observed data for the five different observed data sets. The second column shows the simulated data. The third column shows the absolute value of the mismatch between the simulated and observed data.

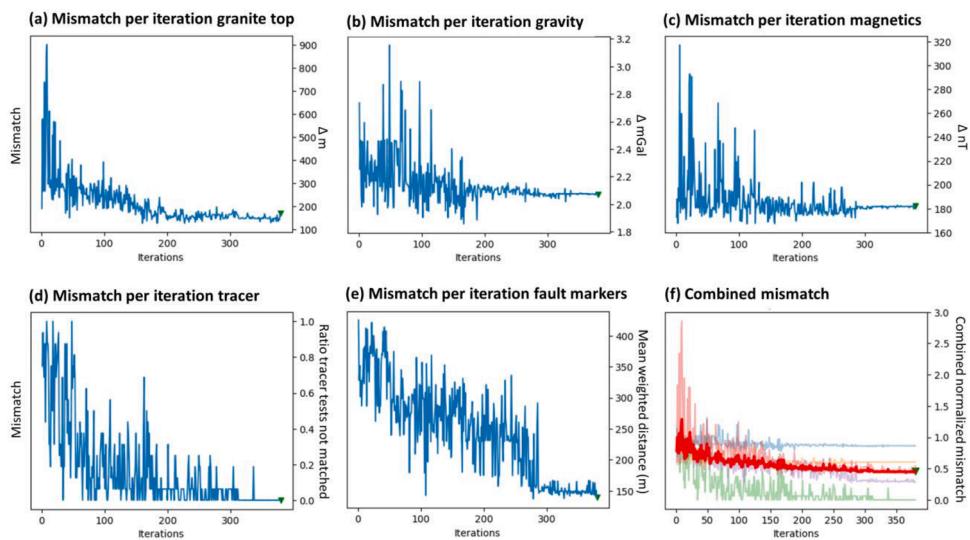


Fig. 13. The optimization progression of mismatch per iteration for (a) granite top markers, (b) gravity, (c) magnetics, and (d) tracer data. (f) The combined normalized mismatch of all data types in red and the individual normalized mismatch in light colors.

of the field, indicating a possible tilt in the subsurface. Therefore, to address the tilting, we added the tilting event. It is possible that this topography on the granite is not due to a tectonic tilting event, but due to the granite intruding to shallower depths to the northeast. For the limited extent of the geological model in this study, having a tilting event achieves the same geological geometry as an uneven intrusion.

The intrusion events were added to match gravity and magnetic data.

Initial runs of the optimization were unable to match three oval-shaped gravity highs in the observed data, shown in Fig. 2, as well as partially collocated magnetic susceptibility highs, shown in Fig. 3. We have therefore added three basaltic intrusions that may explain the oval signatures in the data.

To model the uncertainty regarding the existing fault structures, we tested three methods of initializing faults in the geologic model: expert-

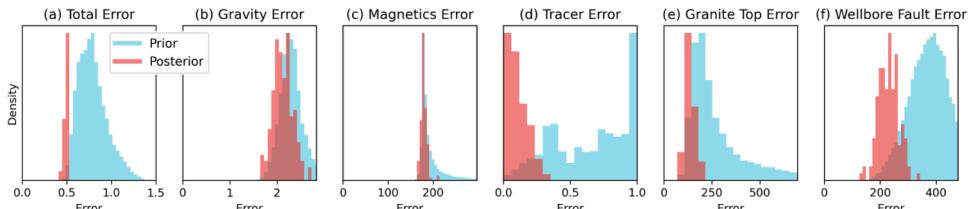


Fig. 14. (a) Distribution of combined misfit values, with light colors indicating the prior models, and the red colors indicating the posterior models. (b-f) Distributions of misfit values for the different data types in this study.

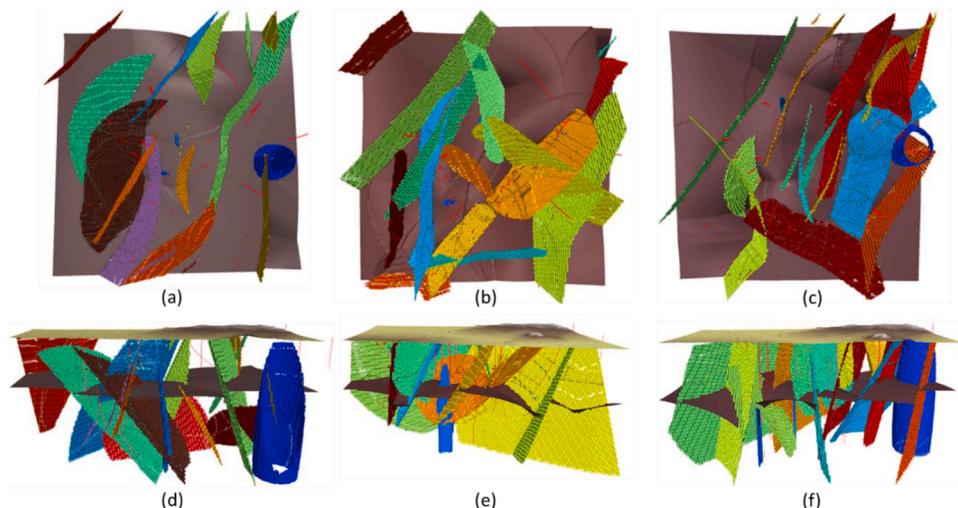


Fig. 15. An example of three posterior structural models from top view (a-c) and side view (d-f). Red lineaments indicate well trajectories, colored surface indicate faults, the dark brown layer is the top of granite surface, the blue cylinder is a volcanic intrusion, and the top layer in the side view is the site topography. The 3D visualizations are made using the open source package vedo ([Musy, 2021](#)).

created faulting scenarios, randomly generated faults, and sampling from a "fault bank." For the expert method, we initialized stochastic realizations with one of several fully defined fault characterization scenarios created by structural geologists (Pollack et al., 2020). We then perturbed the dip, length, and location of these faulting scenarios. This is similar to the method employed by Suzuki et al. (2008) and Cherpeau et al. (2012) who relied on expert interpretation of faulting scenarios to initialize structural models. Using this method, we were unable to get a good fit to the observed data.

We also tested random generation of fault traces, where the model was initially populated by random straight fault traces. As part of the inversion, the control points of these faults were perturbed with spatially correlated offsets. After several iterations, the fault traces could form almost any possible fault trace, which led to unrealistic fault shapes.

In this paper, we use the third "fault bank" method. We created a bank of fault-traces drawn by geologists, see Fig. 7(b), from which to randomly select 10-20 faults to initialize the geomodel realizations while running the inversion algorithm. We created these 107 faults based on large spatial gradients (discontinuities) in the gravity and magnetic data, as well as nearby mapped faults. For example, see the single sampled fault in Fig. 9 that marks a large horizontal gradient in the gravity data. By starting with faults that are already informed by the potential fields data and the local structural geology, the algorithm can converge faster to models with good matches. We chose to use this method because it strikes a balance between having too few degrees of

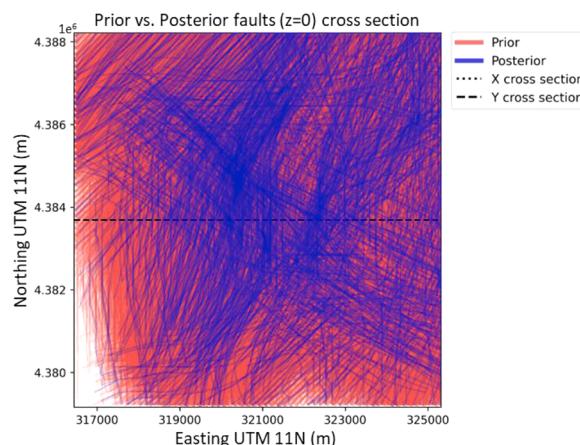


Fig. 16. Map view of prior and posterior faults at a constant-z section (0 m elevation). The black dashed lines show the location of the cross section in the figures below.

freedom and on the other hand, having unrealistic fault traces.

Faults are sampled to ensure an even distribution of faults throughout the model area. Faults are assigned to different regions of the model, shown by the different fault colors in Fig. 7(b). During the initial fault selection process, faults are sampled in a random order from the different zones, such that there will be at least one fault from each of the nine zones and some zones will have multiple faults. A random zone-sampling order is illustrated in Fig. 9.

After sampling fault-traces, three-dimensional faults are constructed from these fault traces by assigning fault dips, fault slips, and elliptical displacement radii lengths to the fault traces according to predefined probabilities. The fault dips are sampled from a uniform distribution between 45° and 90°, based on the possible dips for normal to strike-slip faults (Anderson, 1951) and image log data possibly showing shallower dips. The ratio of fault slip to fault length is sampled from a uniform distribution between 0.05 and 0.2, and this ratio is used to assign fault displacement in proportion to fault slip, see Section 3.3.2 for more detail. During the inversion process, the fault traces are perturbed by rotating, stretching, and translating the fault traces to a limited extent to match the observed data. Fig. 9 illustrates how a sampled fault trace is transformed into a 3D fault by sampling its fault properties.

Other researchers have used different methods of sampling faults. For example, Cherpeau and Caumon (2015) generated fault networks by placing faults in the model such that they pass through one of the "fault sticks" interpreted on a seismic cross section. While the method we use attempts to initialize the structural model with faults that would match the geophysical data, their method generated faults that already honor some of the interpreted fault sticks that are based on seismic data. For our geothermal study, where there are no high-quality seismic data, future work could explore whether initializing fault locations based on the fault well marker data would yield better results. Another great enhancement for our current workflow would be to generate fault traces that consider fault interactions. For example, Bonneau et al. (2013) generate fault realizations by using geometric proxies for mechanics of fracture growth. Their algorithm provides a geometrical proxy for fault linking, stress shadowing and fault termination, that makes fault realizations more geomechanically realistic.

3.3. Heuristics for geologic realism when using PyNoddy

One of the downsides of using Noddy when compared to a physics-based simulator is that the software does not ensure geologic realism, such as reasonable fault slip values and fault geometries that make mechanical sense when considering fault propagation interactions. For example, it is possible to input into Noddy a fault slip value that is larger

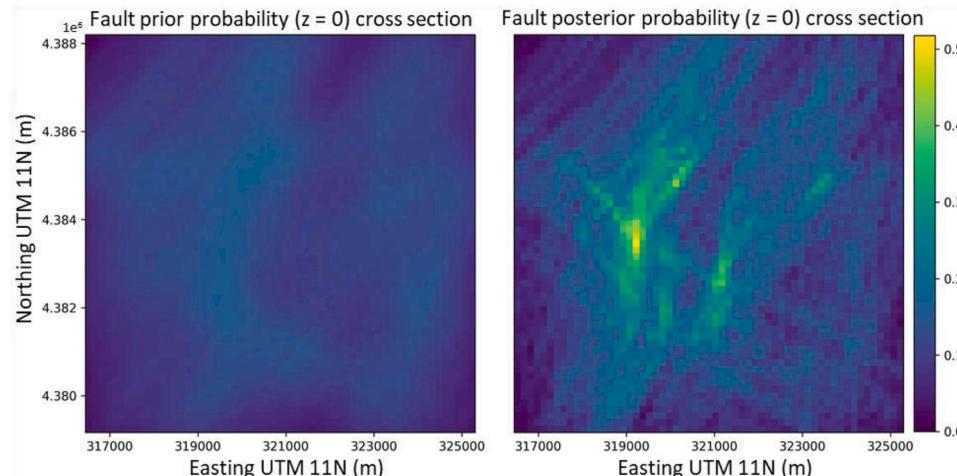


Fig. 17. Map view of the probability of there being a fault for the ensemble of prior models (left) and posterior models (right).

than the fault length value, while such a scenario does not exist in nature. In addition, it is possible to generate small faults that cross large faults, or faults that overlap other faults, while in nature, the smaller faults may terminate against and not cross the large faults. To better ensure geologic realism, we have added several heuristic rules to our workflow:

- 1 The faults should be chosen from a "bank" of possible fault traces drawn by geologists and only moderately modified from those initial fault drawings.
- 2 Faults should not have a constant displacement across their surface, but rather should be defined as elliptic, where the displacement (slip) is maximal at the center and diminishes towards the edges.
- 3 The fault slip should be reasonable given the fault length.
- 4 Fault traces should not overlap for more than a predefined portion of their length.
- 5 East-west faults are considered more recent and therefore should be more likely to cut north-northeast faults.

We discussed the first constraint of using a "fault-traces bank" in Section 3.2.1 and discuss the other four in the sections that follow.

3.3.1. Elliptical faults

Researchers widely accept that while faults themselves might be planar or curviplanar surfaces, the distribution of fault displacement is elliptical, where displacement is maximal at the center of the fault, and diminishes to zero at the edges of the fault, following an ellipsoidal shape (Kim and Sanderson, 2005), see Fig. 10. Noddy has the option to generate not only planar faults, with constant displacement across a surface, but also elliptic faults. This option should be used for all input faults to preserve geologic realism.

3.3.2. Fault slip vs. fault length

There is a relationship between fault length and fault displacement: larger faults have larger displacement and vice versa. Kim and Sanderson (2005) find the maximum fault displacement relates to the length of the fault via the following equation:

$$d_{max} = cL^n \quad (1)$$

where d_{max} is the maximum displacement, c is a constant, L is the fault length, and n is an exponent. Often, an exponent of $n=1$ is observed in field studies (Kim and Sanderson, 2005), leading to a linear relationship between the maximum fault displacement and the fault length, as shown in Fig. 11. For normal-fault settings, this ratio between maximum fault displacement and fault length is typically between 0.001 and 0.1, based on several studies aggregated by Kim and Sanderson (2005) as well as a study by Wells and Coppersmith (1994). In our modeling workflow, we are only representing faults with a length between 100 and 10,000 m and therefore the slip values should always fall within the gray zone indicated in Fig. 11.

3.3.3. No completely overlapping fault traces

The optimization algorithm can randomly select faults that almost completely overlap each other. In reality, such faults should quickly terminate against each other or link into one another. Because such termination capability is not possible via PyNoddy, we remove faults that overlap for more than 25% of their length (with a 100 m buffer to account for the simulation cell size) with another fault. Future work should implement in PyNoddy a capability to terminate faults or link faults together based on geometrical proxies for stress shadowing effects (Bonneau et al., 2013).

3.3.4. Fault age relations

East-west faults in the Patua region are considered part of the Walker Lane complex and typically are more recent than north-northeast faults formed due to Basin and Range extension. Therefore, east-west faults should be more likely to cut north-northeast faults. Therefore, we implemented a mechanism where when an east-west fault intersects a north-northeast fault, there is a 70% chance that the east-west faults will cut the north-northeast faults. We do not have statistics regarding the relative percentage of new faults that are oriented east-west versus north-northeast, but this is an estimate based on knowledge of general trends in this region, that east-west faults are in general more recent, but not always.

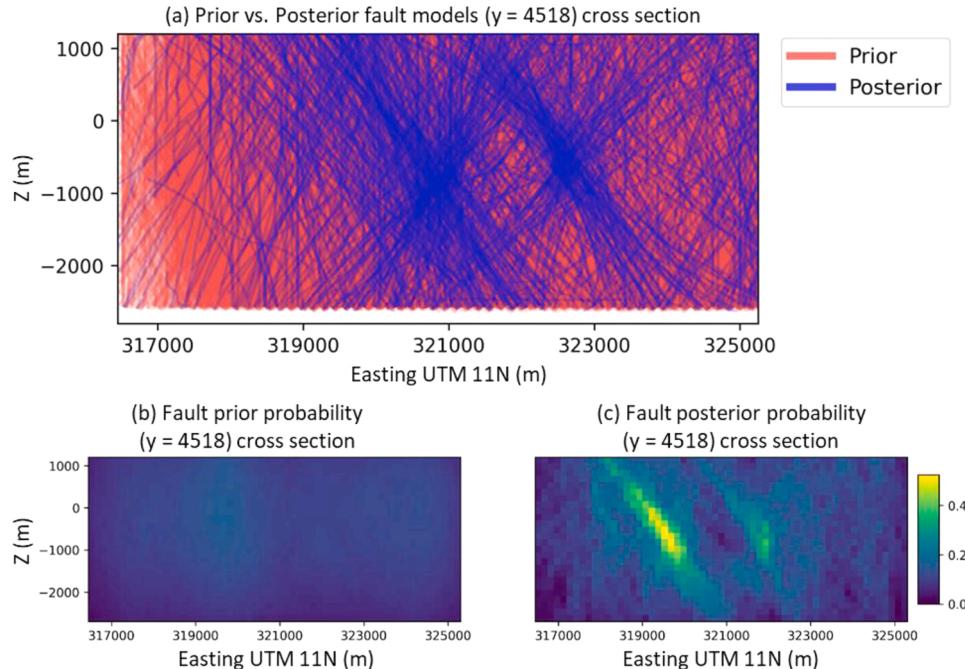


Fig. 18. (a) Side view (constant y section) of the prior and posterior fault lineaments. (b-c) The prior and posterior probabilities of there being a fault at a given cell in the section.

3.4. Scoring or likelihood criterion

We used the scoring criterion of misfit between the observed and simulated data and did not assign any extra scoring for geologic realism or other criteria. In the following, we review the misfit details for each data type.

3.4.1. Gravity data

We used PyNoddy to simulate the gravity, then performed a datum shift so that the simulated gravity data has the same median as the observed gravity data. The misfit is then the mean absolute difference between the observed and the simulated gravity data, which is also called the L1-norm of the misfit.

3.4.2. Magnetic data

The magnetic misfit is the mean absolute difference (L1-norm) between the gridded magnetic data at a resolution of 250 m by 250 m and the magnetic data simulated by PyNoddy following a datum shift.

3.4.3. Granite top data

When simulating the geology in PyNoddy, the top of the granite layer moves up and down at different locations due to faulting and due to the chosen thickness of the layers above the granite. The misfit is the mean absolute difference between the simulated top of granite and the observed top of granite in the well logs.

3.4.4. Wellbore fault markers data

After simulating faults in PyNoddy, the locations where simulated faults intersect the wellbore is calculated. Then the misfit for this data type is the sum of the distances between the observed fault markers to the nearest simulated fault markers in the same wellbore. The misfit is calculated as follows:

$$\text{misfit} = \sum_{\text{mrkr}_{\text{obs}}=0}^{\text{n}_{\text{mrkr}_{\text{obs}}} \text{ confidence}_{\text{mrkr}_{\text{obs}}} \cdot \min(\text{abs}(Z_{\text{mrkr}_{\text{obs}}} - Z_{\text{mrkr}_{\text{sim}}}), \text{MaxErr}_{\text{mrkr}})} \quad (2)$$

where $\text{confidence}_{\text{mrkr}_{\text{obs}}}$ is the confidence associated with each marker, Z_{mrkr} is the Z position (elevation) of the observed or simulated marker, and $\text{MaxErr}_{\text{mrkr}}$ is a hyper parameter that provides a value for the case in which there are no simulated faults that intersect a wellbore. For this study, we chose for it a value of 500 m.

3.4.5. Tracer data

For our inversion workflow, we assumed that all well connectivity is due to faults because the granitic basement that forms the geothermal reservoir has low matrix permeability. By analyzing the output of the PyNoddy topology program (Thiele et al., 2016), we generate a connectivity matrix that indicates which faults are connected. We then calculate which faults each well intersects. Then, using the Dijkstra algorithm (Dijkstra, 1959), we establish which wells are connected to each other via faults. While tracer tests inform about both the existence (or nonexistence) of connectivity between wells as well as the strength of that connection, in this study, we only use the information regarding the presence of connectivity. If a well pair is connected in the observed tracer data, we check for connectivity between those wells in the simulated data. Therefore, the misfit for the tracer data is calculated as follows:

$$\text{misfit} = \sum_{\text{well pair}=0}^{n_{\text{tracer pairs}}} \begin{cases} 1 & \text{if well-pair not connected via faults} \\ 0 & \text{if well-pair is connected via faults} \end{cases} \quad (3)$$

3.4.6. Joint scoring

A significant challenge in a joint inversion is creating a combined fitness metric for variable data types with different ranges of misfit values. For example, the gravity data has an error range approximately

between 0.5 mGal to 2.5 mGal. On the other hand, the magnetic data has an error range between 600 nT and 1000 nT. We used two different methods to join the errors of different data types.

- 1 Normalization: we measured the mean error of each data type for the first $n_{\text{exploration}}$ rounds, and then used this mean error to normalize the different data types and combine them together. We used this normalization method for the MCMC, simulated annealing, and genetic algorithm search methods.
- 2 Dominance: the nondominated sorting algorithm does not necessitate combining the errors. Models are kept if they “dominate” in terms of low misfit for at least one of the data types.

3.5. Search algorithms

In this study, we separately used and compared four search algorithms:

- 1 Markov Chain Monte Carlo (MCMC) with a Metropolis sampler (Mosegaard and Tarantola, 1995),
- 2 simulated annealing (Laarhoven and Aarts, 1987),
- 3 basic genetic algorithm (Back et al., 2000; Holland, 1962), and
- 4 Nondominated Sorting Genetic Algorithm, NSGA II (Deb et al., 2002).

We ran equal number of inversion jobs for each of these methods and checked which method performed best. We then took the resulting posterior models from runs of each method and combined them together to create fault probability maps.

We compared these four inversion methods because each one has advantages over the others. The MCMC search algorithm is the only algorithm that correctly samples the posterior, but its convergence time may be long. While simulated annealing and genetic algorithms do not rigorously sample the posterior, they can be used to generate an ensemble of models that match the observed data and potentially may converge faster or achieve lower misfit with the observed data. NSGA II was chosen because it is supposed to perform well with multiple objectives (Deb et al., 2002).

We did not implement a gradient descent method because the misfit is not differentiable and estimating the gradient via a finite difference approximation would necessitate several evaluations for each step. In addition, the misfit space is highly nonlinear, and it is therefore highly likely that gradient descent methods would easily get stuck in local minima. Many other search and sampling methods may be viable, and a more rigorous comparison of search algorithms is a good topic for future research.

These four search and optimization routines discussed in this paper are not included in PyNoddy. Rather, we have coded these search algorithms either from scratch or based on existing packages and integrated them with PyNoddy. The code calls the PyNoddy simulator and analyzes PyNoddy's simulation results as part of the search algorithm. We describe these optimization algorithms in further detail in Appendix A. Appendix B reviews an additional step of prior falsification that was performed as part of this study's workflow.

4. Results and discussion

We ran 5000 jobs of stochastic inversion using the four different optimization algorithms described in Section 3.5. An example of the result of a single MCMC optimization job with low error is shown in Fig. 12. Each row shows the observed data, the simulated data, and the absolute mismatch between the observed and simulated data for one of the five data types: gravity, granite top markers, magnetic, tracer connection, and fault markers. This example realization has dominantly north-northeast faulting with some east-west faults. As can be observed in the mismatch column, the realization shows a good match with the

granite top, tracer, and fault markers data.

Fig. 13(a-f) show the progression of the mismatch per iteration. As the iterations progressed, the mismatch decreased for all the varying data types. **Fig. 13(d)** shows that the fit for the tracer tests is perfect. The fit for the granite top markers and the fault markers is close to 150 m on average, which is the size of a single cell in the geological model. The error for the magnetic field and gravity field decreased the least. Even though the simulated granite top data shows significant offsets in the top of the granite, the density and magnetic susceptibility values were insufficient to create the strong contrast in gravity and magnetic data values seen in the observed data. **Fig. 13(f)** shows the combined normalized error of all data types, which decreases per iteration.

Fig. 14(a) shows the distribution of the combined error for each job (five samples per job). All models with error lower than 0.51 were classified as matching the data. The optimization and search process found models with half of the typical starting mismatch for the different data types. In the plot, we use Bayesian framework terminology and call these models "posterior models," or models that represent our beliefs about the possible subsurface structures following (post) integrating the observed data.

The distribution of the mismatch values of the five different data types are shown in **Fig. 14(b-f)**. The plots show the tracer data converged the best, often reaching zero error, showing that the connectivity element of tracer data is very easy to match, as it only includes connectivity but no geometric information. The granite top and wellbore fault markers also decreased substantially as part of the optimization process. On the other hand, the gravity and magnetic data have

substantial room for improvement in terms of convergence. In the discussion, we will suggest some ways to better respect these geophysical data types.

Three low mismatch posterior structural models are shown in **Fig. 15**. The top row shows a top-view and the bottom row shows a side view of the geological models. The top of the granitic layer is shown by the dark brown color in the top view, and the colored surfaces are the fault structures. The blue cylinder is a volcanic intrusion. In the side view, the top light brown surface is the site topography. The red lineaments are the well trajectories. One can see in these models that the faults displace the granite top layers as should be the case in a balanced structural geologic model. These complex three-dimensional models are generated completely by the code. A structural geologist can view the generated models and the associated error for each of the data types. These visualizations can help a geologist see alternative scenarios regarding the subsurface geology that are different than the previously conceived single model or few models but are still consistent with the different observed data.

It is possible to see trends of fault structures by plotting the fault traces of all the posterior models on top of the prior models. **Fig. 16** shows the fault lineaments of the posterior models in blue on top of the prior models in red. The prior models basically cover the whole space and do not provide any information regarding the location of faults in the subsurface. In the posterior lineaments, however, it is evident that there are several locations of north-northeast faults that are repeating as well as one or two repeating east-west faults. This can be seen more clearly in **Fig. 17**, which shows the per-cell faulting probability through

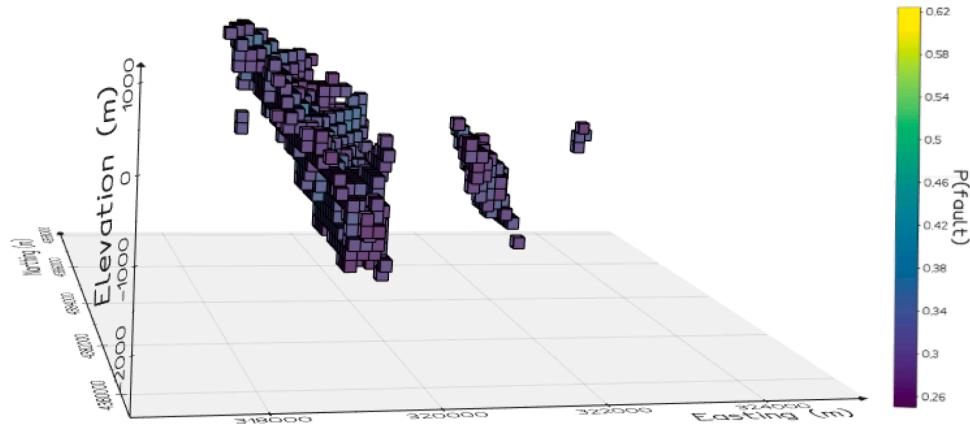


Fig. 19. Posterior fault probabilities, only showing blocks where the probability of a fault is greater than 0.25.

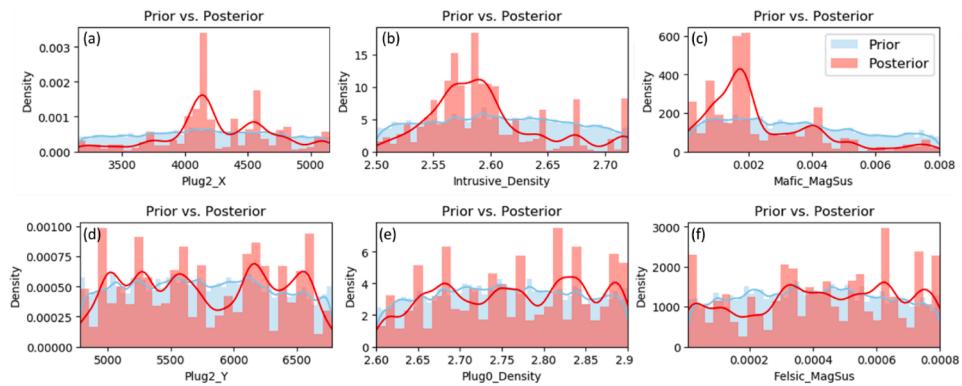


Fig. 20. Prior and posterior probabilities of different properties that are varied during the inversion. (a) The x position of the third intrusion. (b) The density of the granitic basement layer. (c) the magnetic susceptibility of the mafic layer. (d) The y position of plug number 2. (e) The density of the first intrusion. (f) The magnetic susceptibility of the felsic layer.

a section of the model. Fig. 18 shows similar line and probability maps of the prior and posterior for a constant y cross section (perpendicular to the strike of most faults) through the center of the model. While the prior was not very informative, we can see that the posterior shows high probability for several fault structures.

Three-dimensional plots can further illustrate the information obtained during the inversion. Fig. 19 shows voxels with a posterior fault probability above 0.25. In the prior model, the faults can be located everywhere in the modeling domain, and therefore the prior probability of a fault being located in any spot is never higher than 0.25. In the posterior probability cube, however, there are several visible north-northeast structures that have high probabilities. Some voxels have a probability as high as 0.62. These may indicate points in the model where observed data has shown that there are fault intersections and the higher probability at these points show that the posterior models respect the observed data.

4.1. Prior and posterior distribution of various properties

As part of the inversion, the values of many fault and lithology parameters were varied. By comparing the distribution of these properties in all the models (prior) versus the models that do match the data (posterior models), we can see what range of these properties better match the observations. For example, Fig. 20(a) shows the prior and posterior distribution of the x position of the third intrusion (plug2, indicated by the top left rectangle in Fig. 2) in light blue and light red, respectively. While the prior model is mostly uniform, the posterior location of the intrusion tends to be around 4100 m (in model space). Similarly, the plots in Fig. 20(b) and Fig. 20(c) show parameters whose posterior distribution better indicate their range of values. On the other hand, Fig. 20(d) shows the y position of the third intrusion (plug2) and this property shows both a uniform prior and a uniform posterior. The observed data did not reduce the uncertainty regarding this parameter. Similarly, Fig. 20(e) and Fig. 20(f) show the prior and posterior distribution of parameters that do not show uncertainty reduction following the inversion.

4.2. Challenges in matching the geophysical datasets

As we noted previously, the geophysical data, especially the gravity data did not converge substantially towards lower error, with a mean posterior gravity error of 2.14. When we ran the inversion only matching the gravity data, the convergence was much better, with errors commonly as low as 1.2 mGal as shown for example in the realization in Fig. 21(a). When we ran the inversion matching all data types except for the tracer data, the gravity data error reached as low as 1.38 mGal, see

Fig. 21(b). These observations show that as the inversion problem is currently set up, there are interactions between the data types that negatively impact convergence to low mismatch values across all data types.

Overall, we have several hypotheses to explain the challenge in matching the geophysical datasets:

- Incorrect choice of error metric: the cell-wise error calculation may not be the appropriate one since slight translation of the simulated data can cause large errors.
- Homogeneous density for each layer: the density of each layer in our models was constant, and we can see in the well logs that there is significant density variation between layers.
- Incorrect parameterization of stratigraphic layers: in our model, the stratigraphic layers are assumed to begin from a flat layer cake stratigraphy. Perhaps the layers have not been deposited or emplaced as flat layers or have been eroded or folded, leading to a more complex pre-faulting layering.
- Insufficient exploration of parameter space: each run has approximately 170 parameters. If we assume there are ten meaningful splits of each parameter, there are then a total of 10^{170} parameter combinations, while we investigated only around one million possibilities (10^6). We may have not searched the parameter space sufficiently.
- Varying convergence rates: the different data types have different sensitivities to small perturbations. The geophysical data sets change continuously with small changes in the parameter space. On the other hand, the tracer data, for example may change quite drastically with a small perturbation, as a small shift of the location of the faults may cause the faults to no longer be connected. This may lead the optimization to getting stuck in a local minimum dominated by the tracer data, and not sufficiently respecting the other data sets.

We encourage further research into search methods and geological parameterization strategies to address issues with convergence when multiple data types are present. Another avenue for improvement in the optimization results is to create a synthetic case study (or a set of case studies) with a known "truth" to assist in assessing the cause of the geophysical data convergence as well as assist in quantifying the success of different algorithms.

4.3. Method comparison

For each of the 5000 optimization jobs, we determined the lowest mismatch achieved during the optimization run. We collected the lowest mismatch values for each method (1250 values per method) and created

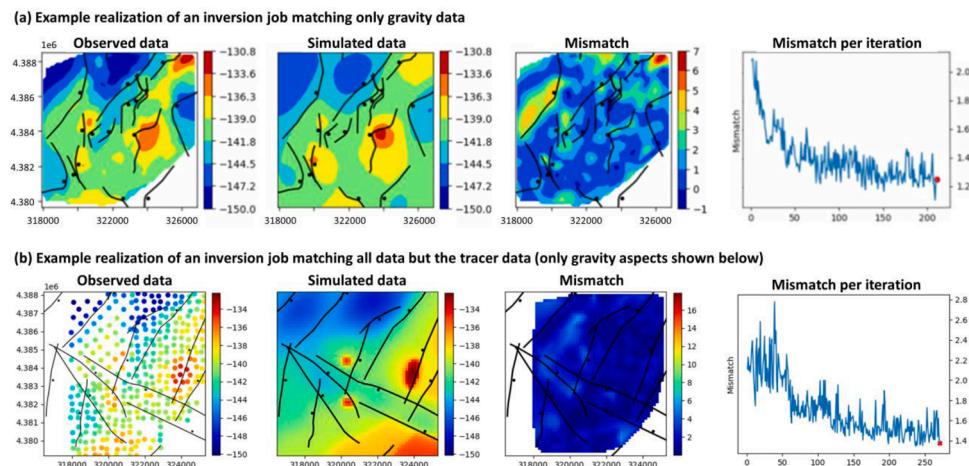


Fig. 21. Example realizations of an inversion job matching (a) only the gravity data and (b) all data types except the tracer data. The figure shows the observed and simulated gravity data (mGal), the mismatch between the observed and the simulated data, and the mismatch per iteration during the inversion.

the box plots in Fig. 22 showing the range of values, the 25th, 50th and 75th percentiles, and the outliers. The genetic algorithm and simulated annealing produced the models with the lowest mismatch values. MCMC had a lower median value than genetic annealing but didn't have the realizations with the lowest mismatch. Though we were expecting the best results from NSGA, the method performed the worst. Perhaps a newer version of the algorithm, NSGA III would provide better results (Deb and Jain, 2014).

4.4. Computational time

The inversion method proposed in this paper can be expensive computationally. Many evaluations are necessary per job and many jobs are necessary to converge to a satisfying posterior. Each evaluation on average takes 6–7 seconds, and each job has at least 150 model evaluations. The different jobs (threads), however, are completely parallel and therefore it is possible to run the stochastic inversion with many threads at once. The results shown in this paper, with 5000 jobs, were completed using a 96 CPU Google Cloud Computing resource within one and a half days.

5. Discussion

Through the inversion process, we were able to reduce the mismatch of models by half from a normalized error of 1 to an error of 0.5. While the prior models showed that faults could be located almost anywhere throughout the geological model, the posterior probability plots highlight areas with a higher likelihood of containing a fault.

Reservoir engineers can now use this resulting ensemble of posterior models, for example, to make a more robust decision on where to drill or stimulate the next well. Typically, engineers would run flow simulations based on one model of the subsurface and the results would be highly dependent on this single model. With this ensemble, engineers could run flow simulations multiple times for each of the posterior models for each of the possible locations. The distribution of the simulated energy output associated with drilling or stimulating in each location describes the financial risk due to geological uncertainty, and this risk information can aid in the decision-making process.

Models generated as part of the workflow presented in this paper do not need to be used directly in flow simulations, but rather can form a starting point for further refinement by geoscientists. Geoscientists may find it difficult to come up with many different three-dimensional scenarios regarding the subsurface, especially when one scenario has been found that satisfactorily matches the observed data, which can often lead to anchoring bias. Using the workflow and code presented in this paper can help suggest many alternative scenarios regarding the subsurface. These can then be further refined by a geoscientist.

This paper puts a strong focus on quantitative evaluation of model fit to observed data. When models are created by hand in a software, it is often difficult to measure the fit of the model to observed data. For example, calculating the match of simulated fault markers to observed fault markers may need to be done manually every time the model is perturbed. An important component of this work is to quantify and visualize the mismatch of multiple data sets in a more streamlined automated algorithm.

This work only suggested the fault locations and does not indicate whether the faults are permeable or not. The gravity, magnetic, and granite top data indicate the position of faults regardless of permeability. Future work could consider only using datasets that indicate permeable fault pathways: such as tracer data, fault intersection data, temperature logs, earthquake data, well tests, and production data.

Inverting using flow related data could be done in conjunction with heuristics that relate fault geometry and slip to fault permeability. For example, research suggests a relationship between fault slip and fault damage zone width, as well as between microfracture density with distance from the fault, and several types of fault zone permeability

structures (Faulkner et al., 2010). In addition, Barton et al. (1995) show that faults that are optimally oriented for failure in their stress field have higher permeability. In future work, fault orientations and fault slip in each geomodel realization could be used to sample more informed fault permeabilities using data and insight from the above cited research. The stochastic inversion workflow would be even more suited for geothermal energy flow simulations and decision making by including permeability as an inversion target.

5.1. Open-source code and benchmark data set

We have put all the data used for this inversion problem as well as the code on GitHub: <https://github.com/ahinoamp/PyNoddyInversion>. The GitHub also includes examples of 15 interactive three-dimensional models that had low mismatch errors that were generated as part of this work and a Binder Jupyter notebook from which readers can run an example code without installing anything. We encourage researchers to download the dataset used in this paper, improve the code with new inversion algorithms, or write a completely new code, and arrive at lower mismatches than those we presented in this paper.

6. Conclusions and future work

Geological uncertainty poses a major barrier for geothermal development. Uncertainty regarding the location of faults carrying hydrothermal fluids directly translates to financial risk and the viability of geothermal development. Stochastic inversion is a crucial tool for quantifying geological uncertainty and generating an ensemble of possible structural geologic models to use in making development decisions.

Traditional stochastic methods in the oil and gas industry, such as sequential Gaussian simulation and other geostatistical methods are not appropriate for the faulted setting of geothermal reservoirs. In this paper, we presented a workflow and code for performing an inversion to create geologically realistic structural models that match five typical geothermal data sets. We tested this workflow on data from the Patua Geothermal Field. We significantly reduced the prior uncertainty regarding fault locations and identified several locations with a high likelihood of containing a fault. We showed several of the resultant

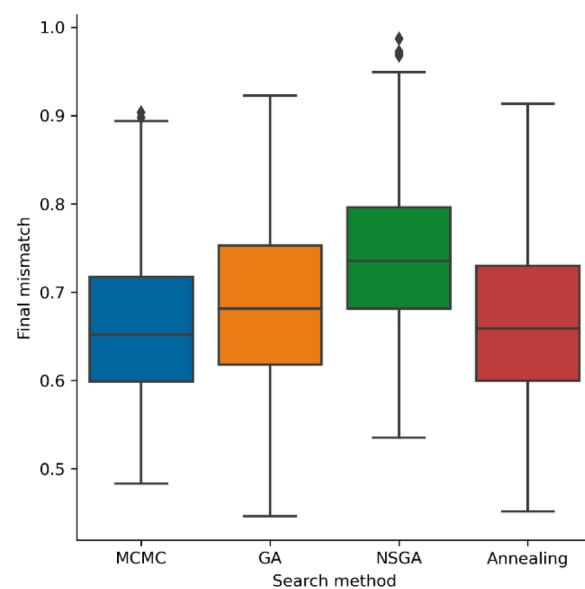


Fig. 22. Comparison of the four search methods tested in this paper via a box plot. The box outlines the median and the 25th and 75th percentile, the lines go to the minimum and maximum, and the diamonds represent outliers.

models and evaluated their geologic realism.

This work used five data types, yet it is possible to add several more data types to the inversion process, such as the temperature logs, the dips from both image logs and surface observations, density data from well logs, and production data. In addition, it is possible to set up the density and magnetic susceptibility distributions as spatially heterogeneous properties.

CRediT authorship contribution statement

Ahinoam Pollack: Conceptualization, Methodology, Software, Formal analysis, Data curation, Visualization, Writing - original draft. **Trenton T. Cladouhos:** Conceptualization, Methodology, Validation, Formal analysis, Data curation, Writing - review & editing. **Michael W. Swyer:** Conceptualization, Methodology, Validation, Formal analysis, Data curation, Writing - review & editing. **Drew Siler:** Conceptualization, Methodology, Validation, Formal analysis, Writing - review & editing. **Tapan Mukerji:** Conceptualization, Methodology, Validation, Formal analysis, Writing - review & editing, Supervision. **Roland N. Horne:** Conceptualization, Methodology, Validation, Formal analysis,

Writing - review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This material is based upon work supported by the National Science Foundation (NSF) Graduate Research Fellowship Program under Grant No. 1656518, as well as sponsors of Stanford Center for Earth Resources Forecasting (SCERF), and funding from Prof. S. Graham, Dean, Stanford School of Earth, Energy and Environmental Sciences. We thank Google Cloud Platform Education Grants for providing research computation credits (EDU Credit kcapozzi 35804832). We thank Prof. Jim Faulds for advice on local Patua geology and data analysis. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Appendix A. Algorithm details

This appendix provides more details on the four algorithms used in this paper.

A.1 Markov Chain Monte Carlo with a Metropolis sampler

Bayesian methods are often used to find the posterior distribution of an event given some prior knowledge. In our case, this means finding the ensemble of accepted structural models given the prior model regarding the geology. Bayes' equation for finding the posterior is formulated as follows:

$$p(m|d) \cong L(d|m)p(m) \quad (5)$$

where m is the structural model, $p(m)$ is the prior probability of a structural model, and $L(d|m)$ is the likelihood of a problem, which in our problem is proportional to the inverse of the misfit. When the likelihood or prior functions are complex and it is impossible to solve the Bayesian equation analytically, Markov Chain Monte Carlo methods are a way to arrive at the posterior in an iterative process. We use the Markov Chain Monte Carlo method with a Metropolis Sampler (Aydin and Caers, 2017; Mosegaard and Tarantola, 1995), which is shown in the pseudo-algorithm in Table 3.

Table 3

Pseudo-algorithm for Markov Chain Monte Carlo with Metropolis Sampler.

m_0	Initialization of the structural model by selecting a subset of faults from the fault-traces bank and randomly selecting fault properties within the prior ranges
for $iter = 1$ to n_{iter} do	Iterate n_{iter} times
$m^* \sim p(m)$	Propose a new step
$u_{acc} \sim U(0, 1)$	Draw acceptance probability
$\alpha_{Metropolis} = \min\left(1, \frac{L(d m^*)}{L(d m)}\right) = \min\left(1, \exp\left(-\frac{\text{misfit}_{new} - \text{misfit}_{old}}{\text{normalizing factor}}\right)\right)$	Compute transition probability by calculating the combined misfit of all the data types
if ($u_{acc} \leq \alpha_{Metropolis}$)	Accept proposed step if random number is less than acceptance probability
$m = m^*$	
else	
$m = m$	Reject proposed step

A.2 Simulated annealing

The simulated annealing algorithm (Laarhoven and Aarts, 1987) is very similar to the MCMC algorithm, with a single difference of substituting an exponentially diminishing parameter instead of the normalization factor, which in simulated annealing literature is referred to as "temperature." In the beginning of the optimization, the temperature is large, and this leads to higher acceptance probabilities. Even if a new perturbation in the optimization chain does not improve upon the previous step, there is a substantial chance that it will be accepted when the temperature value is large. A high temperature value in the beginning of the optimization encourages exploration. As the optimization progresses, the temperature decreases, and it is less likely that the algorithm accepts proposals that do not improve the results. This setup of moving from exploration to exploitation of the search space has been found to improve results. A simplified version of the algorithm is shown in Table 4.

Table 4

Pseudo-algorithm for simulated annealing method.

m_0	Initialization of the structural model by selecting a subset of faults from the fault-traces bank and randomly selecting fault properties within the prior ranges
for $\text{iter} = 1$ to n_{iter} do	Iterate n_{iter} times
$m^* \sim p(m)$	Propose a new step
$u_{\text{acc}} \sim U(0, 1)$	Draw acceptance probability
$T = T_0 * \text{ReductionRate}^{\text{iter}}$	Assign the temperature, T
$\alpha_{\text{Metropolis}} = \min\left(1, \frac{L(d m^*)}{L(d m)}\right) = \min\left(1, e^{-\frac{\text{misfit}_{\text{new}} - \text{misfit}_{\text{old}}}{T}}\right)$	Compute transition probability by calculating the combined misfit of all the data types
if ($u_{\text{acc}} \leq \alpha_{\text{Metropolis}}$)	Accept proposed step if random number is less than acceptance probability
$m = m^*$	
else	
$m = m$	Reject proposed step

A.3 Simple genetic algorithm

MCMC and simulated annealing continuously improve a single structural model by repeated perturbation and evaluation of the structural model. Genetic algorithms (Back et al., 2000; Holland, 1962) emphasize an optimization process using a population of “individuals” (sets of structural model parameters) that share information regarding the best parameters. Genetic algorithms contain three main steps: selection, crossover, and mutation. In the selection step, structural models that have low misfit are preferentially selected to be “parents.” In the crossover step, the chosen “parent” structural models combine their parameters to form “offspring.” In the mutation step, some random noise is added to the process by randomly perturbing some of the parameters of some of the models. The basic genetic algorithm we implemented is shown in Table 5. We modified the python package DEAP: Distributed Evolutionary Algorithms in Python (Fortin et al., 2012) to fit the problem of minimizing the misfit for structural models.

Table 5

Pseudo-algorithm for simple genetic algorithm.

initialize population, pop	Initialize n_{pop} structural models
fitness = evaluate(pop)	Evaluating the misfit of the initial model population
for $\text{iter} = 1$ to n_{iter}	Iterate n_{iter} times
parents = select(pop, fitness)	Select the best models for reproduction
offspring = crossover(parents)	Combine parameters from the best models to form offspring models
offspring = mutate(offspring)	Randomly select some parameters of some models and perturb them
pop = offspring	The offspring become the current population
fitness = evaluate(pop)	Evaluate the mist of the current model population

A.4 Nondominated sorting genetic algorithm (NSGA II)

The algorithm for NSGA II (Deb et al., 2002) is similar to the basic genetic algorithm, except for modifying the selection process to better address problems that have multiple objectives. For example, the inversion problem discussed in this paper has five objectives: minimizing the misfit for gravity, magnetic, tracer, wellbore granite top markers, and wellbore fault markers. In the basic genetic algorithm, the selection algorithm is based solely on the combined normalized misfit for all the objectives. For example, in the “tournament selection” option, $n_{\text{tournament}}$ individual models are chosen from the population and the model with the lowest misfit amongst those in the tournament is appended to the list of parents. However, when there are multiple objectives, many problems arise:

- 1 Normalization: to combine multiple objectives, it is necessary to normalize them before combining. Yet, there is no clear-cut method to combine different objectives into a single misfit. For example, the gravity data misfit and the tracer data misfit have different ranges of misfit values and it is not clear whether these data types should be combined by either normalizing each misfit by the maximum range of the misfit, or by the mean of the misfits calculated during optimization, or by any other normalization criteria.
- 2 Convergence: when multiple objectives are combined, if one of the objectives converges faster than others, the algorithm may quickly end in a local minimum where only the rapidly converging objective is respected.
- 3 Diversity of solutions: solutions may tend to favor only certain objectives and not show models that perform better on other objectives.

Table 6

Pseudo-algorithm of the NSGA selection process.

<i>Selection(pop, multiobjective fitness)</i>	The NSGA II selection algorithm
<i>sort models by dominance per front, sort by crowding distance choose n_{pop} models with highest ranking</i>	Evaluate the dominance of each model and sort the models by their dominance within each dominance front, sort by crowding distance Select the top models from the sorted list of models with the highest dominance ranking

NSGA addresses these issues by selecting models based on two criteria: dominance and crowding distance. The dominance criterion determines which models dominate other models based on the different objectives. A model that is better than all other models in all objectives is assigned to the first “front,” and will be included as a parent in the next generation. Models that are better than all other models in all the objectives, except one, may be in the second “front,” and so on. These fronts form the dominance score. In the algorithm, models are sorted by this dominance (or “nondominated”) score, and this score replaces the typical misfit criteria.

If models have the same dominance score, then a “crowding distance” is used to determine which solution is better. A crowding distance measures the minimum distance between one structural model and all other models in the objective space. This metric encourages models that are more diverse and in less “crowded” areas.

The basic selection process with the NSGA II algorithm is shown in [Table 6](#). This algorithm solves many challenges with multiobjective optimization algorithms. First, it is not necessary to formulate a single misfit for the models. Models are compared to each other in terms of dominance, by comparing the misfit for each objective between different models and determining only whether one model is better than the other on a particular objective, not mattering by how much. The algorithm encourages selection of models whose performance is better in some of the objectives and thus improves both the convergence, since the algorithm is less likely to fall into local minima controlled by one of the objectives and has a higher diversity of solutions.

Appendix B. Prior falsification

Before starting the inversion, we checked whether these prior parameter ranges lead to simulated data that is in the ballpark (statistically) of the observed data. We generated over 1500 prior models using the ranges of uncertainty defined in the previous sections. We then simulated the gravity, magnetic and granite-top data for each of these models. To compare the observed data to the simulated data of the prior models, we transformed the simulated and observed data to a lower dimensional space. We then checked whether the observed data was sufficiently close to the simulated prior data. This procedure is illustrated in [Fig. 23](#) and explained in the work of [Yin et al. \(2020\)](#).

The results of this prior falsification analysis show that the observed data lies sufficiently close to the simulated data of the prior model realizations. [Fig. 24\(a\)](#) shows the observed data and the simulated data in a reduced dimension space. As can be seen, the observed data is on the edge but not an outlier of the simulated data, and several models have data that are close to the observed data. [Fig. 24\(b\)](#) shows the distance of each model from the center of the models. The red line indicated the 97.5 percentile of the distribution of distances from the center of the models. The observed data is below this line, indicating that the observed data is not an outlier, and the prior model definition is sufficiently accurately defined and there is a high chance of finding an ensemble of models that match the observed data.

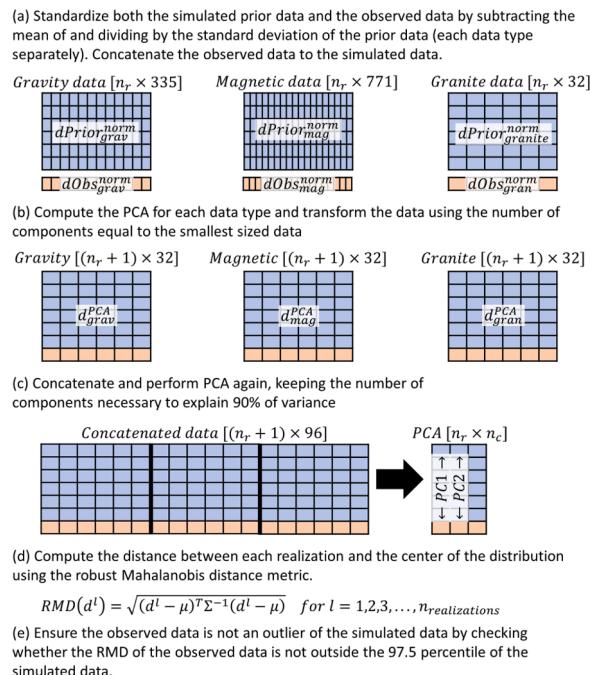


Fig. 23. Steps to perform prior falsification, where n_r is the number of simulated prior realizations, d_{Prior} is the simulated data of the prior model realizations, RMD is the Robust Mahalanobis distance for each realization, μ and σ are the robust estimations of the mean and covariance of the data.

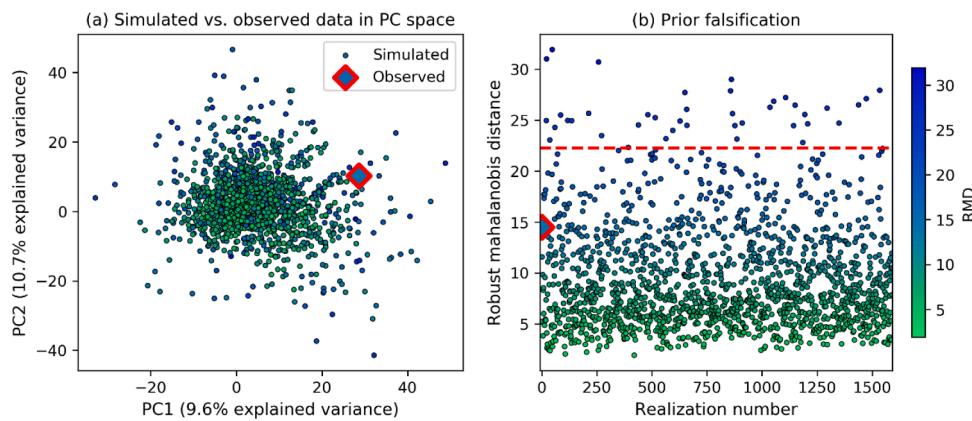


Fig. 24. (a) The location of the observed data relative to the simulated data in a reduced dimensional space, containing the first two principal components of the data. (b) The Robust Mahalanobis distance of the simulated data for the realizations versus the observed data. The red line indicates the 97.5 percentile of the distribution of distances.

References

- Anderson, E.M., 1951. The Dynamics of Faulting and Dyke Formation with Applications to Britain. Oliver and Boyd, Edinburgh.
- Aydin, O., Caers, J.K., 2017. Quantifying structural uncertainty on fault networks using a marked point process within a Bayesian framework. *Tectonophysics* 712–713, 101–124. <https://doi.org/10.1016/j.tecto.2017.04.027>.
- Back, T., Fogel, D.B., Michalewicz, Z., 2000. *Evolutionary Computation 1: Basic Algorithms and Operators*.
- Barton, C.A., Zoback, M.D., Moos, D., 1995. Fluid flow along potentially active faults in crystalline rock. *Geology* 23, 683–686 [https://doi.org/10.1130/0091-7613\(1995\)023<683:FFAPAF>2.3.CO;2](https://doi.org/10.1130/0091-7613(1995)023<683:FFAPAF>2.3.CO;2).
- Bonneau, F., Henrion, V., Caumon, G., Renard, P., Sausse, J., 2013. A methodology for pseudo-genetic stochastic modeling of discrete fracture networks. *Comput. Geosci.* 56, 12–22. <https://doi.org/10.1016/j.cageo.2013.02.004>.
- Bosch, M., Meza, R., Jiménez, R., Höning, A., 2006. Joint gravity and magnetic inversion in 3D using Monte Carlo methods. *Geophysics* 71, 3–6. <https://doi.org/10.1190/1.2209952>.
- Chen, J., Hoversten, G.M., Key, K., Nordquist, G., Cumming, W., 2012. Stochastic inversion of magnetotelluric data using a sharp boundary parameterization and application to a geothermal site. *Geophysics* 77. <https://doi.org/10.1190/geo2011-0430.1>.
- Chen, M., Tompson, A.F.B., Mellors, R.J., Ramirez, A.L., Dyer, K.M., Yang, X., Wagoner, J.L., 2014. An efficient Bayesian inversion of a geothermal prospect using a multivariate adaptive regression spline method. *Appl. Energy* 136, 619–627. <https://doi.org/10.1016/j.apenergy.2014.09.063>.
- Cherpeau, N., Caumon, G., 2015. Stochastic structural modelling in sparse data situations. *Pet. Geosci.* 21, 233–247. <https://doi.org/10.1144/petgeo2013-030>.
- Cherpeau, N., Caumon, G., Caers, J., Lévy, B., 2012. Method for stochastic inverse modeling of fault geometry and connectivity using flow data. *Math. Geosci.* 44, 147–168. <https://doi.org/10.1007/s11004-012-9389-2>.
- Gladouhous, T.T., Uddenberg, M.W.W., Swyer, M.W., Nordin, Y., Garrison, G.H., 2017. Patux Geothermal Geologic Conceptual Model 1057–1075.
- Clark, D.A., Emerson, D.W., 1991. Notes on rock magnetization characteristics in applied geophysical studies. *Explor. Geophys.* <https://doi.org/10.1071/EG991547>.
- Cui, T., Fox, C., Sullivan, M.J.O., O'Sullivan, M.J., 2011. Bayesian calibration of a large-scale geothermal reservoir model by a new adaptive delayed acceptance Metropolis Hastings algorithm. *Water Resour. Res.* 47 <https://doi.org/10.1029/2010WR010352>.
- De La Varga, M., Schaaf, A., Wellmann, F., 2019. GemPy 1.0: Open-source stochastic geological modeling and inversion. *Geosci. Model Dev.* 12, 1–32. <https://doi.org/10.5194/gmd-12-1-2019>.
- de Pasquale, G., Linde, N., Doetsch, J., Holbrook, W.S., 2019a. Probabilistic inference of subsurface heterogeneity and interface geometry using geophysical data. *Geophys. J. Int.* 217, 816–831. <https://doi.org/10.1093/gji/ggz055>.
- de Pasquale, G., Linde, N., Greenwood, A., 2019b. Joint probabilistic inversion of DC resistivity and seismic refraction data applied to bedrock/regolith interface delineation. *J. Appl. Geophys.* 170, 103839 <https://doi.org/10.1016/j.jappgeo.2019.103839>.
- Deb, K., Jain, H., 2014. An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, Part I: Solving problems with box constraints. *IEEE Trans. Evol. Comput.* 18, 577–601. <https://doi.org/10.1109/TEVC.2013.2281535>.
- Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. Evol. Comput.* 6, 182–197. <https://doi.org/10.1109/4235.996017>.
- Dijkstra, E.W., 1959. A Note on Two Problems in Connexion with Graphs. *Numer. Math.* 1, 269–271.
- Farrell, S.M., Jessell, M.W., Barr, T.D., 1996. Inversion of geological and geophysical data sets using genetic algorithms. 1996 SEG Annu. Meet. 1404–1406. <https://doi.org/10.1190/1.1826374>.
- Faulds, J.E., Coolbaugh, M.F., Benoit, D., Oppiger, G., Perkins, M., Moeck, I., Drakos, P., 2010. Structural Controls of Geothermal Activity in the Northern Hot Springs Mountains, Western Nevada: The Tale of Three Geothermal Systems (Brady's, Desert Peak, and Desert Queen). *GRC Transactions*.
- Faulds, J.E., Henry, C.D., 2008. Tectonic influences on the spatial and temporal evolution of the Walker Lane : An incipient transform fault along the evolving Pacific – North American plate boundary. *Ores Orogenes. Circum- Pacific tectonics. Geol. Evol. ore Depos.* Arizona Ge 437–470.
- Faulds, J.E., Ramelli, A.R., Perkins, M.E., 2011. Preliminary Geologic Map of the Hazen Quadrangle, Lyon and Churchill Counties, Nevada.
- Faulkner, D.R., Jackson, C.A.L., Lunn, R.J., Schlische, R.W., Shipton, Z.K., Wibberley, C. A.J., Withjack, M.O., 2010. A review of recent developments concerning the structure, mechanics and fluid flow properties of fault zones. *J. Struct. Geol.* 32, 1557–1575. <https://doi.org/10.1016/j.jsg.2010.06.009>.
- Finch, E., Gawthorpe, R., 2017. Growth and interaction of normal faults and fault network evolution in rifts: insights from three-dimensional discrete element modelling. *Geol. Soc. Spec. Publ.* 439, 219–248. <https://doi.org/10.1144/SP439.23>.
- Fortin, F.A., De Rainville, F.M., Gardner, M.A., Parizeau, M., Gagné, C., 2012. DEAP: evolutionary algorithms made easy. *J. Mach. Learn. Res.* 13, 2171–2175.
- Fullagar, P.K., Pears, G.A., McMonnies, B., 2008. Constrained inversion of geologic surfaces - pushing the boundaries. *Lead. Edge (Tulsa, OK)* 27, 98–105. <https://doi.org/10.1190/1.2831686>.
- Gehringer, M., Loksha, V., 2012. *Handbook on Planning and Financing Geothermal Power Generation. Main findings and recommendations*, Washington, D.C.
- Holland, J.H., 1962. Outline for a logical theory of adaptive systems. *J. ACM* 9, 297–314. <https://doi.org/10.1145/321127.321128>.
- Jardani, A., Revil, A., 2009. Stochastic joint inversion of temperature and self-potential data. *Geophys. J. Int.* 179, 640–654. <https://doi.org/10.1111/j.1365-246X.2009.04295.x>.
- Jessell, M.W., Valenta, R.K., 1996. Structural geophysics: Integrated structural and geophysical modelling. *Comput. Methods Geosci.* 15, 303–324. [https://doi.org/10.1016/S1874-561X\(96\)80027-7](https://doi.org/10.1016/S1874-561X(96)80027-7).
- Jolie, E., Moeck, I., Faulds, J.E., 2015. Quantitative structural-geological exploration of fault-controlled geothermal systems-a case study from the Basin-and-Range Province, Nevada (USA). *Geothermics* 54, 54–67. <https://doi.org/10.1016/j.geothermics.2014.10.003>.
- Kim, Y.S., Sanderson, D.J., 2005. The relationship between displacement and length of faults: a review. *Earth-Sci. Rev.* 68, 317–334. <https://doi.org/10.1016/j.earscirev.2004.06.003>.
- Laarhoven, P.J.M. Van, Aarts, E.H.L., 1987. *Simulated Annealing: Theory and Applications*. Springer, Dordrecht.
- Mellors, R.J., Tompson, A., Dyer, K., Yang, X., Chen, M., Wagoner, J., 2014. Stochastic Joint inversion modeling algorithm of geothermal prospects. In: *Proceedings, Thirty-Ninth World Geotherm. Reserv. Eng.*, pp. 1–7.
- Mosegaard, K., Tarantola, A., 1995. Monte Carlo sampling of solutions to inverse problems. *J. Geophys. Res.* 100 <https://doi.org/10.1029/94jb03097>.
- Musy, M., 2021. Vedo, a python module for scientific analysis and visualization of 3D objects and point clouds. Zenodo. <https://doi.org/10.5281>.
- Pachauri, R.K., Allen, M.R., Barros, V.R., Broome, J., Cramer, W., Christ, R., Church, J.A., Clarke, L., Dahe, Q., Dasgupta, P., Dubash, N.K., Edenhofer, O., Elgizouli, I., Field, C. B., Forster, P., Friedlingstein, P., Fuglestvedt, J., Gomez-Echeverri, L., Hallegatte, S., Hegerl, G., Howden, M., Jiang, K., Cisneros, B.J., Kattsov, V., Lee, H., Mach, K.J., Marotzke, J., Mastrandrea, M.D., Meyer, L., Minx, J., Mulugetta, Y., O'Brien, K., Oppenheimer, M., Pereira, J.J., Pichs-Madruga, R., Plattner, G.-K., Pörtner, H.-O., Power, S.B., Preston, B., Ravindranath, N.H., Reisinger, A., Riahi, K., Rusticucci, M., Scholes, R., Seyboth, K., Sokona, Y., Stavins, R., Stocker, T.F., Tschakert, P.,

- Vuuren, D.van, Ypersele, J.-P.van, 2014. Climate change 2014: synthesis report: summary for policymakers. contrib. work. groups I, II III to Fifth Assess. Rep. Intergov. Panel Clim. Chang. 31 <https://doi.org/10.1017/CBO9781107415324>.
- Pollack, A., Cladouhos, T.T., Swyer, M., Horne, R., Mukerji, T., 2020. Stochastic Structural Modeling of a Geothermal Field: Patua Geothermal Field Case Study. 45th Workshop on Geothermal Reservoir Engineering, Stanford.
- Inc., PRJ, Geophysics, Wave, 2005. Interpretation of Ground Gravity and Airborne Magnetic Data. Hazen KGRA Project, Nevada, U.S.A.
- Salmon, J.P., Meurice, J., Wobus, N., Stern, F., Duame, M., 2012. Guidebook to geothermal power finance. *Geotherm. Power Financ. Guid. Policy Options 1–61*.
- Scheidt, C., Li, L., Caers, J., 2018. Quantifying Uncertainty in Subsurface Systems. American Geophysical Union and John Wiley and Sons, Inc., Hoboken.
- Schon, J.H., 2011. Physical Properties of Rocks, in: Handbook of Petroleum Exploration and Production. Elsevier, Amsterdam, p. 781588.
- Siler, D.L., Hinz, N.H., Faulds, J.E., Queen, J., 2016. 3D analysis of geothermal fluid flow favorability: Brady's, Nevada, USA. 41st Work. Geotherm. Reserv. Eng. Stanford Univ. SGP-TR-209.
- Speer, B., Economy, R., Lowder, T., Schwabe, P., Regenthal, S., 2014. Geothermal exploration policy mechanisms: Lessons for the united states from international applications. *Geotherm. Explor. Dev. Policy Int. Lessons 1–73*.
- Suzuki, S., Caumon, G., Caers, J., 2008. Dynamic data integration for structural modeling: model screening approach using a distance-based model parameterization. *Comput. Geosci. 12*, 105–119.
- Thiele, S.T., Jessell, M.W., Lindsay, M., Ogarko, V., Wellmann, J.F., Pakyuz-Charrier, E., 2016. The topology of geology 1: Topological analysis. *J. Struct. Geol. 91*, 27–38. <https://doi.org/10.1016/j.jsg.2016.08.009>.
- U.S. Energy Information Administration, 2016. September 2016 Monthly Energy Review, Monthly Energy Review. [https://doi.org/DOE/EIA-0035\(2011/02\)](https://doi.org/DOE/EIA-0035(2011/02)).
- U.S. Geological Survey, 2006. Quaternary fault and fold database for the United States [WWW Document]. URL <https://earthquake.usgs.gov/hazards/qfaults/> (accessed 10.28.19).
- Vogt, Christian, Kosack, C., Marquart, G., 2012a. Stochastic inversion of the tracer experiment of the enhanced geothermal system demonstration reservoir in Soultz-sous-Forêts - revealing pathways and estimating permeability distribution. *Geothermics 42*, 1–12. <https://doi.org/10.1016/j.geothermics.2011.11.001>.
- Vogt, C., Marquart, G., Kosack, C., Wolf, A., Clauser, C., 2012b. Estimating the permeability distribution and its uncertainty at the EGS demonstration reservoir Soultz-sous-Forêts using the ensemble Kalman filter. *Water Resour. Res. 48*, 1–15. <https://doi.org/10.1029/2011WR011673>.
- Wellmann, J.F., De La Varga, M., Murdie, R.E., Gessner, K., Jessell, M., 2018. Uncertainty estimation for a geological model of the Sandstone greenstone belt, Western Australia - insights from integrated geological and geophysical inversion in a Bayesian inference framework. *Geol. Soc. Spec. Publ. 453*, 41–56. <https://doi.org/10.1144/SP453.12>.
- Wellmann, J.F., Thiele, S.T., Lindsay, M.D., Jessell, M.W., 2016. Pynoddy 1.0: An experimental platform for automated 3-D kinematic and potential field modelling. *Geosci. Model Dev. 9*, 1019–1035. <https://doi.org/10.5194/gmd-9-1019-2016>.
- Wells, D.L., Coppersmith, K.J., 1994. New empirical relationships among magnitude, rupture length, rupture width, rupture area, and surface displacement. *Bull. Seismol. Soc. Am. 84*, 974–1002.
- Williams, C.F., Reed, M.J., Mariner, R.H., DeAngelo, J., Galanis, S.P.J., 2008. Assessment of moderate-and high-temperature geothermal resources of the United States: U.S. geological survey fact sheet 2008-3082. U.S. Geol. Surv. Fact Sheet, 2008-3082 1–4.
- Witter, J.B., Siler, D.L., Faulds, J.E., Hinz, N.H., 2016. 3D geophysical inversion modeling of gravity data to test the 3D geologic model of the Brady's geothermal area. Nevada, USA. *Geotherm. Energy 4*. <https://doi.org/10.1186/s40517-016-0056-6>.
- Yin, Z., Strelbel, S., Caers, J., 2020. Automated Monte Carlo-based quantification and updating of geological uncertainty with borehole data (AutoBEL v1.0). *Geosci. Model Dev. 13*, 651–672. <https://doi.org/10.5194/gmd-13-651-2020>.