PROCEEDINGS, 50th Workshop on Geothermal Reservoir Engineering
Stanford University, Stanford, California, February 10-12, 2025
SGP-TR-229

Quantile Model Regression of Nameplate Capacities: P50 and P90 Megawatt Hydrothermal Projection

Whitney Trainor-Guitton, Karthik Menon, Paul Pinchuk, Sophie Min Thomson, Nicole Wagoner, Chao Lu, Eli Mlawsky, Cary Lindsey, Mark Coolbaugh, and James Faulds

15013 Denver West Parkway, Golden, CO 80401

Whitney.TrainorGuitton@nrel.gov

Keywords: hidden hydrothermal technical potential, regression, technical and economic assessments (TEA)

ABSTRACT

This paper provides data-driven estimates of hidden hydrothermal technical capacity of the Great Basin. Geothermal (hydrothermal and enhanced geothermal systems) has recently been included in large-scale renewable energy generation models, with the same resolution as wind and solar. Previous estimates for hydrothermal did not include measurements of permeability or uncertainty. The Great Basin area provides the most data-rich location to identify positive and negative hydrothermal sites and data observables (features), which have been used in eXtreme Gradient Boosting (XGBoost) regression to make hydrothermal capacity predictions. This work is novel in its use of 37 nameplate capacities (megawatts) as the positive labels and 248 negative (0-MW) locations. The independent variables or features are 48 geophysical and geologic attributes. Three XGBoost predictions were input into the Renewable Energy Potential (reV) model. These predictions were chosen because of their fits to the 37 geothermal plants and negative locations as well as their reasonable prediction ranges and totals within the INnovative Geothermal Exploration through Novel Investigations Of Undiscovered Systems (INGENIOUS) study area. We present a method for subsampling the negative sites to bring the labels into balance that uses the geologic domain knowledge to proportionally represent negatives. Overall, the distributions of the hydrothermal technical capacity and the site levelized cost of energy are respectively much tighter and lower than the previous estimates for the Great Basin, which only used contiguous U.S. scaled temperature estimates. Percentile (50th and 90th, median and high estimate, respectively) models provide bookends for these metrics.

1. INTRODUCTION

Important advancements in geothermal research and next-generation technologies have greatly expanded geothermal development potential. Historically, hydrothermal resources were discovered only proximal to surficial evidence of a system, such as fumaroles or hot springs, which gave certainty about naturally flowing fluids to bring heat to the surface. Hidden systems, where the permeability and fluids are not apparent at the surface, have been the focus of recent research efforts. Projects focusing on play fairway analysis and machine learning, have resulted in several recent large data collection efforts (INnovative Geothermal Exploration through Novel Investigations Of Undiscovered Systems [INGENIOUS](Ayling et al., 2022), Basin and Range Investigations for Developing GeothermalEnergy [BRIDGE] (Schwering et al., 2022) and better understandings of which data layers are most effective at discovering hidden resources (Craig et al., 2021; Faulds et al., 2015; Faulds et al., 2017).

The potential for substantial growth in hidden geothermal and next-generation geothermal technologies has compelled better representation of geothermal in contiguous U.S. renewable energy generation models. The potential future deployment of renewable energy generation requires accurate representation of potential power production while also considering siting constraints, environmental and ecological impacts, and social acceptance from local communities. This challenge at a nationwide scale has been consistently addressed by the National Renewable Energy Laboratory (NREL) Renewable Energy Potential (reV) model (Lopez et al., 2023; Maclaurin et al., 2021); however, historically, this model has only included wind and solar resource potential and their respective technical and economic assessments (TEA).

The relatively new reV geothermal module allows for several data types to be input as geothermal resource potential, including discrete points of geothermal plant performance and spatially continuous (gridded) depth-discrete temperatures (Pinchuk et al., 2023). The reV geothermal module calls the Geothermal Electricity TechnoEconomic Model (GETEM) to estimate capital and operating costs to produce geothermal electricity (Mines, 2016), and reV geothermal also allows for detailed assessments of the geospatial intersection of the estimated geothermal resource with grid infrastructure and land use characteristics. The geothermal resource estimates arguably have more uncertainty than the solar and wind estimates. Geothermal has fewer data observations because the resource is below the surface. Each well drilled to observe the subsurface requires a significant investment, and the cost increases further when downhole surveying and well flow tests are done. Geothermal resource assessment uncertainty should be included when considering the resource and TEA in reV.

reV requires geothermal power plant megawatt potential to be input. Although play fairway analysis can provide expert-driven assessments of the permeability, heat, and fluids, “favorability” cannot be used in TEA calculations. Temperature regressions also require an extra step of determining how thermal energy can be converted to megawatts. Previous reV geothermal studies have used two published methods that provide widely differing estimates (Augustine, 2017, 2016; Wilmarth et al., 2021). As shown in (Pinchuk et al., 2023) and in Figure 1, the exponential method utilizing contiguous U.S. temperatures from Blackwell et al. (2011), in general results in estimates that are lower than the nameplate capacities (x-axis). More than 70% of the collocated estimates are less than the installed capacities of geothermal plants.

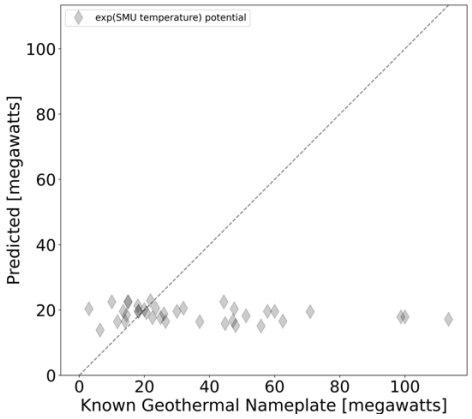

Figure 1: The x-axis value represents the nameplate capacity of 37 geothermal plants within the Great Basin, and the y-value is the colocated estimate derived from only temperature measurements (Pinchuk et al. 2023).

But it is also of concern that there are no "zero" capacity estimations with the exponential method. This is not realistic when estimating hydrothermal because the exponential method does not include an evaluation of permeability and/or fluids. The hydrothermal power potential must include some assessment of permeability, as unlike enhanced geothermal systems (EGS), we are not assuming permeability will be stimulated in the reservoir.

A distinguishing aspect of this work is the presence of independent (e.g., features) and dependent (e.g., labels) variables. The training labels associated with positive and negative hydrothermal geothermal resources are chosen with the end goal of inputting hydrothermal potential into reV; thus, the positive hydrothermal labels are represented by the 37 operating nameplate capacities (megawatts) within the INGENIOUS study area (x-values in Figure 1). Equivalently, the negative sites should be from a "dry" or no-flow well database, but this is not publicly available. Without this information, locations where no hydrothermal systems is expected are determined using estimates absence of convective zones (DeAngelo et al., 2022). We use the data collected by INGENIOUS, which include 48 geologic, geophysical, and geologically derived datasets as independent variables or features (Ayling et al., 2022). The observations include parameters that provide proxies for permeability across the Great Basin at a spatial resolution of 250 meters by 250 meters. The previous reV hydrothermal estimates did not include any indication of permeability and relied only on temperature estimates.

To address the high degree of uncertainty of hydrothermal resource locations and to obtain an achievable representation of regional permeability, we present a methodology and results for representing the hydrothermal resource uncertainty through additional resource variables and statistical modeling to quantify and bound the uncertainty range of the resource potential **with the end goal to use the megawatt estimates within reV**. In this paper, we present a quantile regression using eXtreme Gradient Boosting (XGBoost) to estimate the 50^{th} and 90^{th} percentile megawatt hydrothermal potential across the Great Basin.

The following sections describe the methodology and explorations of how a standard regression and 50^{th} and 90^{th} percentile (P50, P90) potential (megawatt) models for the hydrothermal resource in the Great Basin were obtained. First, in Section 2 (Methodology), we briefly describe how the negative sites were defined from the heat flow residuals and how the two XGBoost regressions that were performed. Next, in Section 3, we describe the final TEA results, which include the technical potential and the levelized cost of energy (LCOE) metrics. Additionally, we demonstrate the sensitivity of the chosen negative sites.

2. METHODOLOGY

This methodology provides data-driven estimates of hidden hydrothermal capacity of the Great Basin. This is possible thanks to the data collection and interpretation efforts of the INGENIOUS team. Additionally, the use of 37 operating geothermal power plants and the development of the reV geothermal module allows for spatially expansive calculations of the potential of hidden geothermal systems, specifically power potential and LCOE. The overall workflow of how the data are turned into TEA matrics is shown in Figure 2.

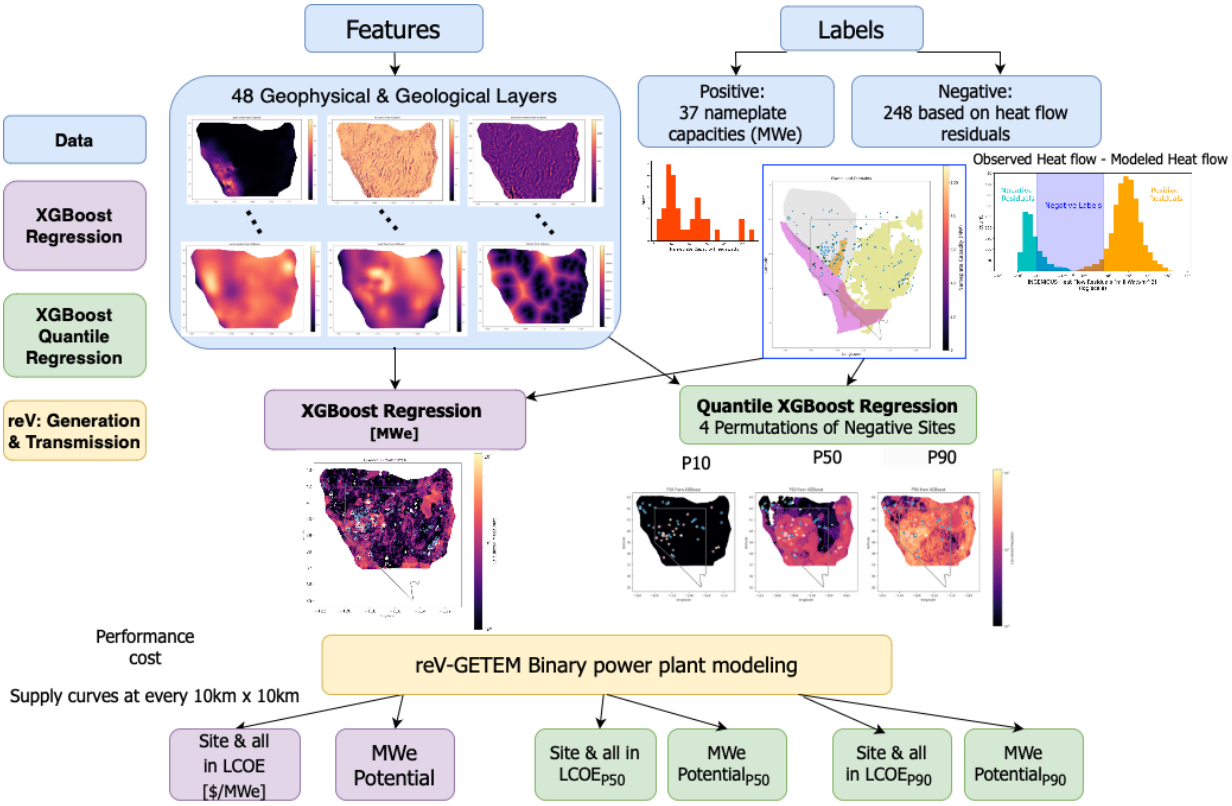

Figure 2: Workflow of methodology.

By coordinating with the INGENIOUS project, we identified negative sites to represent appropriate places as "0 nameplate capacity" for training the XGBoost model. These are critical in identifying and realistically representing the hydrothermal potential in reV generation and TEA modeling.

2.1 Training Sites

We derive the negative sites from the heat flow residuals. These heat residuals were calculated by the U.S. Geological Survey (USGS), and they represent the difference between observed heat flow and calculated regional heat flow at each of the wells used in their analysis (DeAngelo et al., 2022). Large residuals, positive or negative, imply the possible presence of convection. The USGS removed known geothermal sites before their interpolation and performing the residual calculations.

An important distinction is that negative residuals are not equivalent to a "negative" geothermal site. Negative residuals can be indicative of meteoric water recharge, often due to neighboring active hydrothermal sites. Qualitative evidence suggests that in some cases negative temperature anomalies define zones peripheral to upwelling zones related to a geothermal system; thus, in some cases, negative temperature anomalies in wells could indicate the nearby presence of a geothermal system. Negative residuals associated with positive geothermal locations would likely be of relatively large-magnitude compared to negative residuals elsewhere. The objective is to identify negative sites that would not have a pattern of a high positive residual next to a negative residual.

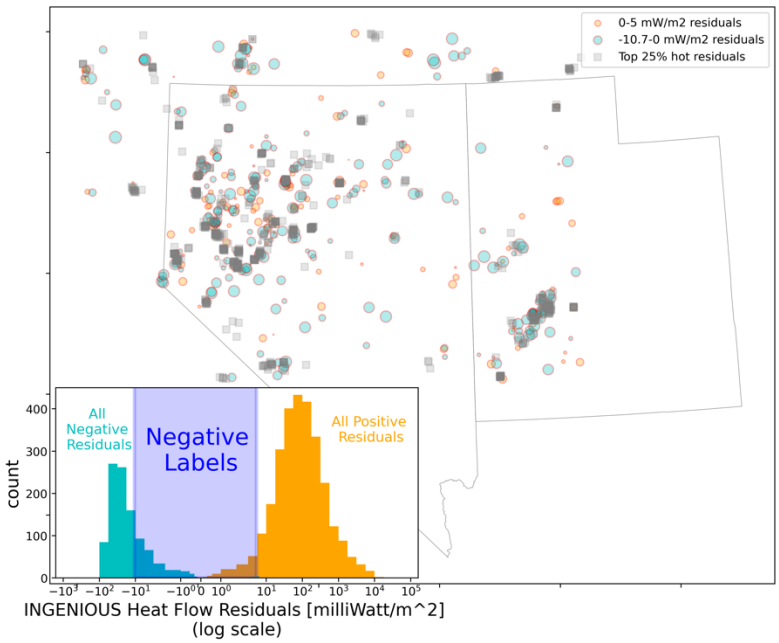

Figure 3: Inset: Distribution of heat flow residuals and the selection of which are negative labels. Note: The x-axis is in log scale. Those within the range of the "negative labels" are plotted on the map, along with the top quartile of the hot residuals.

To achieve this, we statistically examined the magnitudes of the negative residuals to determine where they indicate a lack of convection (a negative geothermal label), as opposed to a downflow or recharge associated with a nearby geothermal system. The residuals where thermal gradient values were judged to be of low quality by subject matter experts were de-emphasized. We used the 75th percentile of the negative residuals (-10.7 milliWatt/m^2) to define the lower-bound magnitude to use for the negative labels. To define the upper-bound magnitude, we used +5 milliWatt/m^2 to include very-low-magnitude positive residuals. These bounds are displayed by the distribution of residuals in Figure 3.

Additionally, we identified the upper quartile (greater than the 75th percentile) of the positive heat residuals as greater than 180 milliWatt/m^2. Any negative residuals within 10 km of these hot residuals were removed to ensure that no downflows were included. This resulted in 248 negative sites. The proportion of negative sites in each of the four domains of (Faulds et al., 2015a) is as follows: 10.4% in the Central Nevada Seismic Belt, 31.4% in the Carbonate Aquifer, 9.2% in the Walker Lane, and 50% in the Western Great Basin. These numbers exceed 100% because of the overlapping domains (Figure 4).

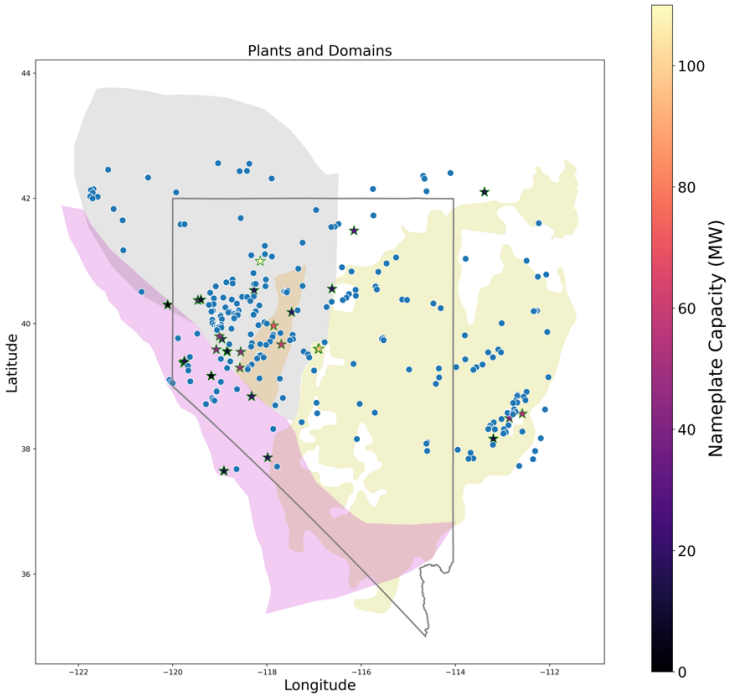

Figure 4: Map of power plants (stars), "negative" sites (blue circles, 0 capacity), and domains of the INGENIOUS study area (pink: Walker Lane, orange: Central Nevada Seismic Belt, gray: Western Great Basin, yellow: Carbonate Aquifer).

The Great Basin has the highest concentration of geothermal power plants in the United States. The positive sites are the 37 existing geothermal plants that are operating in the study area of the current INGENIOUS research project, that includes most of Nevada and parts of California, Utah, and Idaho (Akindipe et al., 2025). These geothermal power plants nameplate capacities allow for a regression to capacity. They are depicted by stars in Figure 4, and their distribution is shown in Figure 5. Figure 4 also depicts a nominal understanding of the subregions: Carbonate Aquifer (yellow), Central Nevada Seismic Belt (orange), Walker Lane (pink), and the Western Great Basin (gray) (Faulds et al., 2015a, 2015b; Smith et al., 2023).

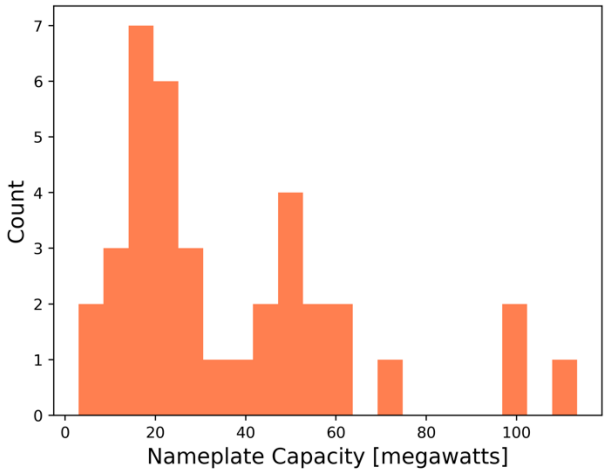

Figure 5: Nameplate capacity of 37 geothermal power plants in the INGENIOUS study area.

2.2 Features: INGENIOUS Data Layers

The datasets from the INGENIOUS project offer a novel view into the high-quality and high-resolution data relevant to hydrothermal systems within the Great Basin. The datasets are represented at a resolution of 250 m by 250 m and contain 48 features. Figure 6 plots six example datasets for the INGENIOUS area. A list of all the features used for the prediction is shown in Table 1. Of keen interest here is to use these features collocated with the 37 geothermal nameplate capacities to glean which layers might help in predicting geothermal

capacity relationships. We use the XGBoost "feature importance function", which measures which data layers statistically have a relationship with nameplate capacity. As shown by the color bars in Figure 6, the features have very different scales. We experimented with different scaling techniques to ensure equal contribution in the regression algorithm.

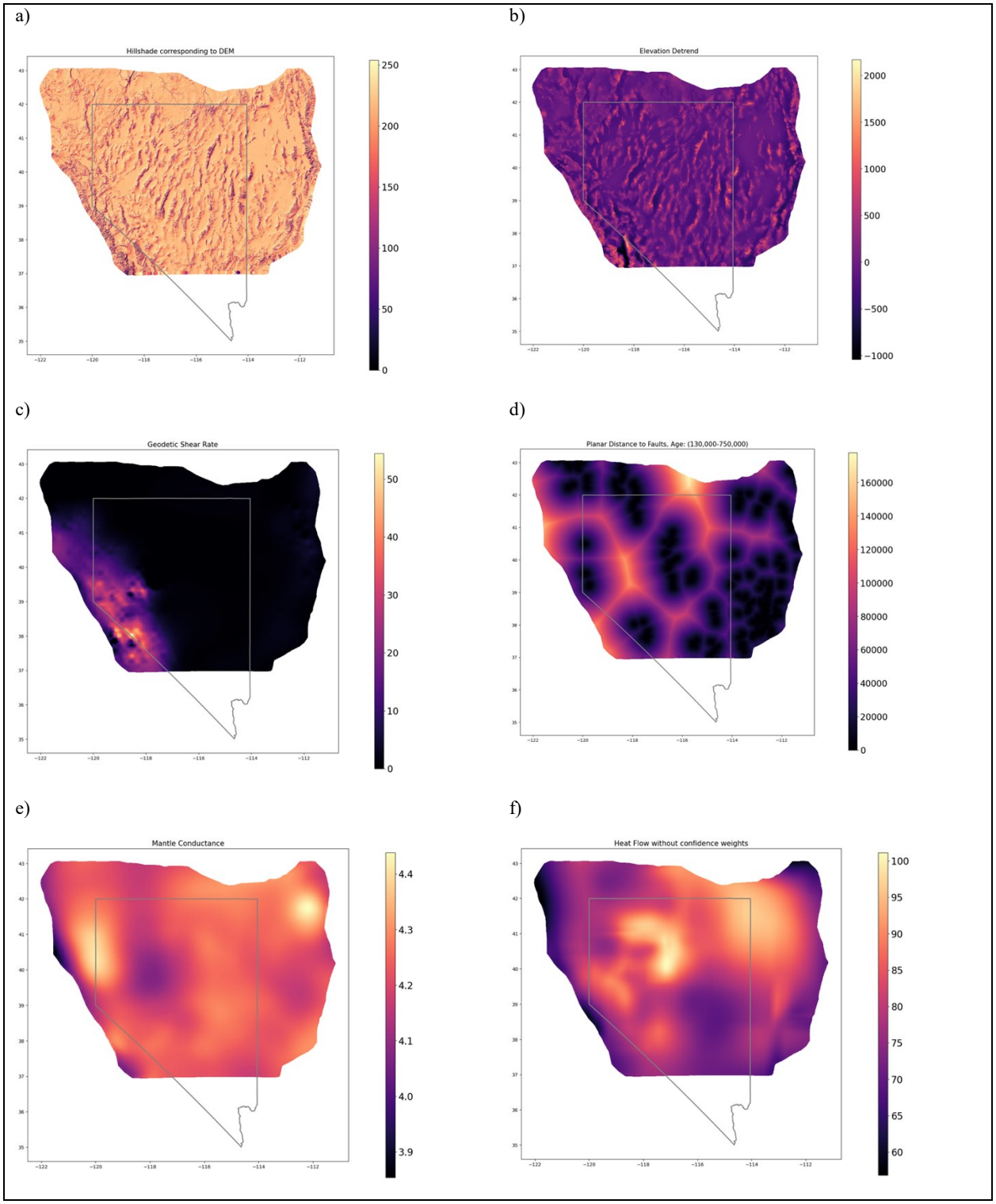

Figure 6: Six example input features for the XGBoost regression of capacity: a) hill shade corresponding to DEM; b) elevation detrended; c) shear stress; d) planar distance to feature: quaternary faults age 130,000–750,000; e) mantle conductance; and f) heat flow.

We explored feature reduction with principal component analysis given the higher number of features versus labels (48 versus 37), and twenty components explain 95% of the variance. But using the 20 components in place of the 48 features did not improve the fit to 37 plants, and predictions at negative sites were significantly worse; therefore, we continued the regression with the original data layers.

Table 1: List of 48 data layers used as features in the XGBoost regression

	Feature				Feature		Feature
1	Conductivity from MT, lower crust (20-50 km depth)	13	Dependent earthquake rate density, n=200 alpha=0.15	25	Elevation trend span01	37	Independent earthquake rate density, n=50 alpha=0.15
2	Conductivity from MT, mantle (90-200 km depth)	14	Dependent earthquake rate density, n=50 alpha=0.05	26	Elevation trend span03	38	Isostatic magnetic anomaly
3	Conductivity from MT, middle crust (12-20 km depth)	15	Dependent earthquake rate density, n=50 alpha=0.10	27	Heat flow map with thermal conductivity quality code (TC) confidence	39	Planar distance to feature: qfaults age 0-15,000
4	Conductivity from MT, surface (2-12 km depth)	16	Dependent earthquake rate density, n=50 alpha=0.15	28	Hillshade corresponding to DEM	40	Planar distance to feature: qfaults age 130,000-750,000
5	Conductivity from MT, upper mantle (50-90 km depth)	17	Geodetic strain second invariant	29	Independent earthquake rate density, n=100 alpha=0.05	41	Planar distance to feature: qfaults age 15,000-130,000
6	Digital elevation model (DEM)	18	Geodetic dilation rate	30	Independent earthquake rate density, n=100 alpha=0.10	42	Planar distance to feature: qfaults age 750,000-2,600,000
7	Depth to basement surface	19	Geodetic shear rate	31	Independent earthquake rate density, n=100 alpha=0.15	43	Planar distance to feature: qfaults age unclassified or misclassified
8	Dependent earthquake rate 9density, n=100 alpha=0.05	20	Isostatic gravity anomaly	32	Independent earthquake rate density, n=200 alpha=0.05	44	Planar distance to feature: all qfaults
9	Dependent earthquake rate density, n=100 alpha=0.10	21	Heat flow map without confidence weights	33	Independent earthquake rate density, n=200 alpha=0.10	45	PRELIMINARY qfault feature set engineering; recency and slip rate combined using logistic regression; scalar indicates favorability with greater values being more favorable
10	Dependent earthquake rate density, n=100 alpha=0.15	22	Heat flow map with thermal gradient quality code (QC) confidence	34	Independent earthquake rate density, n=200 alpha=0.15	46	Magnetics Horizontal Gradient.
11	Dependent earthquake rate density, n=200 alpha=0.05	23	Elevation detrended span005	35	Independent earthquake rate density, n=50 alpha=0.05	47	Detrended elevation surface using span = 0.01
12	Dependent earthquake rate density, n=200 alpha=0.10	24	Elevation trend span005	36	Independent earthquake rate density, n=50 alpha=0.10	48	Detrended elevation surface using span = 0.03

2.3 XGBoost Regression

We chose XGBoost to estimate the hydrothermal technical potential from the collocated INGENIOUS data features, with the goal of improving the permeability representation from previous estimates. XGBoost regression is explored using the 37 geothermal plant capacities, the negative geothermal sites, and the INGENIOUS data layers. We used two variants of XGBoost in this work, where the difference between the two involved the choice of loss function. As with any machine learning algorithm, the objective is to minimize the misfit: the difference between the observed and predicted value, otherwise known as the loss. First, we experimented with sci-kit learn's XGBoost regression model with a mean square loss, and we used the learnings from this for the XGBoost quantile regression, which uses a pinball loss function, which provides the percentile estimates (Pedregosa et al., 2011).

Modelers must make several decisions about the XGBoost algorithm, and these can influence the results. The first is the loss function: how to quantify the difference between the observed and modeled (or predicted) nameplates. We began our exploration by using the mean square error (MSE):

$$MSE_{Loss}(y,\hat{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples-1}} (y_i - \hat{y}_i)^2 \quad (1)$$

where y_i represents the observed nameplate capacities and the negative sites, and $\hat{y}$ represents the corresponding estimate from the XGBoost model. We also conducted training experiments done with the mean absolute error (MAE), but the fits were generally worse, and thus they are not included here. This is to be expected because with scaling, there were no great outliers in the features or labels. We report the mismatch of the model and observations in both MSE and MAE. The loss function also plays a role in the search to identify the optimal hyperparameter values, which are input values that influence the XGBoost learning process. We performed a grid search for five hyperparameters for each scaling and loss function combination. The grid search evaluated the loss of the 285 observations and was performed five times, as a "fivefold" cross-validation that separates the dataset into five distinct training and test sets to evaluate which hyperparameters minimize the MSE loss. Using this cross-validation grid search, we attempted to prevent overfitting by shuffling the training-test set. We used a stratified train-test split to ensure that the positive nameplate capacities are represented in each set.

We compared the XGBoost capacity predictions of five combinations of scaling and loss functions. We used this to guide which scalers and loss functions are more appropriate. Scaling is necessary for features with vastly different scales and potential outliers, which otherwise could reduce the predictive performance of the XGBoost regression. The choice of loss function gives the modeler the preference to penalize the magnitude of overestimates or underestimates;, some behave better with outliers. We performed the hyperparameter search for each unique loss and scaling combination.

Negative predictions (in this case, of nameplate capacity) are known to occur with the XGBoost algorithm because of how XGBoost combines the residuals with every sequential tree. The target for the next tree incorporates the previous loss (misfit or residual) and the learning rate. With many trees, the magnitude of this value can get large enough to produce values outside the range of training data.

The log scaling of the target variable and the Poisson loss function can address these negatives. The log scaling was performed only on the dependent variable (the nameplate capacities and negative sites):

$$y_{scaled} = \log_{10}(y + 1^{-10}) \quad (2)$$

to reduce the negative predictions.

The Poisson loss is typically used for predicting frequencies and counts where only positive predictions are desired:

$$Poisson_{Loss}(y,\hat{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples-1}} 2\left(y_i \log(y_i/\hat{y}_i) - \hat{y}_i - y_i\right) \quad (3)$$

2.4 XGBoost Quantile Regression

For the objective of having multiple, bounding megawatt estimates for hydrothermal in reV, we used the XGBoost-based quantile regression, which can produce estimates for given confidence intervals. To use XGBoost for quantile regression, we used the pinball loss function. By specifying the quantile we want to predict, the pinball loss function will place asymmetric weights on overestimates and underestimates. For a given quantile α, the model will be optimized so that the predicted value will be higher than α% of the training set and lower than (1- α) % of the training set. As a result, it is important to include only the same number of negative labels as there are positive sites, because if there is an imbalance in the number of labels, the predicted value when α = 0.5 would tilt toward the label that has more occurrences in the training set.

In the XGBoost quantile regression model, we generated plots and results by optimizing the P10, P50, and P90 by training separate models for each quantile, each with different hyperparameters. As mentioned earlier, each quantile model used in XGBoost was tuned separately, and we used three different models (P10, P50, and P90) for the predictions.

The pinball loss function, also referred to as the quantile loss, is a metric used to assess the accuracy of a quantile forecast (Pedregosa et al., 2011):

$$Pinball_{Loss}(y,\hat{y}) = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples-1}} \begin{cases} \alpha(y-\hat{y}) & \text{if } y \geq \hat{y} \\ (1-\alpha)(\hat{y}_i - y_i) & \text{if } y < \hat{y} \end{cases} \quad (4)$$

The alpha value in Equation 4 indicates the quantile being estimated. The pinball loss has been named due to its shape—it mimics the trajectory of a ball on a pinball. The function is positive and larger the farther it is from the target or actual value y_i. The slope reflects the imbalance in the quantile forecast, where larger percentiles (P90) penalize overpredictions (e.g. $\hat{y}_i > y$).

3. RESULTS

3.1 Standard Regressions

Similar to Figure 1, Figure 7 displays the predictions on the y-axis versus the 37 nameplate capacities on the x-axis for these five combinations. Figure 7 shows that the quantile scaler in combination with the MSE loss function (cyan dots) provides the smallest loss (95) of the five compared to the 37 nameplate capacities, with the Poisson loss function close behind. Figure 7 provides a visual of the predictions at the locations of the power plants, and Table 2 summarizes the misfit (prediction error) for each model shown in Figure 7.

Recall that the log-scaling of the target variable (orange in Figure 7 and Figure 8) and the Poisson loss function (blue in Figure 7 and Figure 8) are included to address the negative predictions. As shown in Figure 8, this method has no negative prediction, but the predictions at the power plants (Figure 7) were quite sensitive to the chosen factor added (e.g., 1^{-10} in Eqn 2), and they produced higher losses.

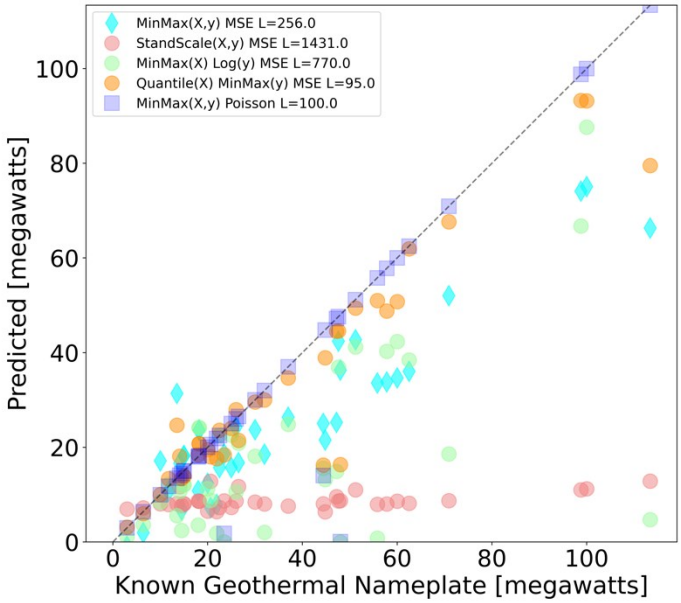

Figure 7: The x-axis shows the observed nameplates, and the y-axis shows the predicted values at the plant locations with specified scaling on either the INGENIOUS features (x) and/or the nameplate capacities (y).

Figure 8 displays the distribution of predictions collocated at the 248 negative sites for the same five combinations shown in Figure 7. The box plots represent the interquartile range, with the whiskers representing a factor of 1.5 of the interquartile range from the nearest quartile. The remaining represent every prediction that lies outside those ranges. The Poisson predictions predicted no negatives (Figure 8). Although the quantile MSE predicts negative values, the box plot is more tightly centered around zero than the MinMax MSE (cyan).

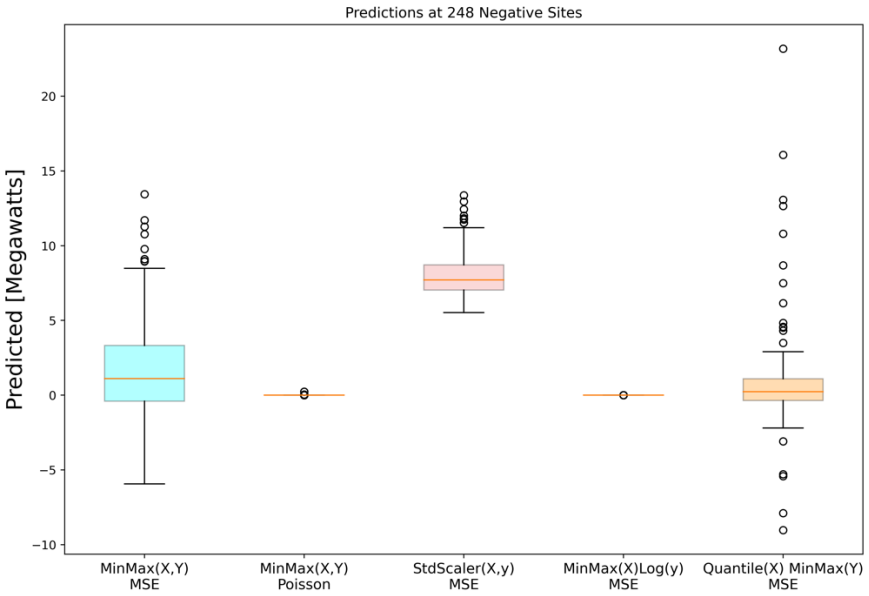

Figure 8: Predictions at negative sites: five combinations of scalers and loss functions.

Table 2: MSE and MAE at 37 Geothermal Plants, 248 Negative Sites, and Prediction Percentiles [MW] for the INGENIOUS Study Area

	Predictions at labels only		**Predictions for the entire Great Basin (INGENIOUS) area**			
Transform & loss function	**MSE with 37 testing plants**	**MSE with 248 negative sites**	**P10**	**P50**	**P90**	**% negative**
Exponential (SMU temperatures)	1,049.4/21.7	21,702/113.8	2.21	2.63	3.02	0
MinMax(x,y) & MSE	262.0/12.7	13.1/2.62	-3.73	2.4	12.5	15.78
Standard scaler (both), MSE	1,431/27.9	65.9/7.99	4.15	5.05	6.6	0
Quantile transform(x), MinMax(y), MSE	95/5	7.88/1.31	-4.11	2.5	140.	16.25
MinMax(x), Log(Y+1e-10)/e-5, MSE	770/19	0/0	-8.2e-11	4.1e-10	1.8e-07	13.74
MinMax & Poisson loss	3	3.5e-3	4.86e-05	2.42e-04	2.08e-02	0

Next, we applied the five combinations to the entire INGENIOUS area, as the prediction distributions are another consideration when choosing the best method to use, in addition to the prediction error with the power plant location. Table 2 summarizes the total, P10, P50, and P90 capacities of those estimated predictions for the entire INGENIOUS area. Recall we use the convention that P10 and P90 are the

smaller and larger percentiles, respectively. For comparison, the SMU dataset is listed in the top row. Whereas Poisson fits the plant data perfectly, the predictions across the INGENIOUS area are incredibly low, with a P90 of 2.08e-02 MW; therefore, the two most promising results given both their fit to the positive and negative sites and their predictions throughout the INGENIOUS are MinMax(x,y) and quantile(x) MinMax(y). These two predictions are shown in .Figure 9 Overall, similar patterns are seen around the Walker Lane (in the southwest) and the northwest of the INGENIOUS region. These patterns are largerly caused by the elevation trends.

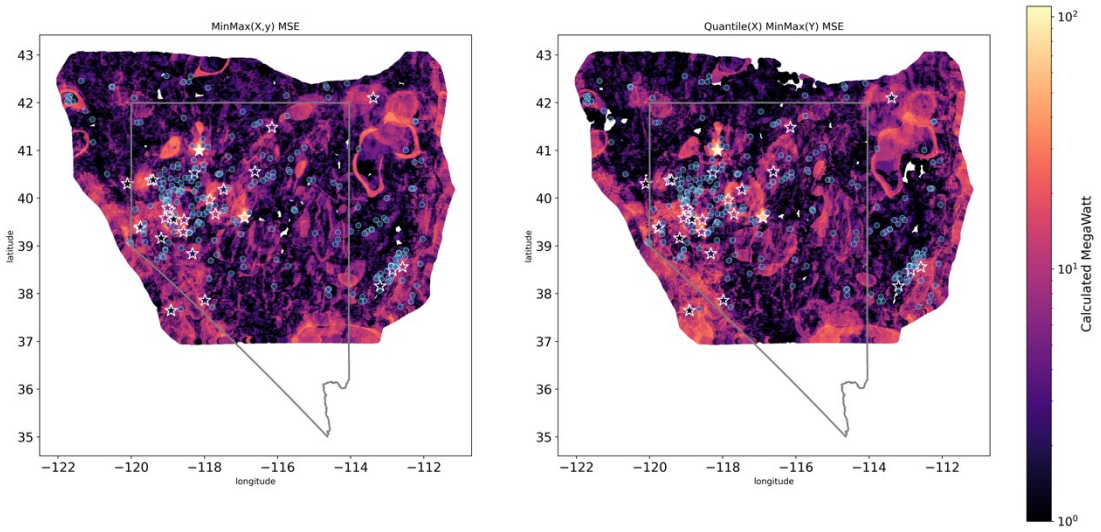

Figure 9: Hydrothermal predictions given the MSE loss function and the (left) Min.Max(x,y) scaler versus the (right) quantile(x) and MinMax(y) scaler.

3.2 Quantile Regressions

For the quantile regression, we explored methods for balancing the positive and negative labels: e.g., use only 37 negative sites at a time to match the number of power plants. Compared to the loss functions that were previously explored, the pinball loss function is more sensitive to the imbalance in the positive and negative labels. We chose to partition the negative sites in four different ways to demonstrate the sensitivity of the predictions to the locations of the negatives and to attempt to include some geologic domain information. The first two methods consider only the absolute distance of any negative site to an existing geothermal power plant until 37 negatives are chosen. The second two methods consider the proportion of how many negatives are in each of the four domains (the four colors in Figure 4) of the 228. Thus, the four methods are:

1. Include the closest 37 negative sites from geothermal power plants.
2. Include the farthest 37 negative sites from geothermal power plants.
3. Include the closest 37 negative sites from plants by proportionally sampling negatives from each domain. Domains (Figure 4) with more negative sites were given a higher count, summing to 37 total negative sites.
4. Include the farthest 37 negative sites from plants by proportionally sampling negatives from each domain. Domains (Figure 4) with more negative sites were given a higher count, but they sum to 37 total negative sites.

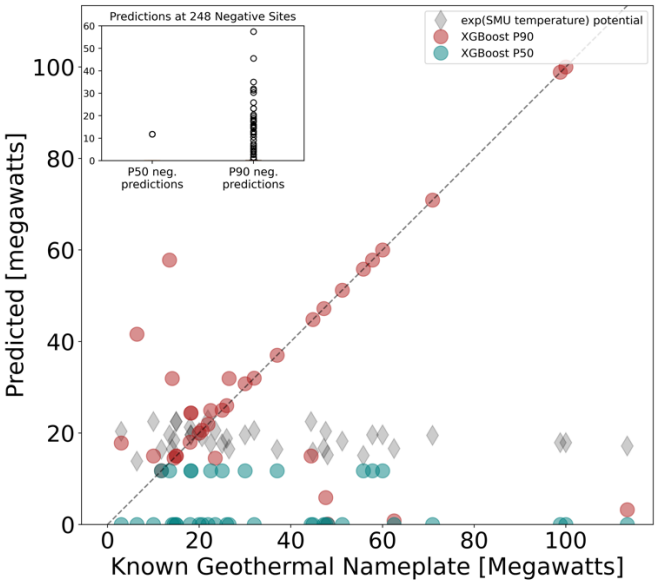

Figure 10: P50 and P90 model predictions at geothermal power plants using all 248 negative labels in fitting the quantile regression.

Figure 10 displays the quantile model predictions against the observed 2023 nameplate capacities of the geothermal power plants, demonstrating how P90 in general fits the power plants better than P50, except for overestimating plants <20 MWe, whereas P50 predicts 0 MW at most of the geothermal power plant locations. All P10 models in the quantile regression produced zero capacity; therefore, they will not be shown.

Table 3 summarizes the P50 and P90 XGBoost quantile model fits to the 37 geothermal nameplate capacities and negative locations for all permutations of the negative sites used. The quantile model using all 248 negatives performs the worst when compared to the power plants. Conversely, the misfit with the negative locations worsens when only 37 negative sites are used in the regression. The visuals of the P50 and P90 predictions for the four negative subsamples are shown in Figure 11.

Table 3: Quantile Regression Error Metrics With Training Data

	MSE/MAE with plants (Scaled X and Y)		**MSE/MAE with all negatives (Scaled X and Y)**	
Negative labels	**P50**	**P90**	**P50**	**P90**
All 248	1,845.4/32.9	670.2/11.8	6.7/0.57	73.17/3.2
Closest 37 to geothermal power plants	751.2/14.9	597.5/16.8	234.0/11.6	1,523.9/37
Farthest 37 from geothermal power plants	1,035.0/21.4	389.0/8.1	202.4/11.6	1182.4/27.5
Closest 37 with proportional sampling of negatives from domains	393.3/6.4	623.9/17.1	368.2/14.5	1,711.9/38.8
Farthest 37 with proportional sampling of negatives from domains	1,026.7/20.9	600.7/17.4	189.7/11.8	1,483.0/34.3

Some geographic and geologic insights about the locations of the negatives in the four scenarios and the misfits are shown in Table 3 and Figure 11. Using only the 37 negatives closest to power plants, P50 better predicts the plants, but P90 and P50 overlap, and both overpredict the 248 negatives (inset box plot). Using the 37 negatives that are farthest from the power plants, the P50 predictions are much lower. These negatives are mostly in the Carbonate Aquifer (where there are no power plants); therefore, the INGENIOUS geologic and geophysical features are likely the most different between positive and negative sites. The predictions using these negatives are lower than those using negatives close to power plants.

The next two sets of subsampled negative sites proportionally grab the negative sites according to how the 248 negatives are distributed in the four domains (Figure 4): 10% in the Central Nevada Seismic Belt, 31% in the Carbonate Aquifer, 9% in Walker Lane, and 50% in the Western Great Basin. In general, negative sites chosen close to plants, results in higher P50 and P90 predictions (left column of Figure 11). Using the domain-proportional and closest 37 negative sites gives the highest predictions where the P50 model fits along the 45-degree line, and the P90 overpredicts, as shown in Figure 11c. Last, for the farthest 37, with proportional sampling of the negatives from the domains, (Figure 11d) P50 is very low (and similar to SMU based prediction), and P90 predicts lower compared with the farthest 37 negatives (Figure 11c).

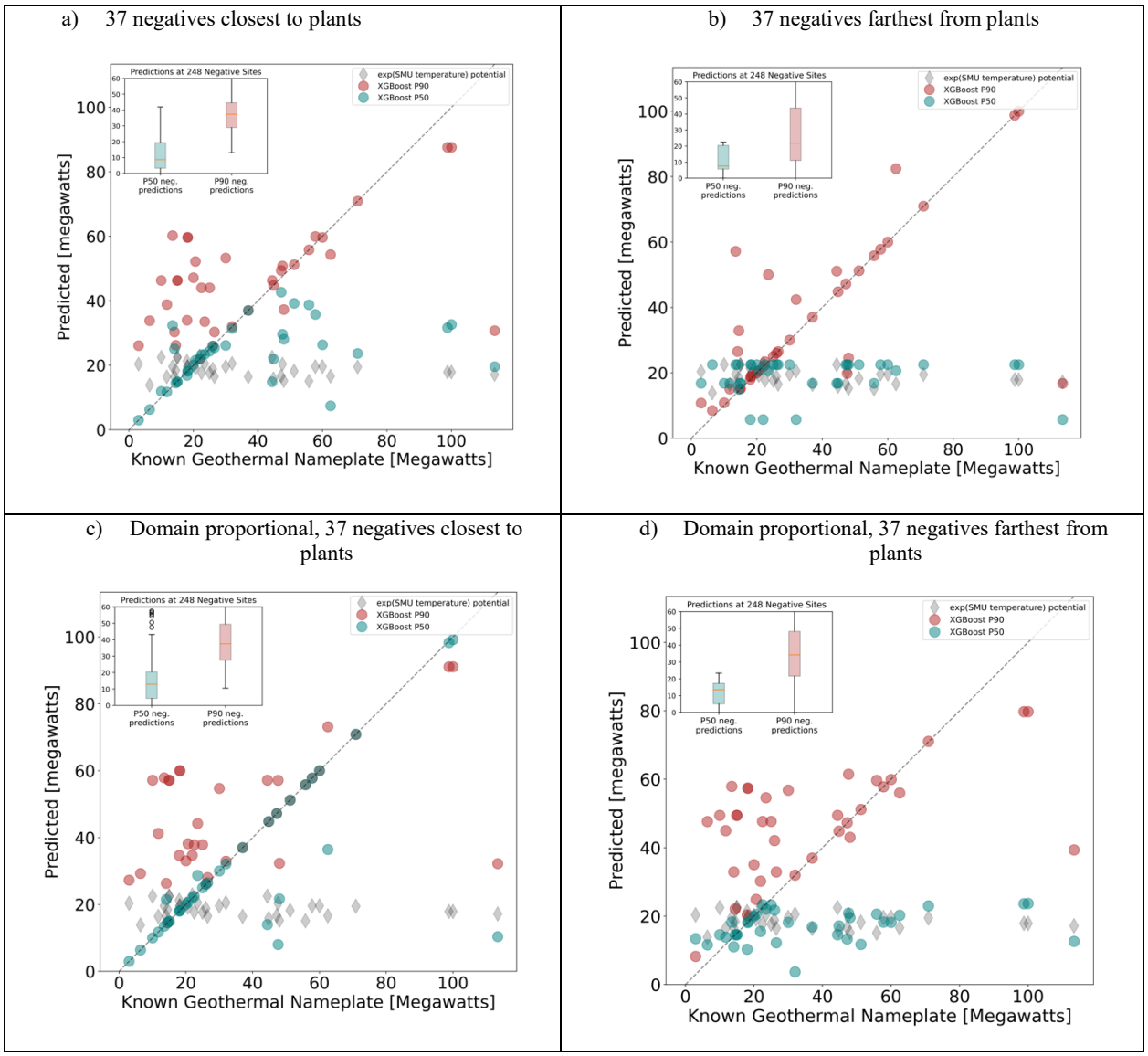

Figure 11: P50 and P90 model predictions at the geothermal plant locations using a) only 37 negatives closest to the power plants; b) only 37 negatives farthest from the power plants; c) 37 negatives proportionally chosen by domain, closest to the plants; and d) 37 negatives proportionally chosen by domain, farthest from the plants.

The next step is to evaluate how these look as predictions across the INGENIOUS area. We considered only the cases where all negatives are used and the two cases with the negatives that farthest from the power plants, as the other two are deemed to overpredict. Figure 12 displays the P50 and P90 predicitons for the farthest, proportional 37 negatives.

Table 4: Percentile Averages (MW) for Quantile XGBoost Models for the INGENIOUS Study Area

	P50 Quantile			P90 Quantile		
Negative labels	**p10**	**p50**	**p90**	**p10**	**p50**	**p90**
All 248	0	0	11.75	0	0.22	32.5
Farthest 37 negatives from power plants	0	5.7	20.6	3.1	20.8	58.7
Farthest 37 with proportional sampling of negatives from domains	-1.12	9.3	18.5	8.19	24.88	50.4

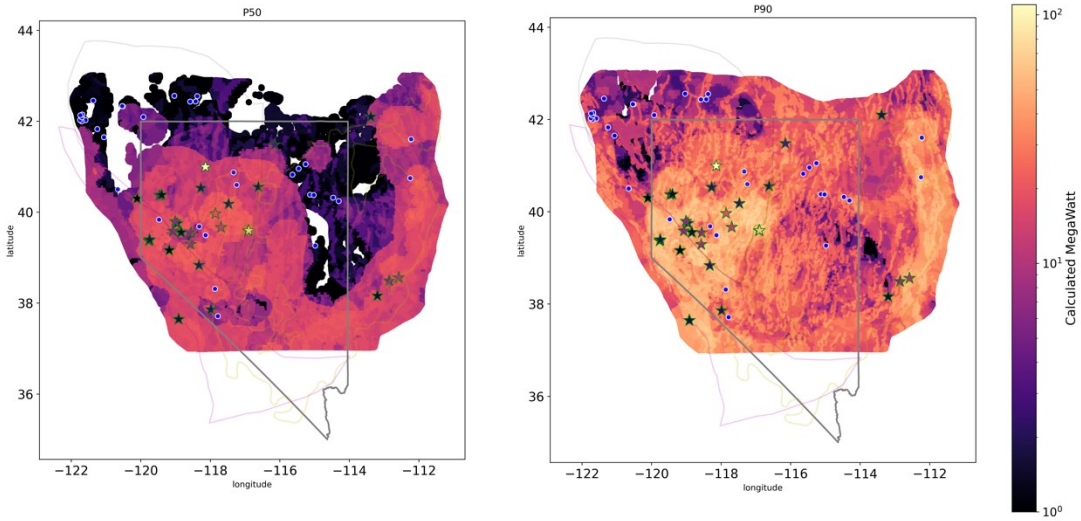

Figure 12: Farthest 37 with proportional sampling of negatives from domains (left) P50 and (right) P90.

3.3 reV Results

The technical potential from XGBoost is passed to reV, which is upscaled from the 250-m by 250-m INGENIOUS resolution to a 10-km by 10-km resolution. This area is based on the resolution of reV's exclusion data layers (Maclaurin et al., 2021). Default GETEM parameters are used to estimate the geothermal capacities and the LCOE across the whole Great Basin region. These include 2035-projected capital and operating costs as determined by the Annual Technology Baseline for predictions of 2035 (NREL, 2023; Pinchuk et al., 2023) and default performance assumptions such as production well flow rate (110 kg/s per well), plant efficiency (80%), and ratio of injection wells to production wells (0.75).

Another useful output of reV is the LCOE for all-in, which considers the transmission interconnection costs ($/MWh). We present both "site" LCOE and "all-in" LCOE for the three XGBoost-derived hydrothermal estimates. The all-in LCOE for each model will depend on the location of the capacity estimate with respect to the existing transmission features. Figure 13 shows the reV technical capacity and the two LCOEs from the quantile scaling MSE loss XGBoost. Compared to the original SMU temperature-derived hydrothermal estimates (in Figure 14 from Pinchuk et al. (2023)), we see that the highest capacities from XGBoost are located closer to Walker Lane and Central Nevada. The LCOE without transmission (site) is much lower in the XGBoost result.

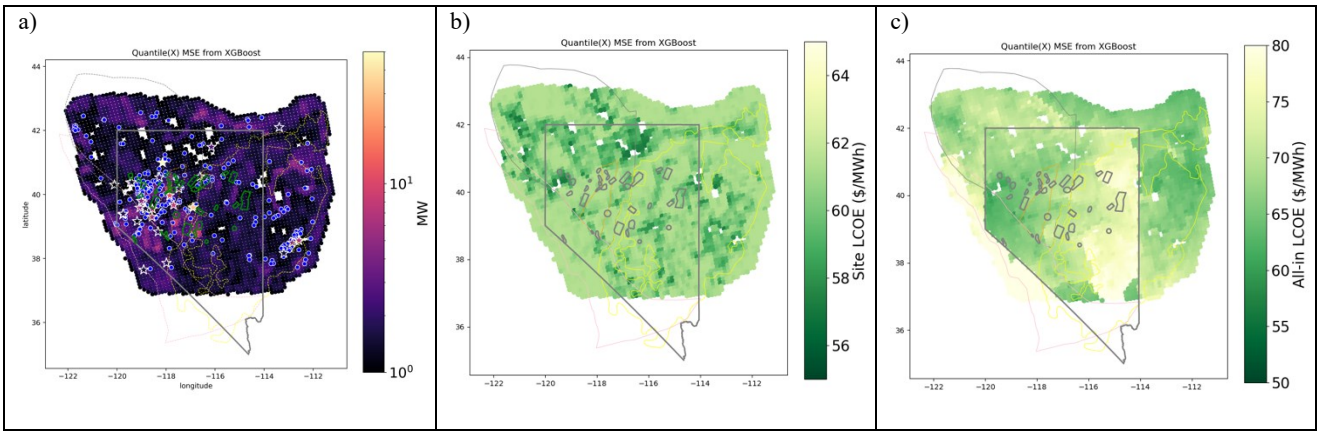

Figure 13: XGBoost quantile(X) MinMax(y) results from reV: a) technical capacity (and negative sites), b) site LCOE, and c) all-in LCOE.

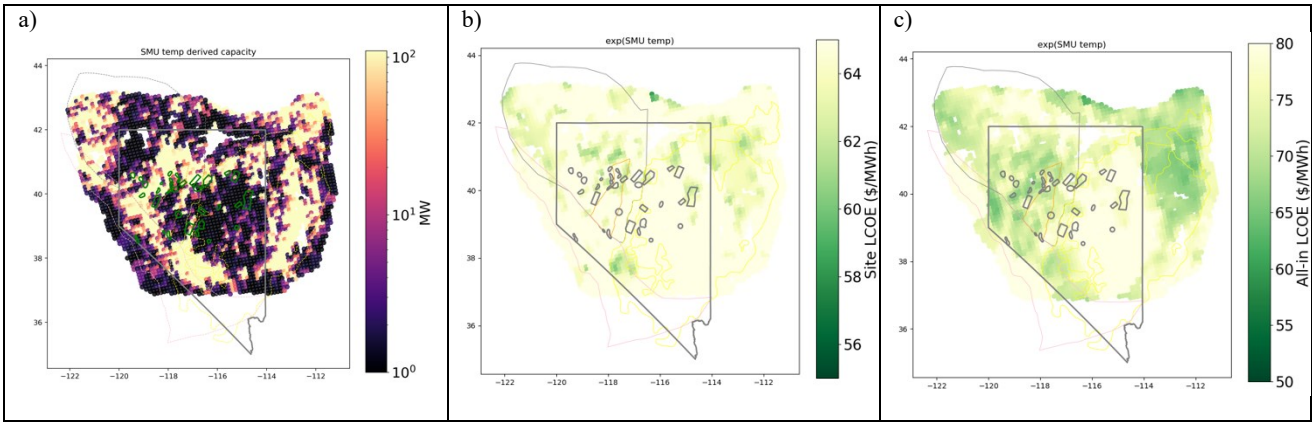

Figure 14: Results from SMU temperatures in reV: a) technical capacity, b) site LCOE, and c) all-in LCOE.

The higher values are obvious in the histogram comparisons in Figure 15, which include the P50 and P90 models as well. The XGBoost site LCOE has a much lower range. The LCOE maps including transmission costs and their distributions (far right column of Figure 15), look much more similar but XGBoost are slightly lower.

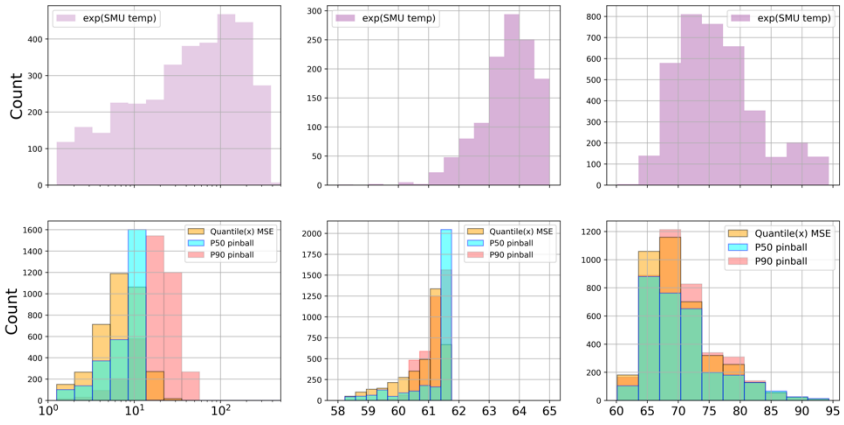

Figure 15: Distributions of capacity (left column), site LCOE (middle), and all-in LCOE (right) of the SMU temperature-derived estimates (top row) compared to the three XGBoost models.

We also explored how these reV metrics compared within areas that have been deemed of high interest with high favorability for hidden systems versus all other locations. These high-interest areas are shown in the gray polygons of Figure 13 and Figure 14. We compare the estimated power density for the three models for the areas located inside and outside these polygons in Table 5. For the MSE loss model, these areas are 40% higher, whereas for the P50 and P90 models, they are 13% and 20% higher.

Table 5: reV Metrics for High Favorability Versus Everywhere Else

	Loss = MSE		P50 pinball		P90 pinball	
	In high-interest areas	Not in high-interest areas	In high-interest areas	Not in high-interest areas	In high-interest areas	Not in high-interest areas
Power density [MW/km^2]	0.069	0.05	0.049	0.043	0.18	0.149
Median site LCOE [$/MWh]	61.06	61.05	61.67	61.7	61.15	61.31
Median LCOE (including LCOT) [$/MWh]	70.40	68.94	69.75	69.28	69.88	69.68
Power density ratio: high interest to everywhere else	1.38		1.13		1.2	

4. Discussion & Conclusions

The labels used here have considerable uncertainty. Geothermal nameplates change over time, given the temperature and pressure drawdown of the field, how well managed the operations are, and how optimized the recharge wells are given the permeability of the field. As mentioned at the beginning of the paper, true negative labels would be indicators of no natural permeability—therefore, wells that have no flow. Given the absence of this information, we used recent heat flow residuals to identify locations where very-low-heat flow would be expected, exclusionding downflow or natural recharge locations.

Although the INGENIOUS layers represent a large advancement in hydrothermal exploration, they still might contain errors, and their representation in XGBoost might not be the ideal way to represent thesem. Convolutional neural nets might better capture the 3D nature of the data, which is what drives these systems, including the structural setting, which is a complex interaction of the faults and strain characteristics.

The objective of this work was to improve the hydrothermal estimates from previous estimates in reV that did not include any indication of permeability by using only an exponential transfer of SMU temperatures (Pinchuk et al., 2023). This exponential relationship was developed by fitting 103 worldwide geothermal power outputs and reservoir temperatures (Wilmarth et al., 2021). To include permeability in hydrothermal estimates, we used 37 operating nameplate capacities in megawatts as reported in 2024 and 48 geologic and geophysical data layers that can serve as proxy indicators of permeability.

Two variations of XGBoost regressions were employed: The first used quantile scaling and an MSE loss function, whereas the second attempted to quantify the uncertainty using a pinball loss function to define the50th and 90th percentile models. For both XGBoost variants, all 37 of the positive hydrothermal labels were included; however, the label imbalance affected the estimates for the percentile models more. We explored ways subsampling the negative sights to bring the labels into balance and using the geologic domain knowledge to proportionally represent the negatives.

Three XGBoost predictions were input into the reV model. These models were chosen because of their fits to the 37 geothermal plants and negative locations as well as their reasonable prediction ranges and totals within the INGENIOUS area. The MSE loss model, which used all 248 negative sites, had 5 and 1.31 MAE misfits with the positives and negatives, respectively. The P50 and P90 models with negatives proportionally sampled from geologic regions had approximately 20 and 17 MAE's with the positive and negative sites, respectively.

The final reV metrics demonstrated how all three XGBoost models significantly reduced the range of technical capacities and site LCOEs compared to the hydrothermal estimates based solely on the temperature estimates. The MSE loss model, which used all 248 negative sites, had comparable capacities and LCOEs to the P50 model.

ACKNOWLEDGMENTS

This work was authored by the National Renewable Energy Laboratory, operated by Alliance for Sustainable Energy, LLC, for the U.S. Department of Energy (DOE) under Contract No. DE-AC36-08GO28308. Funding provided by U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Geothermal Technologies Office. The views expressed in the article do not necessarily represent the views of the DOE or the U.S. Government. The U.S. Government retains and the publisher, by accepting the article for publication, acknowledges that the U.S. Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for U.S. Government purposes. We thank Sean Porse and Andy Adams (U.S. Department of Energy Office of Energy Efficiency and Renewable Energy Geothermal Technologies Office) for their insight and support of the project.

REFERENCES

Akindipe, D., Pauling, H., Martinez Smith, F., Witter, E., Podgorny, S., Rane, J., Harmon, J., Javed, S., Shlosberg, D., Holt, E., Trainor-Guitton, W., Levine, A., 2025. 2025 U.S. Geothermal Market Report (No. NREL/TP-5700-91898.). National Renewable Energy Laboratory.

Augustine, C., 2017. A Methodology for Calculating EGS Electricity Generation Potential Based on the Gringarten Model for Heat Extraction From Fractured Rock (No. NREL/CP-6A20-66485). National Renewable Energy Lab. (NREL), Golden, CO (United States).

Augustine, C., 2016. Update to Enhanced Geothermal System Resource Potential Estimate, in: Geothermal Resources Council Transactions. Presented at the GRC, Geothermal Resources Council Transactions, pp. 673–677.

Ayling, B., Faulds, J.E., Rivera, A., Koehler, R., Kreemer, C., Mlawsky, E., Kleber, E., 2022. INGENIOUS - Great Basin Regional Dataset Comilation. Study Area Boundary. https://doi.org/10.15121/1881483

Blackwell, D., Richards, M., Frone, Z., Ruzo, A., Dingwall, R., Williams, M., 2011. Temperature-At-Depth Maps For the Conterminous US and Geothermal Resource Estimates, in: GRC. Presented at the 1029452.

Craig, J.W., Faulds, J.E., Hinz, N.H., Earney, T.E., Schermerhorn, W.D., Siler, D.L., Glen, J.M., Peacock, J., Coolbaugh, M.F., Deoreo, S.B., 2021. Discovery and analysis of a blind geothermal system in southeastern Gabbs Valley, western Nevada, USA. Geothermics 97, 102177. https://doi.org/10.1016/j.geothermics.2021.102177

DeAngelo, J., Burns, E.R., Gentry, E., Batir, J.F., Lindsey, C.R., Mordensky, S.P., 2022. Heat flow maps and supporting data for the Great Basin, USA. https://doi.org/10.5066/P9BZPVUC

Faulds, J.E., Hinz, N.H., Coolbaugh, M.F., DePolo, C.M., Siler, D.L., Shevenell, L.A., Hammond, W.C., Kreemer, C., Queen, J.H., 2015a. Discovering Blind Geothermal Systems in the Great Basin Region : An Integrated Geologic and Geophysical Approach for Establishing Geothermal Play Fairways, Department of Energy Final Report.

Faulds, J.E., Hinz, N.H., Coolbaugh, M.F., Sadowski, Andrew J, Shevenell, Lisa A, Mcconville, Emma, Craig, J.W., Sladek, Chris, Siler, D.L., 2017. Progress Report on the Nevada Play Fairway Project: Integrated Geological, Geochemical, and Geophysical Analyses of Possible New Geothermal Systems in the Great Basin Region, in: Proceedings. Presented at the 42nd Workshop on Geothermal Reservoir Engineering, Stanford University.

Faulds, J.E., Hinz, N.H., Coolbaugh, M.F., Shevenell, L.A., Siler, D.L., Craig, M., Hammond, W.C., Kreemer, C., Oppliger, G., Wannamaker, P.E., Queen, J.H., Visser, C.F., 2015b. Integrated Geologic and Geophysical Approach for Establishing Geothermal Play Fairways and Discovering Blind Geothermal Systems in the Great Basin Region , Western USA : A Progress Report. Geothermal Research Council Transactions 39, 691–700.

Lopez, A., Cole, W., Sergi, B., Levine, A., Carey, J., Mangan, C., Mai, T., Williams, T., Pinchuk, P., Gu, J., 2023. Impact of siting ordinances on land availability for wind and solar development. Nature Energy. https://doi.org/10.1038/s41560-023-01319-3

Maclaurin, G., Grue, N., Lopez, A., Heimiller, D., Rossol, M., Buster, G., Williams, T., 2021. The Renewable Energy Potential (reV) Model: A Geospatial Platform for Technical Potential and Supply Curve Modeling.

Mines, G., 2016. GETEM User Manual (No. INL/EXT-16-38751). Idaho National Laboratory.

NREL, (National Renewable Energy Laboratory), 2023. 2023 Annual Technology Baseline.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, É., 2011. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12, 2825–2830.

Pinchuk, P., Thomsen, S.-M., Trainor-Guitton, W., Buster, G., Maclaurin, G., 2023. Development of a Geothermal Module in reV: Quantifying the Geothermal Potential while Accounting for the Geospatial Intersection of the Grid Infrastructure and Land Use Characteristics, in: Geothermal Resources Council Transactions. Presented at the Geothermal Rising 2023, Geothermal Rising, Reno, NV, pp. 2564–2586.

Schwering, P., Lowry, T., Hinz, N., Matson, G., Sabin, A., Blake, K., Zimmerman, J., Sewell, S., Cumming, W., 2022. The BRIDGE Project ? Hidden Systems Reconnaissance in Western Nevada., in: Proposed for Presentation at the Geothermal Rising Conference 2022 Held August 28-31, 2022 in Reno, Nevada. Presented at the Proposed for presentation at the Geothermal Rising Conference 2022 held August 28-31, 2022 in Reno, Nevada., US DOE. https://doi.org/10.2172/2004444

Smith, C.M., Faulds, J.E., Brown, S., Coolbaugh, M., DeAngelo, J., Glen, J.M., Burns, E.R., Siler, D.L., Treitel, S., Mlawsky, E., Fehler, M., Gu, C., Ayling, B.F., 2023. Exploratory analysis of machine learning techniques in the Nevada geothermal play fairway analysis. Geothermics 111, 102693. https://doi.org/10.1016/j.geothermics.2023.102693

Wilmarth, M., Stimac, J., Ganefianto, G., 2021. Power Density in Geothermal Fields, 2020 Update, in: Proceedings World Geothermal Congress 2020+1.