Proceedings World Geothermal Congress 2026
Calgary, Canada, June 8 – 11, 2026

Identifying Regions Of Hydrothermal Upflow In The Great Basin Using Supervised Machine Learning

Stanley P. Mordensky[1*], John J. Lipor[2], Jeff D. Pepin[3], Jacob DeAngelo[4], Erick R. Burns[4]

[1]U.S. Geological Survey, Bozeman MT 99201, USA, [2]Portland State University, Portland OR 97201, USA, [3]U.S. Geological Survey, Denver CO 80225, USA, [4]U.S. Geological Survey, Moffett Field CA 94035, USA

*smordensky@usgs.gov

Keywords: geothermal resource assessment, hydrothermal favorability, Great Basin, geological province, machine learning

ABSTRACT

The U.S. Geological Survey (USGS) is updating the conventional hydrothermal resource assessment for the Great Basin. As part of this updated assessment, a hydrothermal favorability map is produced to show where geologic characteristics are favorable for hydrothermal resources, thereby allowing the use of conditional statistics (i.e., the likelihood of a resource being present in an area based on its predicted favorability and extent of exploration) to estimate the total (i.e., discovered and undiscovered) geothermal resources. If geologic characteristics vary across the Great Basin, then the factors controlling hydrothermal systems likely vary as well. Therefore, dividing the region based on the geologic characteristics that influence hydrothermal circulation could lead to a more accurate estimate of total geothermal resources. Geoscientists have previously developed an understanding of the regions within the Great Basin with marked differences in behavior (e.g., structurally, hydrologically, and geothermally) using expert knowledge and/or unsupervised machine learning analyses. The unsupervised analyses (e.g., principal component analysis) have been used to identify groups of data with similar behavior and describe characteristics of the regions. Use of only unsupervised methods fundamentally answers the question *which areas are different?*, after which expert input can identify if these differences are meaningful for predicting a condition of interest. To reduce the need for post-analysis expert input, the use of a supervised method explicitly incorporates what one is trying to predict (i.e., the known resources) and answers the question *which areas have differences that are important for predicting a condition of interest?* Herein, hydrothermal upflow is the condition of interest, and a supervised method is tested to delineate regions within the Great Basin based on geologic characteristics that promote or impede hydrothermal upflow.

We group hydrothermal systems and their surrounding areas in the Great Basin based on similar decision tree paths using Extremely Randomized Trees (ExtraTrees), a supervised machine learning algorithm. Tree-based machine learning methods are a series of branches through which choices are made (e.g., heat flow above or below a specified magnitude) to categorize training sites. Sites that follow the same paths through the decision trees (down the same branches) are similar in a quantifiable way. This approach identifies nine distinct regions with varying resource characteristics (e.g., frequency of occurrence and reservoir properties). This workflow delineating regions by the characteristics that control hydrothermal upflow allows for a region-by-region estimation of undiscovered hydrothermal resources through conditional statistics and thus a more precise estimation of the total geothermal resources within the Great Basin.

1 INTRODUCTION

The U.S. Geological Survey (USGS) produces periodic national and regional geothermal energy resource assessments (e.g., White and Williams, 1975; Muffler, 1979; Reed, 1983; Williams and DeAngelo, 2008; Williams et al., 2008b, 2009; Burns et al., 2025) and is directed to update these assessments "as the availability of data and developments in technology warrants" (U.S. Congress, 2005). Conventional hydrothermal assessments have historically developed and used maps of hydrothermal favorability (i.e., a relative measure of the predicted presence of the geologic conditions associated with the presence of hydrothermal upflow). These maps are subsequently coupled with the statistics of power-producing hydrothermal systems (i.e., a computation workflow termed as *conditional statistics*) to estimate the total (i.e.,

discovered and undiscovered) geothermal resources. That is, USGS geothermal resource assessments use hydrothermal favorability maps and conditional statistics to answer *what is the availability of the total resources for a range of favorability predictions given the known resources in and extent of exploration for the corresponding area*?

The most recent geothermal resource assessment was completed for the western U.S. (Williams and DeAngelo, 2008; Williams et al., 2008a,b), and the conceptual understanding, data, and analysis methods for regions of the western U.S. have greatly evolved since that assessment. Particularly, the Great Basin, an endorheic basin characterized by extensional tectonics, resulting in a landscape of north-south trending fault-block mountain ranges and intervening arid basins (e.g., Harry et al., 1993; Dickinson et al., 2006), has been a focus of intense data collection and study for geothermal resources (e.g., refer to data references in Table A1 as well as Faulds and Hinz, 2015; Faulds et al., 2016a,b, 2017, 2020, 2021a; Brown et al., 2020; Craig et al., 2021; Brown and Smith, 2022; Hart-Wagoner et al., 2023; Burns et al., 2024a,b). Early favorability and prospectivity mapping machine learning (ML) methods for the Great Basin (e.g., Coolbaugh et al., 2005a,b; Williams and DeAngelo, 2008) have evolved, employing a range of strategies for a range of purposes (e.g., Faulds et al., 2021b; Smith et al., 2023; Caraccioli et al., 2023, 2024, 2025; Mordensky et al., 2023a,b,c, 2025a,b; Hart-Wagoner et al., 2024a; Trainor-Guitton et al., 2025). Mordensky et al. (2023b) found that tree-based methods outperform most other methods given the limitations imparted by two qualities intrinsic to geologic datasets associated with geothermal resources which are: 1) thresholding characteristics necessary for favorable conditions (e.g., sufficient permeability and heat), and 2) requiring relatively few training sites for learning (e.g., 100s to 1000s of training sites as is commonly the case for conventional hydrothermal systems). Although tree-based algorithms perform robustly, overfitting can be a problem, and for geological/geothermal datasets, ExtraTrees is a tree-based algorithm that is less prone to overfitting compared with XGBoost, a common alternative tree-based approach (Mordensky et al., 2025b).

We seek to delineate regions of the Great Basin so that the conditional statistics for hydrothermal resource assessments can be conducted on regions of the Great Basin that capture the key physical attributes controlling the occurrence of hydrothermal systems, thereby producing more precise estimates. Experts have divided the Great Basin into regions for a range of purposes (e.g., Faulds et al., 2004; Brooks et al., 2014). More recently, works have attempted to employ unsupervised ML methods to identify patterns in the many new and refined datasets created for favorability and prospectivity mapping. These methods typically depend upon mathematical methods generally described as flat, centroid-based (i.e., linear) clustering methods. For example, Ahmmed et al. (2021) and Ahmmed and Vesselinov (2022) used non-negative matrix factorization with k-means clustering and groundwater geochemistry data to identify hidden hydrothermal systems in the Great Basin. Smith et al. (2023) used principal component analysis (PCA) coupled with k-means clustering to identify regions in a 96,000-km^2 section of the Great Basin. Later, Hart-Wagoner et al. (2024b) used a different approach that again used PCA on geologic datasets for the Great Basin, but they applied k-means clustering to the loading vectors for known hydrothermal systems to categorize the systems into different types. Herein, we detail a workflow that differs from the previous data-driven approaches by using a non-linear (i.e., a supervised tree-ensemble) approach and data related to hydrothermal systems covering a 488,379-km^2 area to aid in the assignment of regionality dependent upon the datasets and threshold values within those datasets that are most relevant for predicting hydrothermal favorability.

2 METHODS

Delineating regions from a tree-based supervised ML model first requires fitting the model to the training data to create a resource favorability map (Fig. 1a→b). Then, the training sites are clustered by their similarity in their paths taken through the tree-ensemble (i.e., by their leaf-distance), where the distance used here is the Hamming distance (Hamming, 1950) (Fig. 1a→c). Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBScan; Campello et al., 2013) then uses this distance as an input to create clusters (Fig. 1c→d), and the identified clusters become a recipe for creating a map which is then subsequently generalized into regions (Fig. 1d→e). Below, we provide more detail about constructing the supervised tree-ensemble, computing the Hamming distance (i.e., the leaf-distance), conferring cluster assignments, delineating regions from the clusters, and comparing the regions.

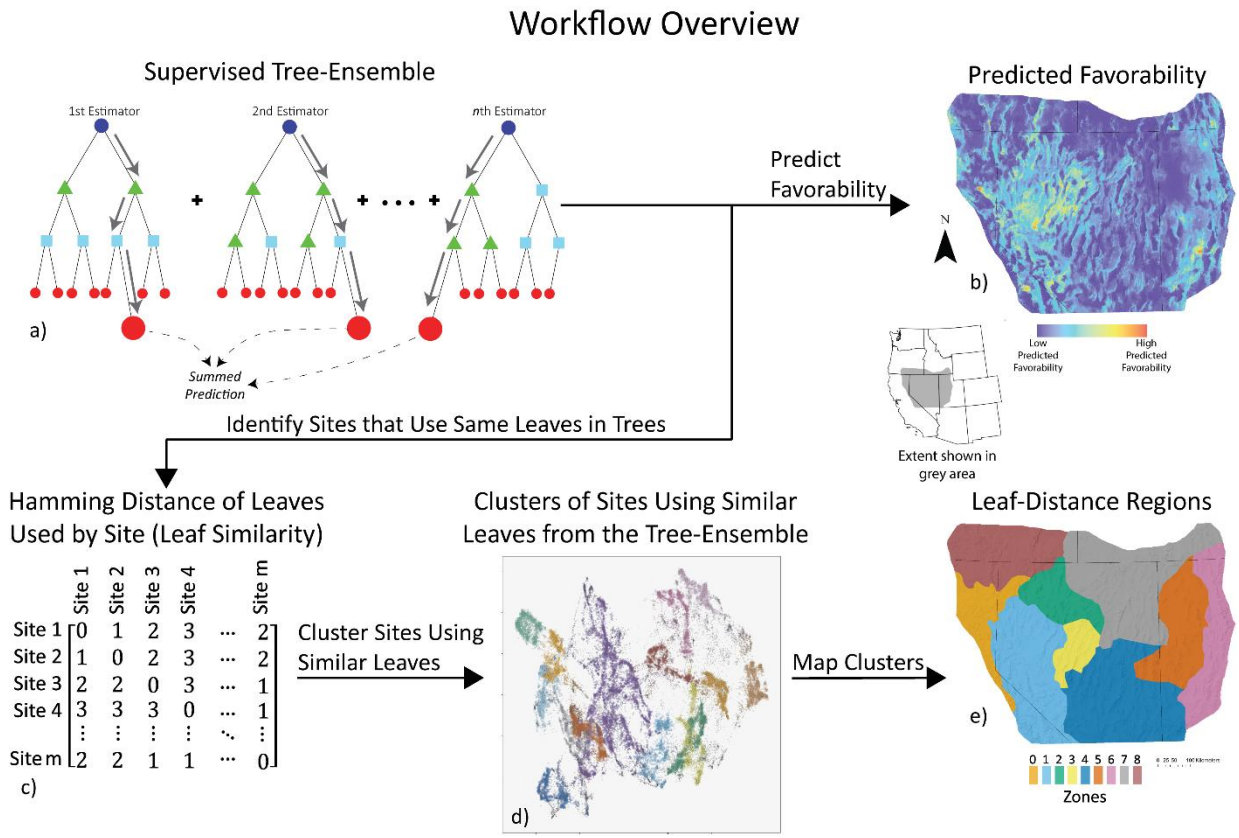


Figure 1. Overview of the leaf-distance workflow. Delineating regions by leaf-distance requires (a) fitting a supervised ML tree-ensemble that can (b) predict hydrothermal favorability for each training site following a path through each tree of the ensemble (e.g., grey arrows in [a]). The similarity of these paths is measured by (c) comparing which training sites use the same leaves (e.g., large red circles in [a]). Then, (d) hierarchical density-based (i.e., non-linear) clustering groups the training sites that use similar leaves, which then allows (e) delineating regions from leaf-distance clusters. In (c) *m* corresponds to the maximum number of training sites. Hillshade from USGS 3D Elevation Program (U.S. Geological Survey, 2024).

2.1 Constructing the Supervised ML Tree-Ensemble

Construction of the favorability map is described in detail in Mordensky et al. (2025b) using Extremely Randomized Trees (ExtraTrees; Geurts et al., 2006) as the supervised ML algorithm, a modeling approach that is shown to be an improvement over Coolbaugh et al. (2005a,b) and Williams and DeAngelo (2008) (Mordensky et al., 2025a,b). The favorability map herein is the result of 600 80/20 train-test splits (i.e., realizations) using 2,192 training sites from DeAngelo (2022, 2023a) as supplemented with a random sampling of negatives first used by Mordensky et al. (2023b). The 53 input features (i.e., evidence layers) are provided with their data references in Table A1 and are compiled into the USGS data release that accompanies this conference paper (Mordensky and DeAngelo, 2026)

2.2 Computing Hamming Distance

We select Hamming distance (Hamming, 1950) as the measure of distance between leaves used by the training sites within the ExtraTrees ensemble.

Hamming distance is the measure of different elements between two vectors of equal dimension (i.e., how many changes are necessary to make two lists of items identical). In the context of our study, the Hamming distance is the count of how many times two sites use different leaves in a single realization (e.g., 0 indicates

that two sites use all the same leaves in the ensemble). We select the 2,192 training sites from DeAngelo et al. (2022, 2023a), in addition to 40,000 of the randomly selected training sites. We then compute the pairwise Hamming distance between all 42,192 sites, resulting in a 42,192-by-42,192 matrix in which each element is the number of times a pair of training sites uses different leaves (refer to Fig. 1c for a conceptual depiction). The Hamming distance matrix for the entire model is the sum of the Hamming distance matrices from the 600 realizations through matrix addition.

The 42,192-by-42,192 matrix for Hamming distance can be viewed as tabular data containing 42,192 training sites. In this light, each row of length 42,192 corresponds to a new feature vector for the corresponding site that can then be used as inputs to a clustering algorithm as discussed in Section 2.3.

2.3 Conferring Cluster Assignments

To identify clusters of training sites using similar leaves, we apply dimensionality reduction to the matrix for Hamming distance and then apply a clustering algorithm.

Having as many fields of data as training sites can create several different problems in ML models (e.g., the curse of dimensionality; Hughes, 1968). When clustering data with too many fields, Euclidean distances between training sites become nearly equidistant. Therefore, we apply uniform manifold approximation and projection (UMAP; Sainburg et al., 2021) for dimensionality reduction. UMAP is a non-linear dimensionality reduction method that preserves local and global structure. That is, UMAP can reduce the number of fields for Hamming distance while preserving the relationships between the training sites. UMAP functions by constructing a high-dimensional graph with dimensionality equal to the number of fields and then looking for the nearest neighboring training sites for each training site to reduce the high dimensionality to lower dimensional embeddings (i.e., the new, fewer dimensions of the data). Having experimented with using 25 to 500 nearest neighbors and reducing to three to 24 embeddings, we elect to reduce the dimensionality of the 42,192 features to three embeddings using 100 nearest neighbors to balance preserving local and global relationships between the training sites.

Using these three embeddings, we cluster the training sites. Recognizing that flat, centroid-based clustering methods like k-means clustering may not capture the complex relationships between training sites, we seek a hierarchical and/or density-based clustering approach. Hence, we select Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBScan; Campello et al., 2013) as the clustering method. HDBScan is a clustering algorithm that groups similar data points based on how densely they are packed together and then builds a hierarchy of clusters by varying density thresholds. Clusters are partially defined by the minimum cluster size (i.e., the smallest size a group must have to be considered a valid cluster), and the minimum samples value (i.e., how many training sites are needed near a point for it to be classified as a training site that can build a cluster), and epsilon (i.e., a radius specifying how far around a training site to search for neighboring training sites to link together). After experimenting with these hyperparameters, we selected a minimum cluster size of 1,000, a minimum samples value of 10, and epsilon of 0.1 to create 15 clusters from the 42,192 points in the study area.

2.4 Delineating of Regions

The clusters identified from HDBScan are examined to identify and outline geographic regions.

HDBScan creates a clustering formula that assigns every provided location on a map to a cluster, but the resulting HDBScan cluster map can be complex, with a region comprised of two or three dominant cluster types. When clusters are presented as distinct colors, this complexity can be thought of as a color pattern (e.g., a red and green pattern in one area, and a blue and purple pattern in another area), so the final step when assigning regions is to generalize the patterned areas. Because relative topography is a controlling input dataset (i.e., mountains are high and valley bottoms are low; refer to Appendix A; Caraccioli et al., 2024, 2025), it will be demonstrated that masking valley bottoms emphasizes bedrock patterns allowing easier delineation of regions. Regions are then delineated manually based on the resulting visual patterns.

2.5 Comparing the Regions

With the regions delineated, we then inspect the occurrence of power-producing systems within each region. We consider the developed power-producing systems from Mlawsky and Ayling (2021) (Fig. 2a) and the identified electric-grade systems described by Williams et al. (2008a,b) available in DeAngelo and Williams (2010) (Fig. 2b) with the corresponding reservoir depths estimated using the three-dimensional temperature model of the Great Basin of Burns et al. (2024a) and the reservoir temperatures from Williams et al. (2008a,b).

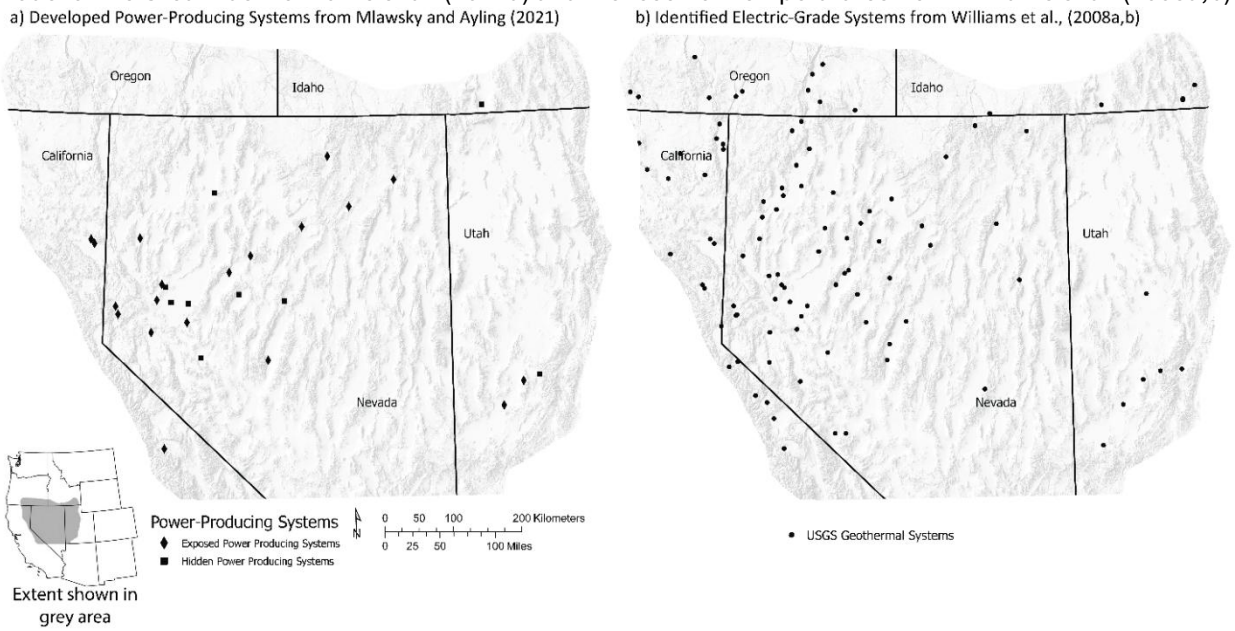


Figure 2. Locations of a) developed power-producing systems from Mlawsky and Ayling (2021) and b) identified electric-grade systems identified by Williams et al. (2008a,b) available in DeAngelo and Williams (2010). Hillshade from USGS 3D Elevation Program (U.S. Geological Survey, 2024).

3 RESULTS

The workflow using ExtraTrees produces a tree-based ensemble model that predicts favorability for hydrothermal resources across the study area (refer to Appendix B for the favorability maps, model fit, and model performance). From the leaf-distances of the 42,192 selected sites, HDBScan identifies 15 distinct clusters (Fig. 3) using three UMAP embeddings of the Hamming distances. These 42,192 points are mapped and colored by cluster (Fig. 4a). Areas between the selected points are filled using the nearest cluster assignment to create a contiguous cluster map (Fig. 4b). Regionalization of patterns may already be apparent to the eye (Fig. 4b), but applying a mask to valley lowlands highlights the patterning in the bedrock uplands, allowing clearer delineation of regions (Fig. 4c), which are then assigned region numbers (Fig. 4d). The regions range in area from 17,643 km^2 (Region 3) to 93,577 km^2 (Region 7) with a median of 50,930 km^2 (Region 5; Fig. 5).

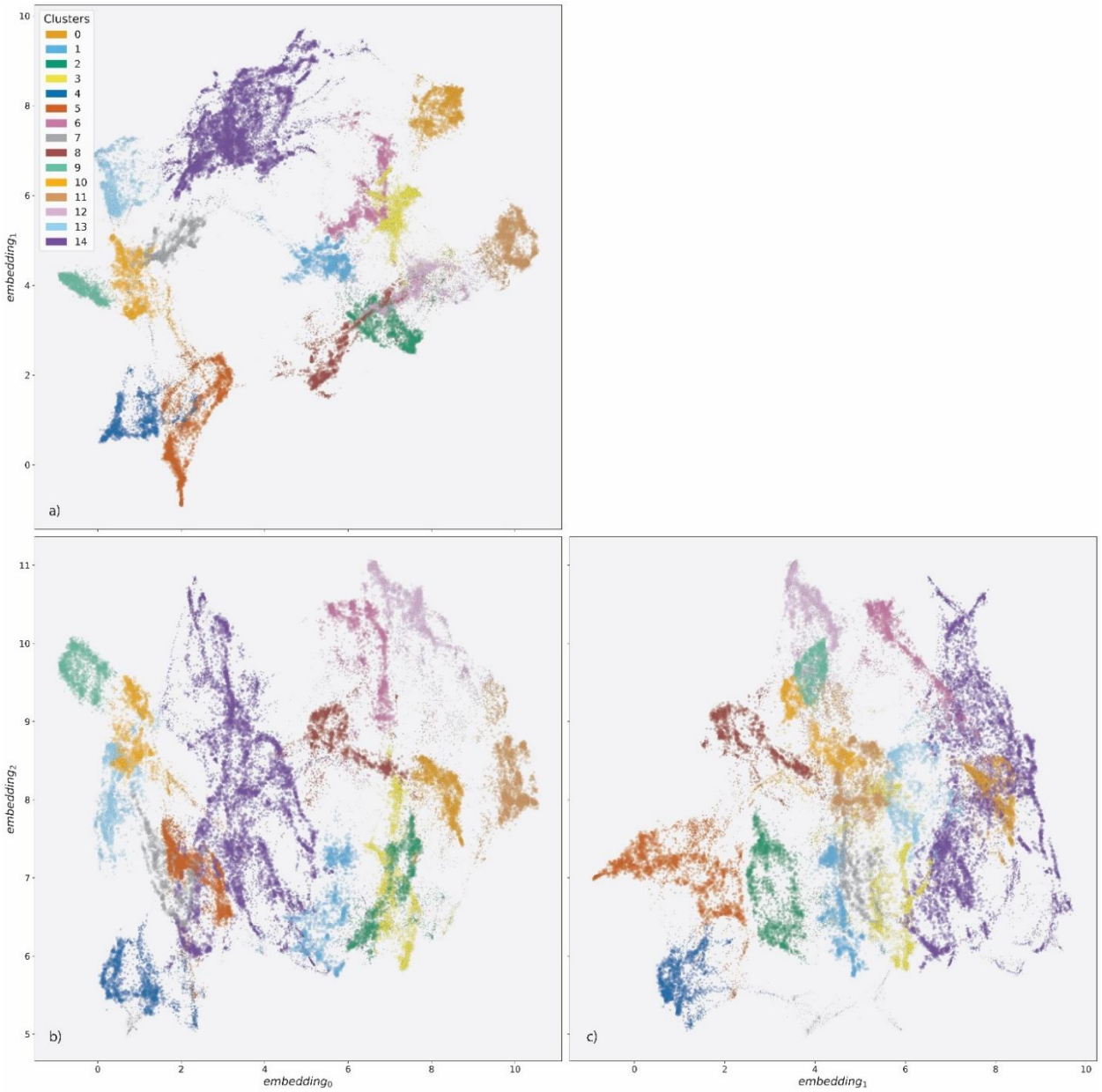


Figure 3. Leaf-distance clusters from HDBScan in unitless embedding space after reducing the dimensionality of the Hamming-distance matrix to three embeddings with UMAP. Colors correspond to clusters identified by HDBScan. The order and color of the clusters do not reflect any relative relation to other clusters.

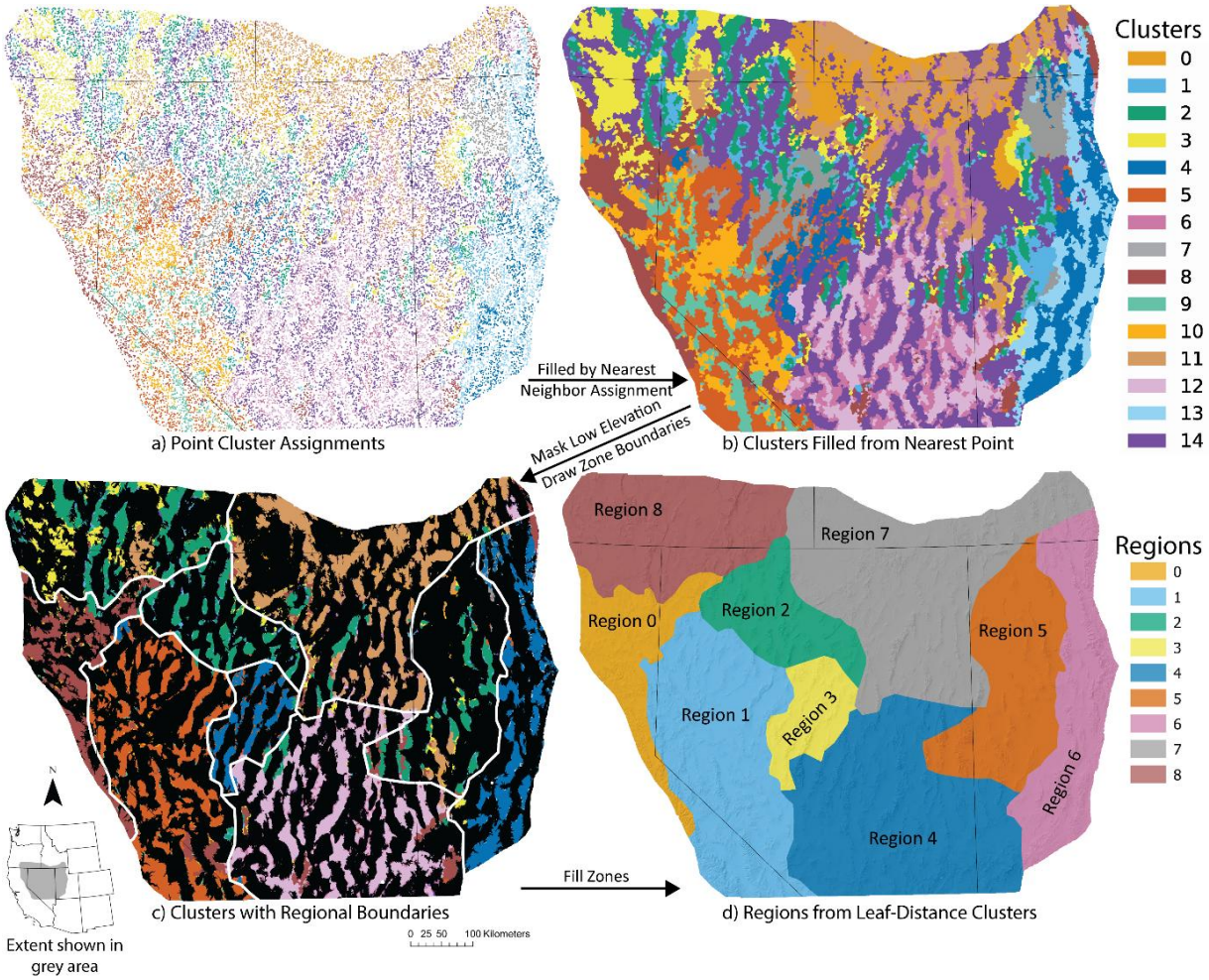


Figure 4. Processing the leaf-distances between the selected 42,192 sites to regions required three processing steps: 1) After HDBScan assigned clusters to the selected sites (refer to Fig. 3), the remaining cells without a cluster assignment were filled by the nearest cell value with an HDBScan cluster assignment (a) to produce a complete map of clusters (b). 2) Because some clusters were predominated by low detrended elevation values (valley lowlands), which are found across the entirety of the map (e.g., Cluster 5), a black mask was applied to areas with low detrended elevation (< 50 m), thereby allowing 3) boundaries (white line) to be drawn by regional qualities (c), leading to the final regional map (d). Cluster colors in Fig 4a,b,c correspond to same clusters in Fig. 3. Hillshade from USGS 3D Elevation Program (U.S. Geological Survey, 2024).

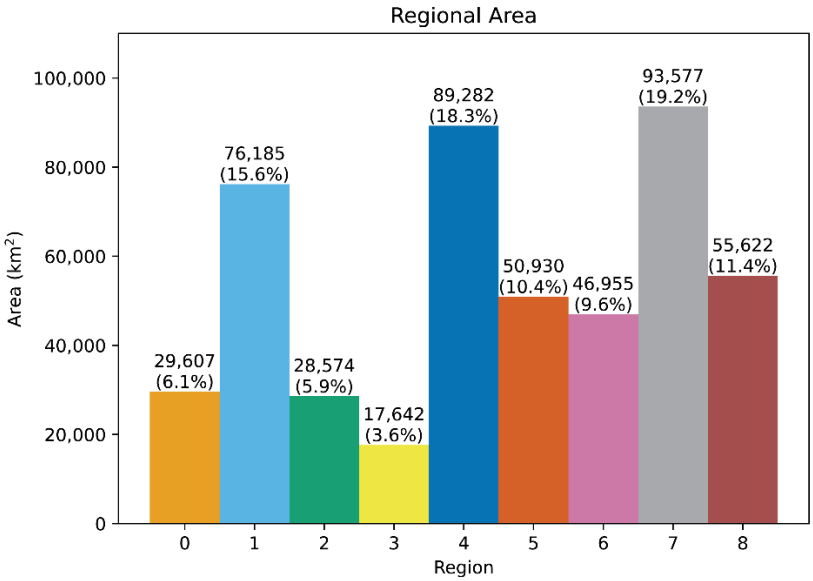


Figure 5. The area of regions. The area (km²) for each region is given as the upper value above each bar. The percentage of the study area for each region is the lower value in parentheses. Region colors correspond to the same regions in Fig. 4d.

4 DISCUSSION

It might be hypothesized that these regions represent hydrothermal systems controlled by different (albeit perhaps mostly similar) geologic characteristics. Under this hypothesis, the properties of hydrothermal systems in one region should be different than the properties of hydrothermal systems in another region (e.g., more, bigger, and/or hotter). If this hypothesis is true, the resulting regional map could prove useful when prospecting, assessing favorability, or computing conditional statistics to improve resource assessments (e.g., individually considering the known resources in a decile of predictions respective to each region when estimating the undiscovered resources rather than considering resources in a decile of favorability predictions from the entire study area).

The regions delineated by leaf-distance (Fig. 4d) show value for conditional statistics with the regions having different resource properties. For example, developed resources (Figs. 6a, 7a, 8) and identified electric-grade resources (Figs. 6b, 7b, 8) vary by region. Region 1, much of which coincides with Walker Lane as defined by Faulds and Henry (2008), contains the majority of the developed systems and a plurality of the identified systems (Figs. 6, 7, 9). The general density of identified electric-grade systems varies not only substantially by region (Figs. 6, 7, 9), but reservoir properties (i.e., reservoir temperature (Figs. 9a, 10a), reservoir volume (Figs. 9c, 10c), circulation depth (Figs. 9d, 10d) also vary by region. These differing qualities (i.e., the density of hydrothermal systems and reservoir properties) between regions indicate that the conditional probabilities of resource availability will also vary between regions. Consequently, the estimated power potential varies too (Fig. 8, 9b, 10b).

The relative abundance of hydrothermal systems in Region 1 is likely at least partially attributed to the relatively higher strain of the region (refer to Zeng, 2022; Kreemer and Young, 2023) because strain facilitates active deformation, which contributes to the permeability of active hydrothermal systems (e.g., Faulds and Hinz, 2015; Kennedy et al., 2020). Similarly, the higher strain rate of Region 1 may contribute to hotter, larger-volume systems than those found in other regions (Figs. 9a,c, 10a,c) and these hotter, larger-volume systems thereby contribute to a higher density of systems with greater power potential than found in the other regions (Figs. 8, 9b, 10b).

Curiously, three regions identified by Williams et al. (2008a,b) as having several identified electric-grade systems have no presently developed power-producing systems (i.e., Regions 0, 5, and 8, corresponding respectively to the eastern margin of the Sierra Nevada Batholith, the central eastern Great Basin, and the Modoc Plateau) with two of these regions (i.e., Regions 0 and 8) estimated as containing power potential exceeding several 100s of MWe (Figs. 8, 9b, 10b). The reason for the discrepancy between the developed and identified power production in Region 0 may relate to its systems generally being cooler and smaller than regions with greater developed capacity (e.g., Region 1; Figs. 9a,c, 10a,c). The estimated reservoir temperatures and volumes in Region 8, however, more closely resemble the systems of Region 1 than Region 0, suggesting that qualities other than system temperature and volume disincentivize development there. Fundamentally, we ask — *what are the driving mechanisms differentiating the regions?* Measuring the importance of the input features on a region-by-region basis could aid in identifying which input features have differing importance in the different regions.

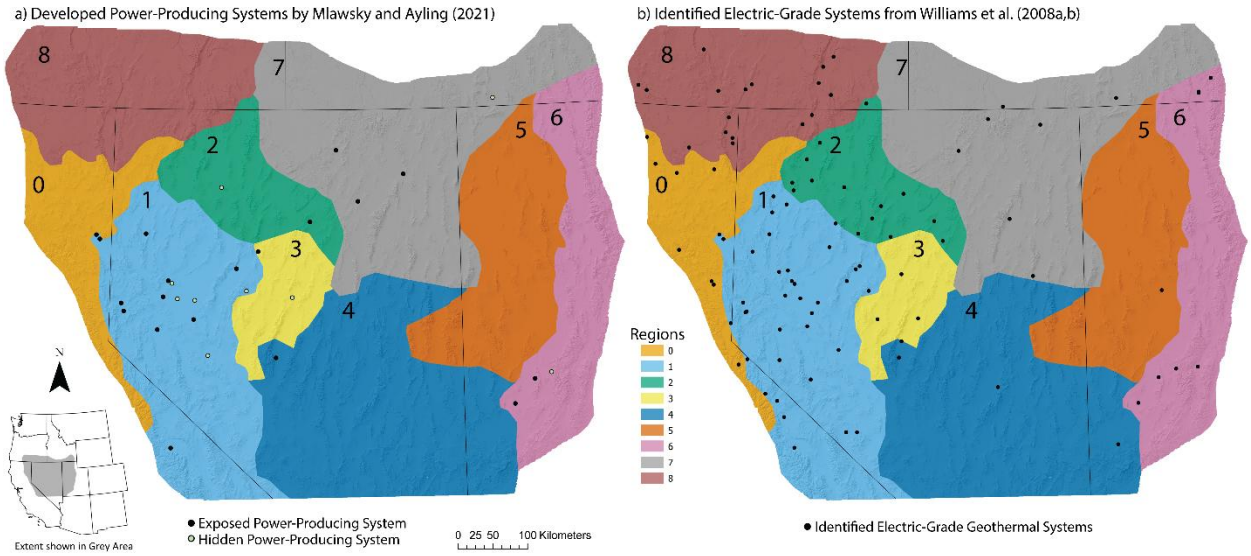


Figure 6. Distributions of a) developed power-producing systems (Mlawsky and Ayling, 2021) and b) identified electric-grade systems by region from Williams et al. (2008a,b) and available in DeAngelo and Williams (2010). Region colors correspond to the same regions in Fig. 4d. Overlaid numbers identify region. Hillshade from USGS 3D Elevation Program (U.S. Geological Survey, 2024).

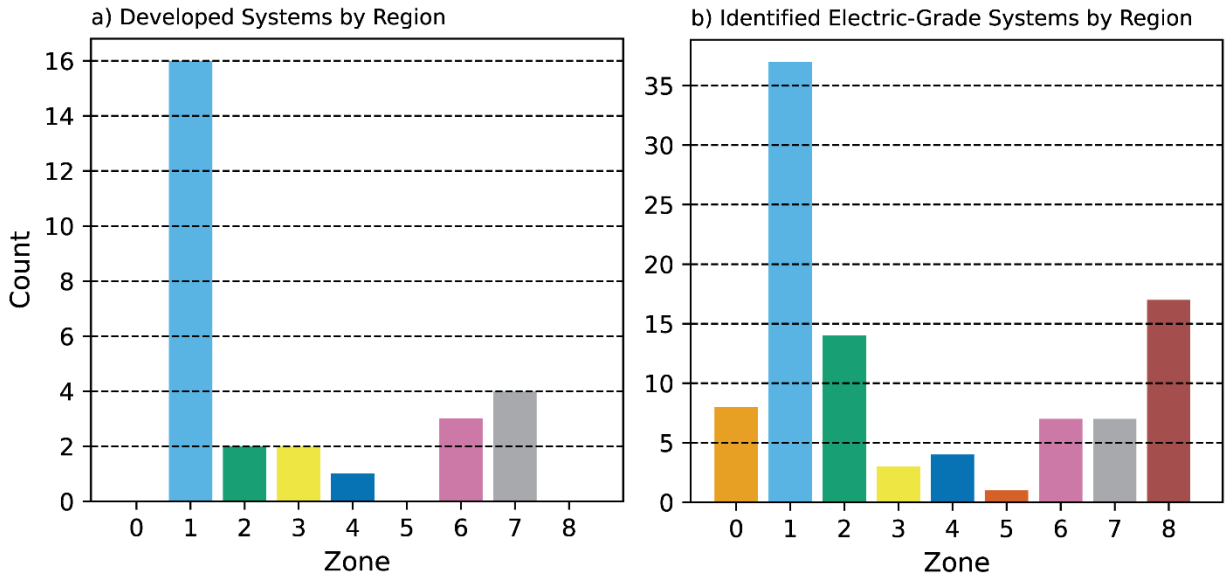


Figure 7. Frequency of a) developed power-producing systems by Mlawsky and Ayling (2021) and b) identified electric-grade systems from Williams et al. (2008a,b) available in DeAngelo and Williams (2010) by region. Region colors correspond to the same regions in Fig. 4d.

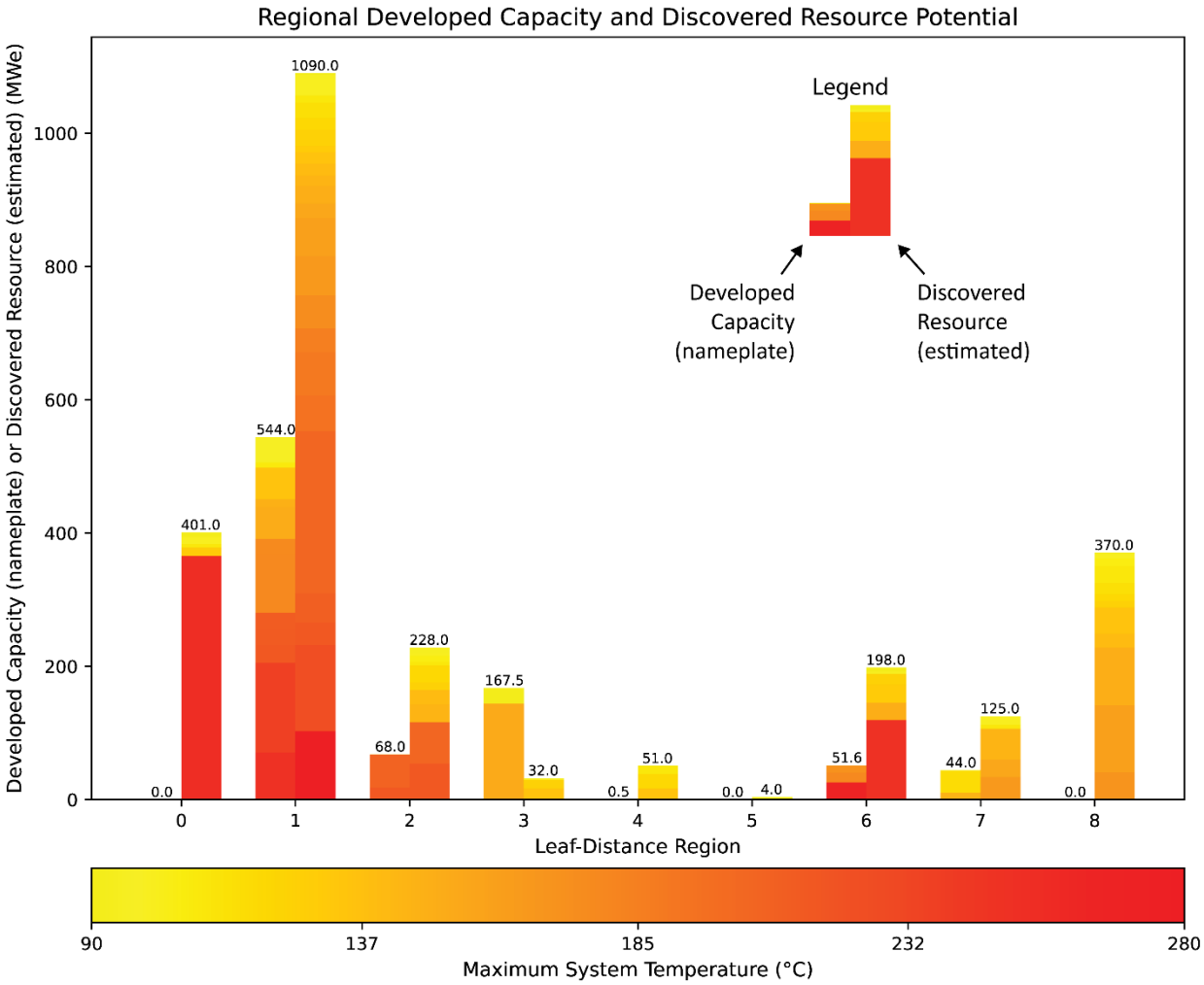


Figure 8. Comparison of developed capacity (left bar; Mlawsky and Ayling, 2021) versus estimated resource (right bar; from Williams et al. (2008a,b) available in DeAngelo and Williams, 2010) by region. Color scale provides the estimated megawatts electric (MWe) of resource at its maximum estimated temperature. The cumulative developed capacity and estimated discovered resources are rounded to the nearest tenth of a MWe and provided at the top of the bars. Note that Region 3 has more developed capacity than estimated discovered resource because of the discovery and development of the McGinness Hills geothermal system since the completion of the 2008 USGS geothermal resource assessment.

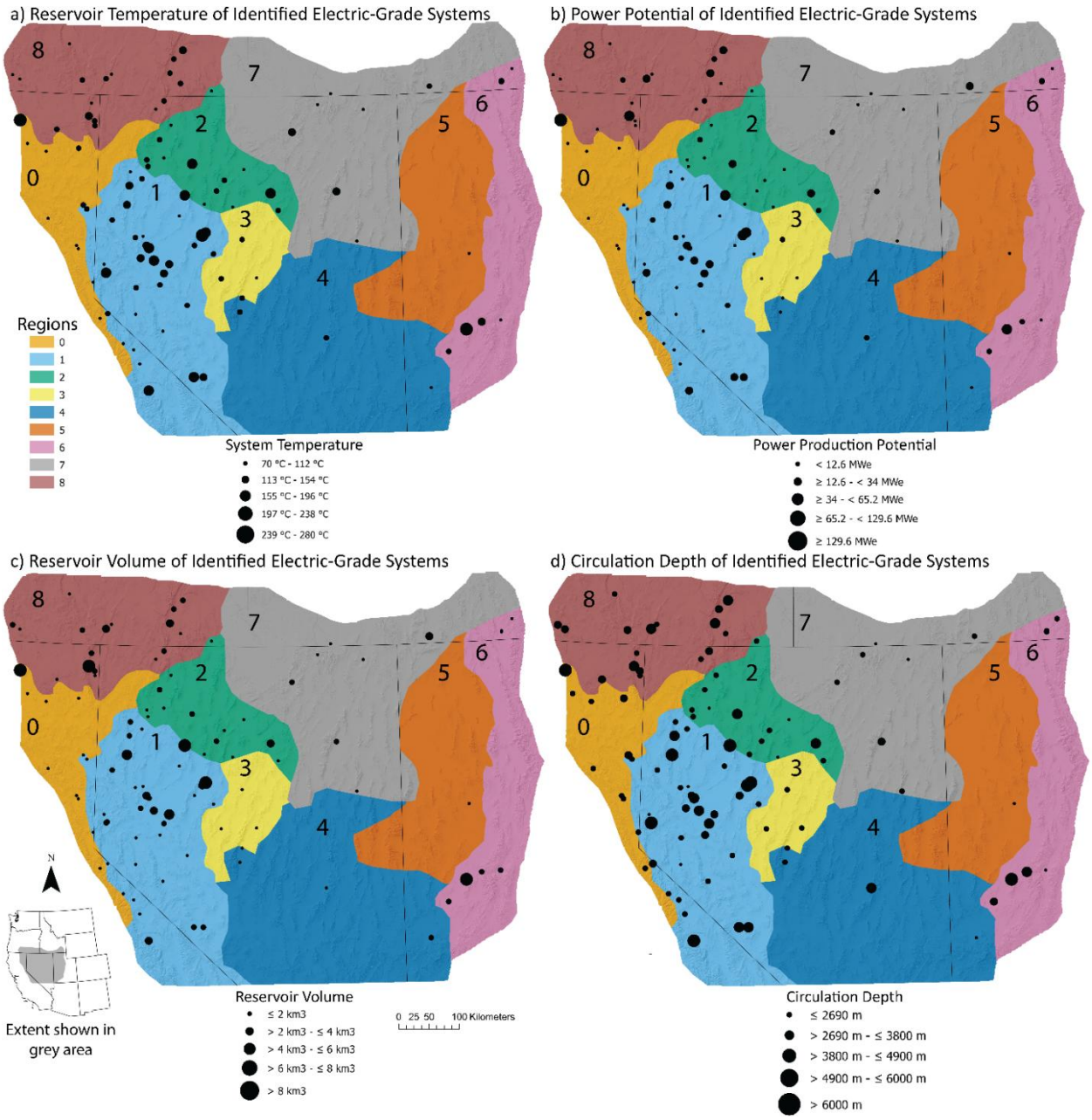


Figure 9. Estimates of most likely (i.e., mean) a) reservoir temperature, b) power production potential, and c) reservoir volume for identified electric-grade systems from Williams et al. (2008a,b) available in DeAngelo and Williams (2010) with d) estimated circulation depth using the method of Burns et al. (2024a) based on the reservoir temperature. Region colors correspond to the same regions in Fig. 4d. Overlaid numbers identify region. Hillshade from USGS 3D Elevation Program (U.S. Geological Survey, 2024).

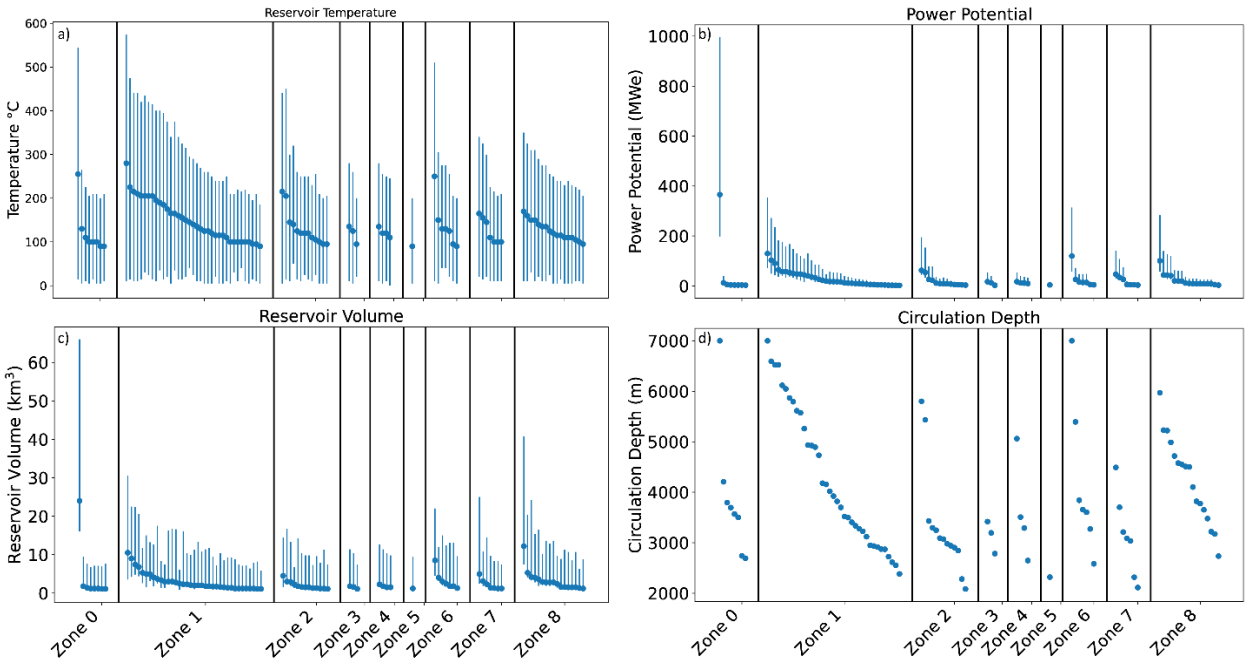


Figure 10. Low (farthest extent of bottom whisker), most likely (point), and high (furthest extent of top whisker) estimates of a) system temperature, b) production potential, c) reservoir volume, and d) circulation depth for known geothermal systems in descending order by region. Most likely estimates are the mean estimates from Williams et al. (2008a,b) available in DeAngelo and Williams (2010). Low and high estimates correspond respectively to the minimum and maximum system estimates for system temperature and reservoir volume and the 5th and 95th percentile estimates for circulation depth. The circulation depth is estimated using the three-dimensional temperature model of the Great Basin from Burns et al. (2024a) based on the reservoir temperature. Circulation depths greater than 7,000 m are truncated to 7,000 m.

5 CONCLUSION

In preparation for an updated geothermal resource assessment of the Great Basin, we present a workflow for delineating regions that can reveal key differences in resource occurrence or type, by using non-linear clustering informed by the supervised tree-ensemble (Extremely Randomized Trees [ExtraTrees]) machine learning (ML) algorithm. Using a supervised ML algorithm allows delineation of regions by geologic characteristics that control hydrothermal upflow. This novel approach divides the Great Basin into nine regions with varying resource characteristics (e.g., frequency of occurrence and reservoir properties) by identifying which training sites use the same branching paths in the trees constituting the tree-ensemble. The majority of the currently developed power-producing systems and plurality of the identified electric-grade hydrothermal systems from Williams et al. (2008a,b) are found within one region. Two of the three regions with the most discovered resources have no developed power capacity. The identified regions may include areas where different strategies may need to be employed to develop these resources. This research mapping regions in the Great Basin may lend new insights into the conditional statistics when assessing geothermal resources.

APPENDIX A – INPUT FEATURES

We use 53 input features (i.e., independent variables or evidence layers) to delineate the regions of the Great Basin as defined by leaf-distance from a supervised machine learning (ML) tree-ensemble (Table A1). For completeness and convenience, these input features are compiled into the USGS data release that accompanies this conference paper (Mordensky and DeAngelo, 2026). Distances to the nearest Quaternary fault or magmatic body were log transformed using Equation A1 to allow the ML algorithm to place greater emphasis on distances nearer these geologic phenomena than farther away under the assumption that, for

example, being 3 km versus 7 km from a fault or magmatic body is more likely to have an effect on hosting a hydrothermal system than being 30 km versus 34 km from a fault or magmatic body.

$$\begin{aligned} transformed\ distance \\ = (0.5200314775450414 \\ \times \log_{1.4695608302071574} distance\ in\ meters + 125) - 6.5 \end{aligned} \quad \text{(A1)}$$

Table A1: The input features used in this study and their sources.

Input Feature	Source
Conductive Heat Flow	DeAngelo et al. (2022, 2023a)
First Derivative of Conductive Heat Flow	Modified from DeAngelo et al. (2022, 2023a)
Curie Depth	Bouligand et al. (2009)
Electrical Surface Conductance (2 - 12 km)	Peacock and Bedrosian (2022)
Electrical Middle Crust Conductance (12 - 20 km)	Peacock and Bedrosian (2022)
Electrical Lower Crust Conductance (20 - 50 km)	Peacock and Bedrosian (2022)
Electrical Upper Mantle Conductance (50 - 90 km)	Peacock and Bedrosian (2022)
Electrical Mantle Conductance (90 - 200 km)	Peacock and Bedrosian (2022)
Distance to Nearest Quaternary Magmatic Activity	Ayling et al. (2022)
Magnetic Field	Glen et al. (2022)
First Derivative of Magnetic Field	Modified from Glen et al. (2022)
Isostatic Gravity	Glen et al. (2022)
First Derivative of Isostatic Gravity	Glen et al. (2022)
Depth to Basement	Glen et al. (2022)
Aphi Stress Indicator	Lundstern (2022, 2023)
Distance to Nearest Quaternary Fault	Ayling et al. (2022)
Distance to Nearest Fault < 130 ka	Ayling et al. (2022)
Distance to Nearest Fault < 15 ka	Ayling et al. (2022)
DEM (30 m)	USGS (2023)
Detrended Topography (span = 0.001)	DeAngelo et al. (2023b, 2023c)
Detrended Topography (span = 0.005)	DeAngelo et al. (2023b, 2023c)
Detrended Topography (span = 0.01)	DeAngelo et al. (2023b, 2023c)
Detrended Topography (span = 0.03)	DeAngelo et al. (2023b, 2023c)
Elevation Trend (span = 0.001)	DeAngelo et al. (2023b, 2023c)
Elevation Trend (span = 0.005)	DeAngelo et al. (2023b, 2023c)
Elevation Trend (span = 0.01)	DeAngelo et al. (2023b, 2023c)
Elevation Trend (span = 0.03)	DeAngelo et al. (2023b, 2023c)
Shear Strain Rate	Zeng (2022)
Shear (Conditional Extension) Strain Rate	Zeng (2022)
Second Invariant of Strain-Rate Tensor	Zeng (2022)
Dilation Strain Rate	Zeng (2022)
Dilation (Extensional Only) Strain Rate	Zeng (2022)
Shear Strain Rate	Kreemer and Young (2023)
Dilation Strain Rate	Kreemer and Young (2023)
Second Invariant of Strain-Rate Tensor	Kreemer and Young (2023)
Independent Seismic Density (> 2M, ≤ 30-km depth, n = 200, α = 0.15)	Kreemer and Young (2023)
Independent Seismic Density (> 2M, ≤ 30-km depth, n = 200, α = 0.10)	Kreemer and Young (2023)
Independent Seismic Density (> 2M, ≤ 30-km depth, n = 200, α = 0.05)	Kreemer and Young (2023)

Independent Seismic Density (> 2M, ≤ 30-km depth, n = 100, α = 0.15)	Kreemer and Young (2023)
Independent Seismic Density (> 2M, ≤ 30-km depth, n = 100, α = 0.10)	Kreemer and Young (2023)
Independent Seismic Density (> 2M, ≤ 30-km depth, n = 100, α = 0.05)	Kreemer and Young (2023)
Independent Seismic Density (> 2M, ≤ 30-km depth, n = 50, α = 0.15)	Kreemer and Young (2023)
Independent Seismic Density (> 2M, ≤ 30-km depth, n = 50, α = 0.10)	Kreemer and Young (2023)
Independent Seismic Density (> 2M, ≤ 30-km depth, n = 50, α = 0.05)	Kreemer and Young (2023)
Foreshock/Aftershock Seismic Density (> 2M, ≤ 30-km depth, n = 200, α = 0.15)	Kreemer and Young (2023)
Foreshock/Aftershock Seismic Density (> 2M, ≤ 30-km depth, n = 200, α = 0.10)	Kreemer and Young (2023)
Foreshock/Aftershock Seismic Density (> 2M, ≤ 30-km depth, n = 200, α = 0.05)	Kreemer and Young (2023)
Foreshock/Aftershock Seismic Density (> 2M, ≤ 30-km depth, n = 100, α = 0.15)	Kreemer and Young (2023)
Foreshock/Aftershock Seismic Density (> 2M, ≤ 30-km depth, n = 100, α = 0.10)	Kreemer and Young (2023)
Foreshock/Aftershock Seismic Density (> 2M, ≤ 30-km depth, n = 100, α = 0.05)	Kreemer and Young (2023)
Foreshock/Aftershock Seismic Density (> 2M, ≤ 30-km depth, n = 50, α = 0.15)	Kreemer and Young (2023)
Foreshock/Aftershock Seismic Density (> 2M, ≤ 30-km depth, n = 50, α = 0.10)	Kreemer and Young (2023)
Foreshock/Aftershock Seismic Density (> 2M, ≤ 30-km depth, n = 50, α = 0.05)	Kreemer and Young (2023)

APPENDIX B – FAVORABILITY MAPS, MODEL FIT, AND MODEL PERFORMANCE

ExtraTrees predicts 89.69%, 10.31%, and <0.01% of the study area as having low, high, and very high hydrothermal favorability, respectively (Fig. B1). The favorability map from ExtraTrees predicts the western, southwestern, and southeastern Great Basin as having high favorability and the south-central and central-eastern Great Basin as having low favorability (Fig. B1a). High favorability generally follows high-granularity lineaments along the interface of basins and ranges. Low-favorability terrain is predicted across every cardinal extent of the Great Basin at high detrended elevation values (Fig. B1b).

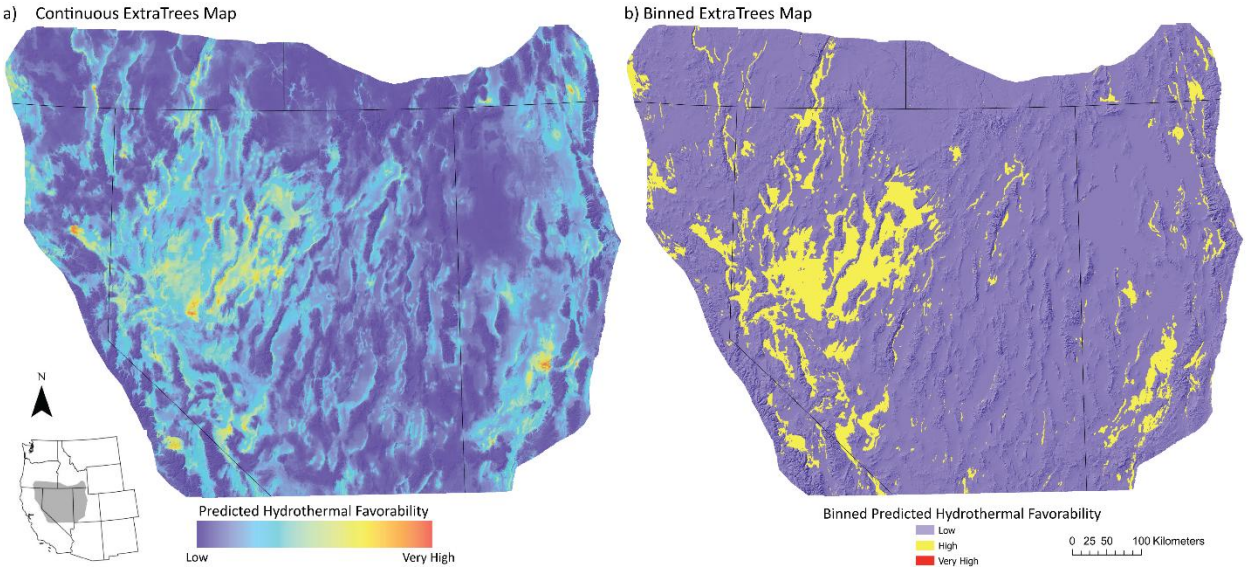


Figure B1. Continuous (a) and binned (b) favorability maps from ExtraTrees derived from the median predictions of the 600 train-test splits (i.e., realizations). Cooler colors (blue) represent lower hydrothermal favorability. Warmer colors (red) represent high favorability. Hillshade from USGS 3D Elevation Program (U.S. Geological Survey, 2024).

To evaluate model fit and performance, we compare the area under the curve (AUC) for developed power-producing systems. Graphically plotting each of the developed power-producing systems in order of their descending prediction percentiles with the area under that resulting curve as the AUC value provides a numeric metric of model performance that is a measure for which model predicts developed power-producing

systems in the smallest, most-favorable fraction of the map. When the predictions for the developed power-producing systems are from the favorability map, a higher AUC indicates that the model predicts developed power-producing systems in higher percentiles. By comparing the AUC from the favorability map to the AUC when the developed power-producing systems are used only as testing data, we can evaluate if a model is overfit because 80% of the grid predictions at developed power-producing systems in the favorability map are from training data and the remaining 20% are from testing data. An overfit model will predict substantially more developed power-producing systems with higher favorability in the training data than in the testing data. We compare our results to legacy models that cover our study area (i.e., models from Coolbaugh et al., 2005a,b and Williams and DeAngelo, 2008). Developed power-producing systems are identified using two-letter initials for the new modeling results, but for clarity, only lines are shown for the 57 previous models (56 from Williams and DeAngelo [2008] and 1 from Coolbaugh et al. [2005a,b]).

The favorability map from the ExtraTrees model predicts 27 of the 28 developed power-producing systems as having high or very high favorability (i.e., above the 89th prediction percentile; Fig. B2a). The AUC for the favorability map (i.e., the AUC when the developed power-producing systems are used as training and testing data; Fig. B2a) is 98.34. To inspect for overfitting, we compare the AUC for the favorability to the AUC of the same developed power-producing systems when they are used strictly in the testing data (93.62; Fig. B2b).

Given the scarcity of the hydrothermal resources, a perfect model would predict developed power-producing systems in the top 1% of predictions. A perfect model would also predict equally well with data used during training and data not used during training. In these regards, the workflow using 53 input features we present herein outperforms previous models for the region (e.g., from Mordensky et at., 2025a,b) by producing high AUC values for developed power-producing systems using predictions from both the favorability map (98.34; Fig. B2a) and the testing data (93.62; Fig. B2b).

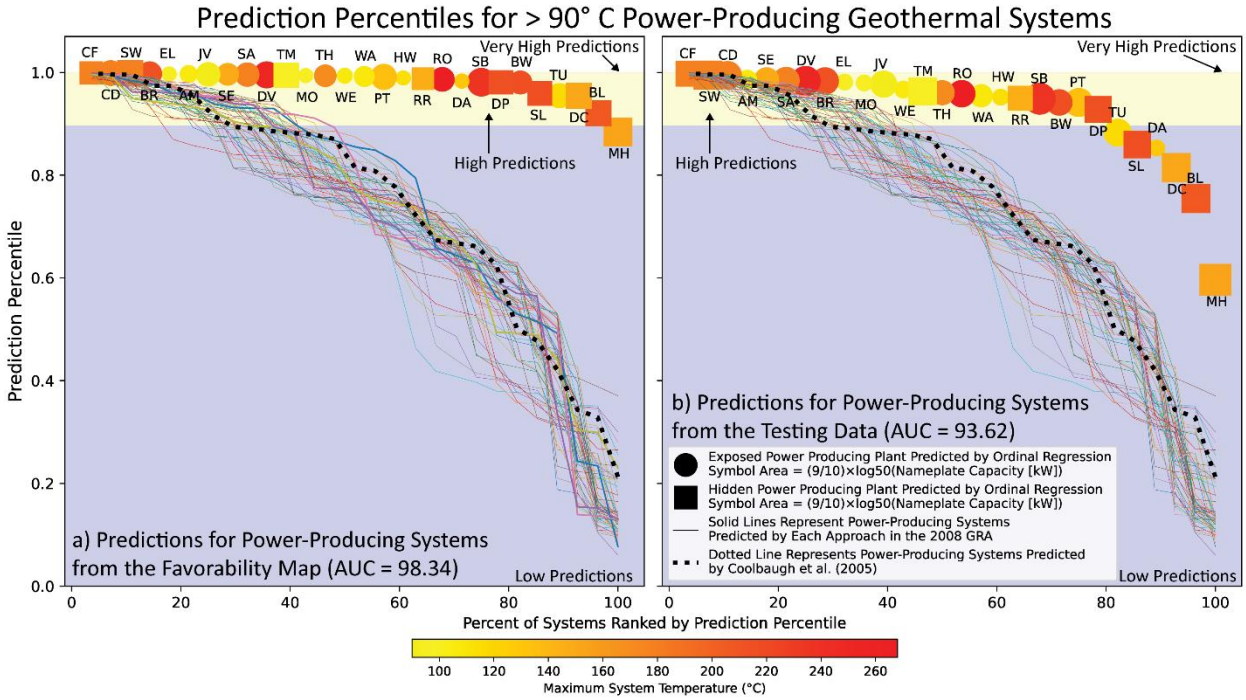


Figure B2. Predictive skill for developed power-producing systems when developed power-producing systems are (a) from the favorability map (i.e., used in the training and testing data) and (b) from the testing data. The difference in AUC between a) and b) (4.72 = 98.34 – 93.62) suggests some minor overfitting. Predictions from previous modeling studies (i.e., Coolbaugh et al. 2005a,b and Williams and DeAngelo 2008) are provided for comparison. The yellow background corresponds to prediction

percentiles predicted as high and very high favorability. The blue background corresponds to prediction percentiles predicted as low favorability. Circles represent the prediction quantiles for exposed developed power-producing systems from the model we present. Squares represent the prediction percentiles for hidden developed power-producing systems from the model we present. Symbol area corresponds to the $\log_{50}$ of the nameplate capacity in kW. Symbol color corresponds to the maximum recorded or modeled temperature at that developed power-producing system. Each solid line represents one of the 56 models from Williams and DeAngelo (2008). The dotted line represents the model from Coolbaugh et al. (2005a,b). The developed power-production data are available through Mlawsky and Ayling (2021). Abbreviations: AM – Amedee; BL – Blue Mountain; BR – Brady's; BW – Beowawe; CD – Casa Diablo; CF – Cove Fort; DA – Darroughs Hot Spring; DC – Don Campbell; DP – Desert Peak; DV – Dixie Valley; EL – Elko Hot Springs; HW – Humboldt Wells; JV – Jersey Valley Hot Springs; MH – McGinness Hills; MO – Moana; PT – Patua Hot Springs; RO – Roosevelt Hot Springs; RR – Raft River; SA – Salt Wells; SB – Steamboat; SE – San Emidio; SL – Soda Lake; SW – Stillwater; TH – Thermo Hot Springs; TM – Tungsten Mountain; TU – Tuscarora; WA – Wabuska; WE – Wendel.

Acknowledgements

This work was supported by the U.S. Geological Survey Energy Resources Program, the U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy (EERE), Geothermal Technologies Office (GTO) under Contract No. DEAC02-05CH11231 with Lawrence Berkeley National Laboratory, Conformed Federal Order No. 7520443 between Lawrence Berkeley National Laboratory and the U.S. Geological Survey (Award Number DE-EE0008105), and Standard Research Subcontract No. 7572843 between Lawrence Berkeley National Laboratory and Portland State University. Additional support for John Lipor was provided by the National Science Foundation award NSF CAREER CIF-2046175. We appreciate the data made available through the INnovative Geothermal Exploration through Novel Investigations of Undiscovered Systems project (INGENIOUS), which is supported by the U.S. Department of Energy - Geothermal Technologies Office under award DE-EE0009254 to the University of Nevada, Reno. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

DATA AVAILABILITY

The data for the input features, clusters, and regions are available through a U.S. Geological Survey data release available at https://doi.org/10.5066/P14QM39R.

REFERENCES

Ahmmed, B., Vesselinov, V., Mudunuru, M., Middleton, R., Fractures, S., 2021. Machine Learning on the Geochemical Characteristics of Low-, Medium-, and Hot-Temperature Geothermal Resources in the Great Basin, USA, World Geothermal Congress. Reykjavik, Iceland.

Ahmmed, B., Vesselinov, V.V., 2022. Machine Learning and Shallow Groundwater Chemistry to Identify Geothermal Prospects in the Great Basin, USA. Renewable Energy 197, 1034-1048.

Ayling, B., Faulds, J., Rivera, A.M., Koehler, R., Kreemer, C., Mlawsky, E., Coolbaugh, M., Micander, R., dePolo, C., Kraal, K., Wagoner, N., Siler, D., DeAngelo, J., Glen, J., Batir, J., Gentry, E., Berti, C., Lifton, Z., Clark, A., Kirby, S., Hardwick, C., Kleber, E., 2022. Innovative Geothermal Exploration through Novel Investigations of Undiscovered Systems (Ingenious) - Great Basin Regional Dataset Compilation. Great Basin Center for Geothermal Energy (GBCGE), Nevada Bureau of Mines and Geology (NBMG), University of Nevada, Reno (UNR), https://gdr.openei.org/submissions/1391.

Bouligand, C., Glen, J.M.G., Blakely, R.J., 2009. Mapping Curie Temperature Depth in the Western United States with a Fractal Model for Crustal Magnetization. Journal of Geophysical Research: Solid Earth 114, 1-25.

Brooks, L.E., Masbruch, M.D., Sweetkind, D.S., Buto, S.G., 2014. Steady-State Numerical Groundwater Flow Model of the Great Basin Carbonate and Alluvial Aquifer System, Scientific Investigations Report 2014–5213. U.S. Geological Survey, Reston, VA.

Brown, S., Smith, C.M., 2022. Python Codebase and Jupyter Notebooks - Applications of Machine Learning Techniques to Geothermal Play Fairway Analysis in the Great Basin Region, Nevada. https://gdr.openei.org/submissions/1402.

Brown, S.R., Coolbaugh, M.F., DeAngelo, J., Faulds, J.E., Fehler, M., Gu, C., Queen, J.H., Treitel, S., Smith, C.M., Mlawsky, E., 2020. Machine Learning for Natural Resource Assessment: An Application to the Blind Geothermal Systems of Nevada, Geothermal Resources Council. Geothermal Resources Council Transactions, Virtual.

Burns, E.R., DeAngelo, J., Williams, C.F., 2024a. Three-Dimensional Temperature Model of the Great Basin, USA. U.S. Geological Survey Data Release, https://doi.org/10.5066/P149FR54.

Burns, E.R., DeAngelo, J., Williams, C.F., 2024b. Updated Three-Dimensional Temperature Maps for the Great Basin, USA, Stanford Workshop on Geothermal Reservoir Engineering. Stanford University, Stanford, CA.

Burns, E.R., Williams, C.F., DeAngelo, J., 2025. Enhanced Geothermal Systems Electric-Resource Assessment for the Great Basin, Southwestern United States, USGS Fact Sheet. U.S. Geological Survey, Reston, VA, p. 4.

Campello, R.J.G.B., Moulavi, D., Sander, J., 2013. Density-Based Clustering Based on Hierarchical Density Estimates. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 160-172.

Caraccioli, P.D., Mordensky, S.P., DeAngelo, J., Burns, E.R., Lipor, J.J., 2024. Separating Signals in Elevation Data Improves Supervised Machine Learning Predictions for Hydrothermal Favorability, Geothermal Resources Council Transactions.

Caraccioli, P.D., Mordensky, S.P., DeAngelo, J., Burns, E.R., Lipor, J.J., 2025. How Selecting Representative Training Sites and Features Can Improve Machine Learning Models for Hydrothermal Resource Favorability. Geothermics.

Caraccioli, P.D., Mordensky, S.P., Lindsey, C.R., DeAngelo, J., Burns, E.R., Lipor, J.J., 2023. Don't Let Negatives Hold You Back: Accounting for Underlying Physics and Natural Distributions of Hydrothermal Systems When Selecting Negative Training Sites Leads to Better Machine Learning Predictions, Geothermal Resources Council Transactions. pp. 1672-1693.

Coolbaugh, M., Zehner, R., Kreemer, C., Blackwell, D., Oppliger, G., 2005. A Map of Geothermal Potential for the Great Basin, USA: Recognition of Multiple Geothermal Environments, Geothermal Resource Council Transactions. pp. 223 - 228.

Coolbaugh, M.F., Zehner, R.E., Blackwell, D., Oppliger, G., Sawatzky, D., Blewitt, G., Pancha, A., Richards, M., Helm-Clark, C., Shevenell, L., Raines, G., Johnson, G., Minor, T., Boyd, T., 2005. Geothermal Potential Map of the Great Basin Region, Western United States. Nevada Bureau of Mines and Geology.

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.

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. U.S. Geological Survey Data Release, https://doi.org/10.5066/P9BZPVUC.

DeAngelo, J., Burns, E.R., Gentry, E., Batir, J.F., Lindsey, C.R., Mordensky, S.P., 2023a. New Maps of Conductive Heat Flow in the Great Basin, USA: Separating Conductive and Convective Influences, 48th Workshop on Geothermal Reservoir Engineering. Stanford, CA, USA.

DeAngelo, J., Burns, E.R., Mordensky, S.P., Lindsey, C.R., 2023b. Detrending Great Basin Elevation to Identify Structural Patterns for Identifying Geothermal Favorability, Geothermal Resources Council Transactions. Reno, NV, USA, pp. 1694-1702.

DeAngelo, J., Burns, E.R., Mordensky, S.P., Lindsey, C.R., 2023c. Maps of Elevation Trend and Detrended Elevation for the Great Basin, USA. U.S. Geological Survey Data Release, https://doi.org/10.5066/P9MQRCBY.

DeAngelo, J., Williams, C.F., 2010, Identified Moderate and High Temperature Geothermal Systems: U.S. Geological Survey data release, https://doi.org/10.5066/P1YJWMFC.

Dickinson, W.R., 2006. Geotectonic Evolution of the Great Basin. Geosphere 2(7), 353-368.

Faulds, J.E., Brown, S., Coolbaugh, M., Deangelo, J., Queen, J.H., Treitel, S., Fehler, M., Mlawsky, E., Glen, J.M., Lindsey, C., Burns, E., Smith, C.M., Gu, C., Ayling, B.F., 2020. Preliminary Report on Applications of Machine Learning Techniques to the Nevada Geothermal Play Fairway Analysis, 45th Workshop on Geothermal Reservoir Engineering. Stanford, California, pp. 1-6.

Faulds, J.E., Brown, S.R., Smith, C.M., Queen, J.H., Treitel, S., 2021a. Machine Learning Model Geotiffs - Applications of Machine Learning Techniques to Geothermal Play Fairway Analysis in the Great Basin Region, Nevada. Geothermal Data Repository, https://gdr.openei.org/submissions/1351.

Faulds, J.E., Coolbaugh, M., Blewitt, G., Henry, C.D., 2004. Why Is Nevada in Hot Water? Structural Controls and Tectonic Model of Geothermal Systems in the Northwestern Great Basin, Geothermal Resources Council. Geothermal Resources Council Transactions, 24, pp. 649-654. Indian Wells, California. 29 August – 1 September 2004.

Faulds, J.E., Craig, J.W., Hinz, N.H., Coolbaugh, M.F., Glen, J.M., Earney, T.E., Schermerhorn, W.D., Peacock, J., Deoreo, S.B., Siler, D.L., 2017. Discovery of a Blind Geothermal System in Southern Gabbs Valley, Western Nevada, through Application of the Play Fairway Analysis at Multiple Scales, Geothermal Resource Council Transactions. Salt Lake City, UT, USA.

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, in: Spencer, J.E., Titley, S.R. (Eds.), Ores and Orogenesis: Circum-Pacific Tectonics, Geologic Evolution, and Ore Deposits: Arizona Geological Society Digest. pp. 437-470.

Faulds, J.E., Hinz, N., Coolbaugh, M., dePolo, C., Siler, D., Shevenell, L., Hammond, W., Kreemer, C., Queen, J.H., 2016a. Discovering Geothermal Systems in the Great Basin Region: An Integrated Geologic, Geochemical, and Geophysical Approach for Establishing Geothermal Play Fairways, Workshop on Geothermal Reservoir Engineering, 41. Stanford, California. 22 – 24 February, 2016.

Faulds, J.E., Hinz, N.H., 2015. Favorable Tectonic and Structural Settings of Geothermal Systems in the Great Basin Region, Western USA: Proxies for Discovering Blind Geothermal Systems, World Geothermal Congress 2015. Melbourne, Australia. 19 – 25 April 2015.

Faulds, J.E., Hinz, N.H., Coolbaugh, M.F., Craig, J.W., Glen, J.M., Ayling, B.F., Sadowski, A.J., Siler, D.L., DeOreo, S., 2021b. The Nevada Geothermal Play Fairway Project: Exploring for Blind Geothermal Systems through Integrated Geological, Geochemical, and Geophysical Analyses, World Geothermal Congress. Reykjavik, Iceland.

Faulds, J.E., Hinz, N.H., dePolo, C., Hammond, W., Kreemer, C., Shevenell, L., Coolbaugh, M.F., Queen, J.H., Siler, D.L., Visser, C., Wannamaker, P.E., 2016b. Nevada Great Basin Play Fairway Analysis - Reports & Appendices. US Department of Energy Geothermal Technologies Office, https://gdr.openei.org/submissions/756.

Geurts, P., Ernest, D., Wehenkel, L., 2006. Extremely Randomized Trees. Machine Learning 63, pp. 3-42.

Glen, J.M., Earney, T.E., Zielinski, L.A., Schermerhorn, W.D., Dean, B.J., Hardwick, C., 2022. Regional Geophysical Maps of the Great Basin, USA. U.S. Geological Survey Data Release, https://doi.org/10.5066/P9Z6SA1Z.

Hamming, R.W., 1950. Error Detecting and Error Correcting Codes. Bell System Technical Journal 29, 147-160.

Harry, D.L., Sawyer, D.S., Leeman, W.P., 1993. The Mechanics of Continental Extension in Western North America: Implications for the Magmatic and Structural Evolution of the Great Basin. Earth and Planetary Science Letters 117(1), 59-71.

Hart-Wagoner, N.R., Coolbaugh, M., Faulds, J.E., Mlawsky, E., 2023. Feature Engineering of Fault Attributes for Play Fairway Analysis, Great Basin Region, USA, Geothermal Rising Conference. Geothermal Resources Council Transactions, Reno, NV, USA.

Hart-Wagoner, N.R., Coolbaugh, M., Faulds, J.E., Mlawsky, E., Lindsey, C., Trainor-Guitton, W., 2024a. Preliminary Regional Play Fairway Workflow for the Great Basin Region, USA, Stanford Geothermal Workshop. Stanford Geothermal Workshop, Stanford, CA, USA.

Hart-Wagoner, N.R., Lindsey, C., Trainor-Guitton, W., Coolbaugh, M., Faulds, J.E., 2024b. Exploring Geothermal Potential of Great Basin Sub-Regions, Geothermal Rising Conference. Geothermal Resource Council Transactions, Waikoloa, HI, USA, pp. 1291-1309.

Hughes, G., 1968. On the Mean Accuracy of Statistical Pattern Recognizers. IEEE Transactions on Information Theory 14(1), 55-63.

Kennedy, B.M., Farquhar, A., Hilderman, R., Villeneuve, M.C., Heap, M.J., Mordensky, S.P., Kilgour, G., Jolly, A., Christenson, B., Reuschlé, T., 2020. Pressure Controlled Permeability in a Conduit Filled with Fractured Hydrothermal Breccia Reconstructed from Ballistics from Whakaari (White Island), New Zealand. Geosciences (Switzerland) 10, 1-19.

Kreemer, C., Young, Z.M., 2023. Crustal Strain Rates in the Western United States and Their Relationship with Earthquake Rates. Seismological Research Letters 93(6), 2990-3008.

Lundstern, J.-E., Zoback, M.D., 2022. State of Stress in Areas of Active Unconventional Oil and Gas Development in North America. AAPG Bulletin 106(2), 355–385.

Lundstern, J., Zoback, M.D., 2023. Maximum Horizontal Stress Orientation and Relative Stress Magnitude (Faulting Regime) Data Throughout North America. U.S. Geological Survey Data Release, https://www.sciencebase.gov/catalog/item/6469516bd34e3a6027e2f527.

Mlawsky, E., Ayling, B.F., 2021. The Gbcge Subsurface Database Explorer. Nevada Bureau of Mines and Geology, Great Basin Center for Geothermal Energy, University of Nevada Reno, https://gdr.openei.org/submissions/1486.

Mordensky, S.P., Burns, E.R., DeAngelo, J., Lipor, J.J., 2023a. Predicting Large Hydrothermal Systems, Geothermal Resources Council Transactions. Geothermal Rising, Reno, NV, USA, pp. 1763-1796.

Mordensky, S.P., Burns, E.R., Lipor, J.J., DeAngelo, J., 2025a. Favorability Mapping for Hydrothermal Power Resource Assessments the Great Basin, USA. Geothermics 133 (103450). https://doi.org/10.1016/j.geothermics.2025.103450.

Mordensky, S.P., DeAngelo, J., 2026. Feature, Label, Prediction, and Zone Data for Great Basin Zone Classification and Hydrothermal Favorability Modeling Using Supervised Machine Learning: U.S. Geological Survey Data Release, https://doi.org/10.5066/P14QM39R.

Mordensky, S.P., Lipor, J.J., Burns, E.R., DeAngelo, J., 2025b. Preventing Overfitting When Using Tree-Based Methods for Mapping Hydrothermal Favorability. Geothermal Resources Council Transactions, 49, 179-203, Reno, Nevada, 26-29 October 2025.

Mordensky, S.P., Lipor, J.J., DeAngelo, J., Burns, E.R., Lindsey, C.R., 2023b. When Less Is More: How Increasing the Complexity of Machine Learning Strategies for Geothermal Energy Assessments May Not Lead toward Better Estimates. Geothermics 110 (102662). https://doi.org/10.1016/j.geothermics.2023.102662.

Muffler, L.P.J., 1979. Assessment of Geothermal Resources of the United States-1978, U.S. Geological Survey Circular 790. pp. 163.

Peacock, J.R., Bedrosian, P.A., 2022. Electrical Conductance Maps of the Great Basin, USA. U.S. Geological Survey Data Release, https://doi.org/10.5066/P9TWT2LU.

Reed, M.J., 1983. Assessment of Low-Temperature Geothermal Resources of the United States- 1982, Geological Survey Circular 892. pp. 73.

Sainburg, T., McInnes, L., Gentner, T.Q., 2021. Parametric Umap Embeddings for Representation and Semisupervised Learning. Neural Computation 33(11), 2881-2907.

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., 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.

Trainor-Guitton, W., Menon, K., Pinchuk, P., Thomson, S.M., Hart-Wagoner, N., Lu, C., Mlawksy, E., Coolbaugh, C., Lindsey, C., Faulds, J., 2025. "Hidden" hydrothermal technical potential & technoeconomics: Revealing permeability & fluids with more data. Geothermics 133 (10373). https://doi.org/10.1016/j.geothermics.2025.103473.

U.S. Congress, 2005. Energy Policy Act of 2005. Public Law 109-58, 109th Congress, Available at: https://www.congress.gov/bill/109th-congress/house-bill/106 [Accessed 8 August 2025].

U.S. Geological Survey, 2024. 3d Elevation Program Hillshade Model. U.S. Geological Survey, https://www.usgs.gov/3d-elevation-program.

White, D.E., Williams, D.L., 1975. Assessment of Geothermal Resources of the United States, U.S. Geological Survey Circular 726, pp. 155.

Williams, C.F., DeAngelo, J., 2008. Mapping Geothermal Potential in the Western United States, Geothermal Resources Council Transactions. Reno, NV, USA, pp. 181-188.

Williams, C.F., Reed, M.J., DeAngelo, J., Galanis, S.P., 2009. Quantifying the Undiscovered Geothermal Resources of the United States, Geothermal Resource Council Transactions. Reno, NV, USA, pp. 882-889.

Williams, C.F., Reed, M.J., Mariner, R.H., 2008a. A Review of Methods Applied by the U.S. Geological Survey in the Assessment of Identified Geothermal Resources. Open-File Report 2008–1296, 30 p.

Williams, C.F., Reed, M.J., Mariner, R.H., DeAngelo, J., Galanis, S.P., 2008b. Assessment of Moderate-and High-Temperature Geothermal Resources of the United States, U.S. Geological Survey Fact Sheet 2008-3082. pp. 1-4.

Zeng, Y., 2022. Gps Velocity Field of the Western United States for the 2023 National Seismic Hazard Model Update. Seismological Research Letters 93(6), pp. 3121-3134.