Value of Information (VOI) App V 1.0 User Manual:

https://voigeothermalrising.streamlit.app

http://voi.geothermal.nrel.gov/

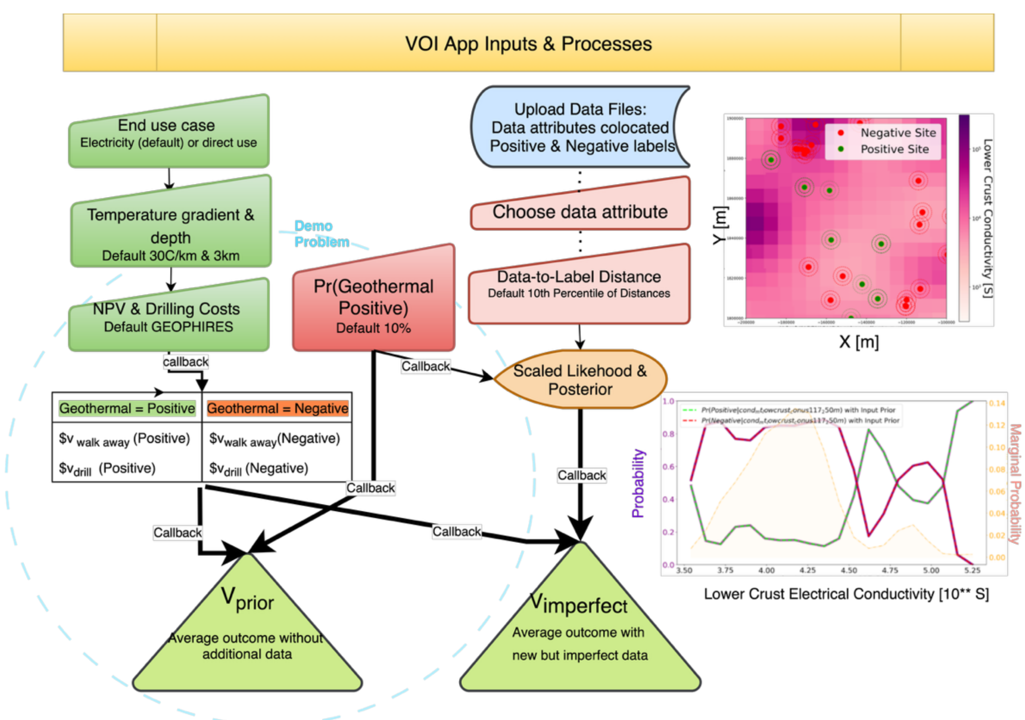


Figure 1: VOI App workflow, identifying: both V_{prior} and $V_{imperfect}$ use the same 4 value outcomes from the binary decision-binary geothermal uncertainty and rely on prior probability; the VOI demo problem doesn't consider imperfect data; and user's data should have data attributes with a distance associated to both positive and negative geothermal labels. If $V_{imperfect} > V_{prior}$, then it is a sound decision to purchase that data attribute before making the decision.

Introduction:

The Value of Information (VOI) app is an open-source tool designed for computing the value of geoscience datasets in geothermal project development. VOI, a decision theoretical concept that originates from decision analysis, is based on the idea that new information can influence a decision when several alternatives exist (Clemen and Reilly, 2014). If this information leads to a change in the initial decision, the resulting improvement represents it's 'value'. Conventionally, VOI can be of two types- Value of Perfect Information (assuming the information is perfect i.e. provides a complete picture of what we are trying to model) and Value of Imperfect Information (assuming the information is incomplete, ambiguous or non-unique). For the geothermal VOI App, imperfect specifically means certain values of the data (e.g. electrical conductivity of 40 Ohm-m), is associated equally with both the existence (positive) and absence (negative) of a geothermal system. Most, if not all, geoscience data is imperfect, and thus the goal of this app is to provide a straightforward environment for computing the value of such imperfect information.

The VOI App currently only models a binary decision (walk away or drill) where the geothermal resource is also binary: positive (exists) or negative (doesn't exist). Therefore, there are four possible outcomes of this binary-binary decision and outcome are depicted in **Table 1**.

Table 1: Four value outcomes from two geothermal possibilities (positive/negative) and two decision actions (drill/walk away)

Actions	**Positive geothermal scenario**	**Negative geothermal scenario**
Walk away	$v_{a=walk\ away}(\Theta = positive) = \$0$	$v_{a=walk\ away}(\Theta = negative) = \$0$
Drill	$v_{drill}(\Theta = positive) = \$\,revenue - drilling\ costs$	$v_{drill}(\Theta = negative) = -\$drilling\ costs$

Getting Started:

To access the app, navigate your browser to https://voigeothermalrising.streamlit.app or http://voi.geothermal.nrel.gov/. The page should appear as shown below (without colored boxes and annotations):

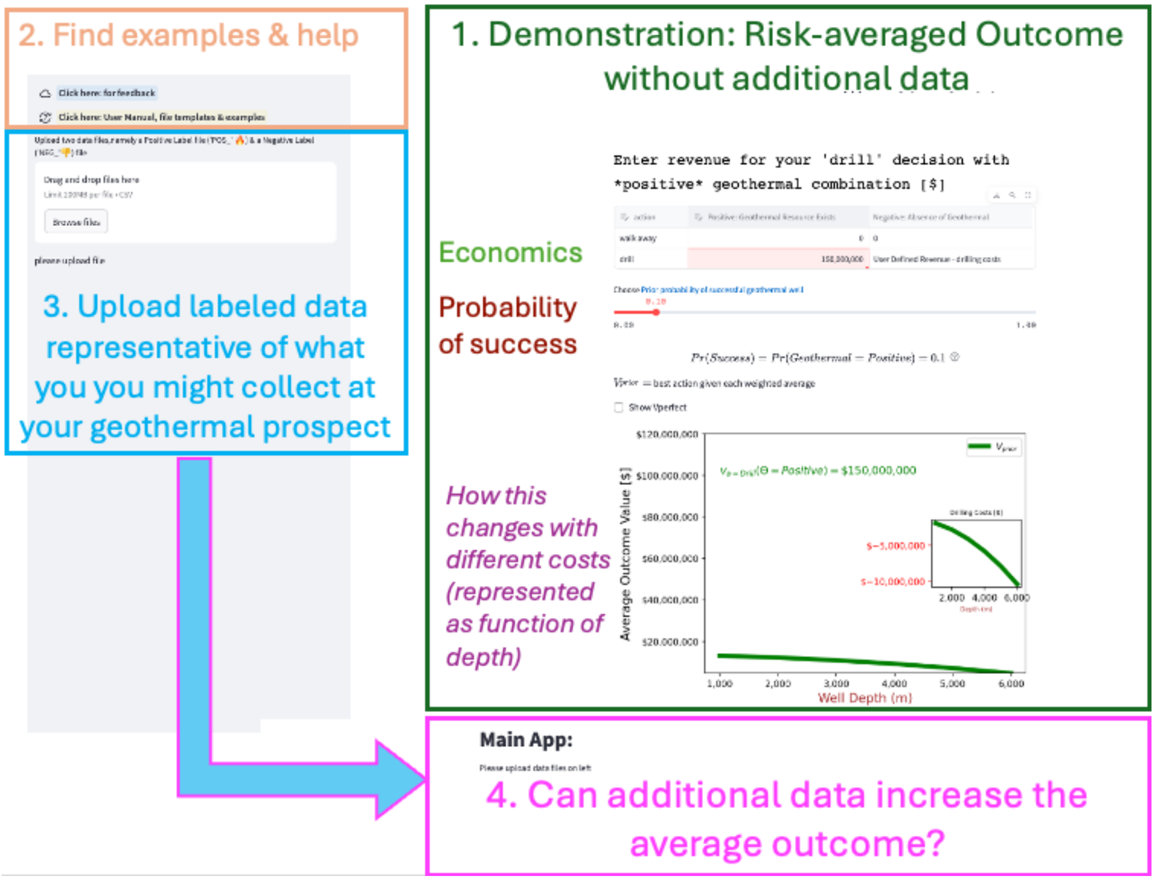


Figure 2: VOI App Landing Page. The upper left section provides links to this manual and example files.

There are four main sections of the app as shown by the colored boxes in Figure 2.

1. Demo problem (green box): This section helps users become familiar with the concept of average decision outcome (V_{prior}) and Value of Perfect information (which assumes unique relationships exist between a data attribute and the presence or absence of a geothermal resource).
2. Helpful links (orange box, upper left): Here, users can access:
   - This VOI App Manual
   - Example data files from the INGENIOUS project:
     - Data Attributes at **Positive** Labels (geothermal resource exists)
       - POS Entire INGENIOUS area: POS_COMBINED_NEW_JULYRESID_INGENIOUS_updated.csv
       - POS Only Walker Lane Region: POS_Walker_INGENIOUS_onlypublished.csv
     - Data Attributes at Negative Labels (absence of geothermal resource)
       - NEG Entire INGENIOUS area: NEG_COMBINED_NEW_JULYRESID_INGENIOUS_update.csv

- [NEG Only Walker Lane Region](#): NEG_Walker_INGENIOUS_onlypublished.csv

3. Upload Section (blue box): This is where users can upload their labeled data files.
   - As denoted below, ***no data are*** retained by the App.
4. Evaluation of uploaded data starts here (pink box): This is where imperfect information starts and will only be visible once data is uploaded.

The rest of the manual will describe these sections in detail.

VOI Demo Problem:

The VOI app includes an example problem at the very beginning, highlighting the various terminologies used in the app, and introducing the user to basic terms in Bayesian Statistics, a field of statistics that uses probability to update prior knowledge with new evidence and assist in decision-making where the outcome is uncertain. It also covers a brief overview of how the main VOI app calculations work.

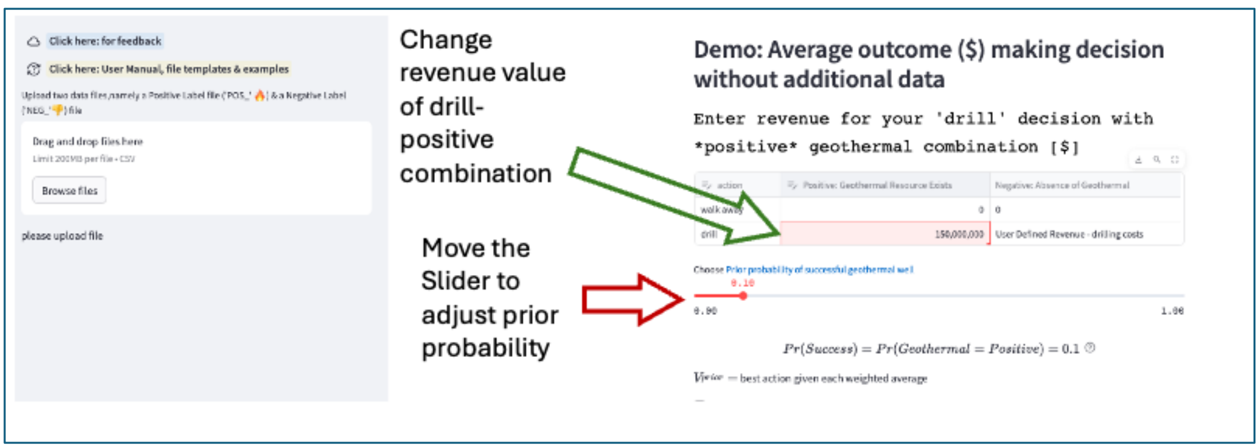


Figure 3: Inputs for demo problem: the green arrow indicates economic outcome when geothermal resource exists and an action taken (e.g. drill, v_{drill}(Success)). The red arrow indicates prior probability of success (geothermal resource exists).

In Figure 3 above, the two user-adjustable 'toggles' are indicated by the green and red arrows; only these two values can be modified in the demo problem. The green arrow indicates where the user can adjust the revenue expected if a geothermal resource exists AND the decision was made to drill. The action to drill (last row) comes with costs: drilling costs as a function of depth. They are displayed along the x-axis (depth) in Figure 4. A range of negative values, representing unsuccessful outcome, are assigned to match reasonable drilling cost per well at each depth. If no geothermal resource exists, only negative economic values are used (the lower right cell of table in Figure 4).

The other decision alternative is to "walk away" (row below the header): not to drill or do nothing. Therefore, the value outcomes of this row are both $0: they represent no further investment or action.

The red arrow indicates where the prior probability can be toggled (Figure 3). This the ***current*** decision maker's estimate of a geothermal resource existing (positive), e.g. using current understanding of the prospect. For this binary problem, it can be referred to as the prior probability of success *Pr(Positive*, and it ranges from 0 to 1. Changing this allows the user to observe changes in both the prior expected value (V_{prior}) and the Value of Perfect Information ($VOI_{perfect}$).

As shown in the App, the V_{prior} is the action that has the highest (weighted) average, using the prior probability as the weights:

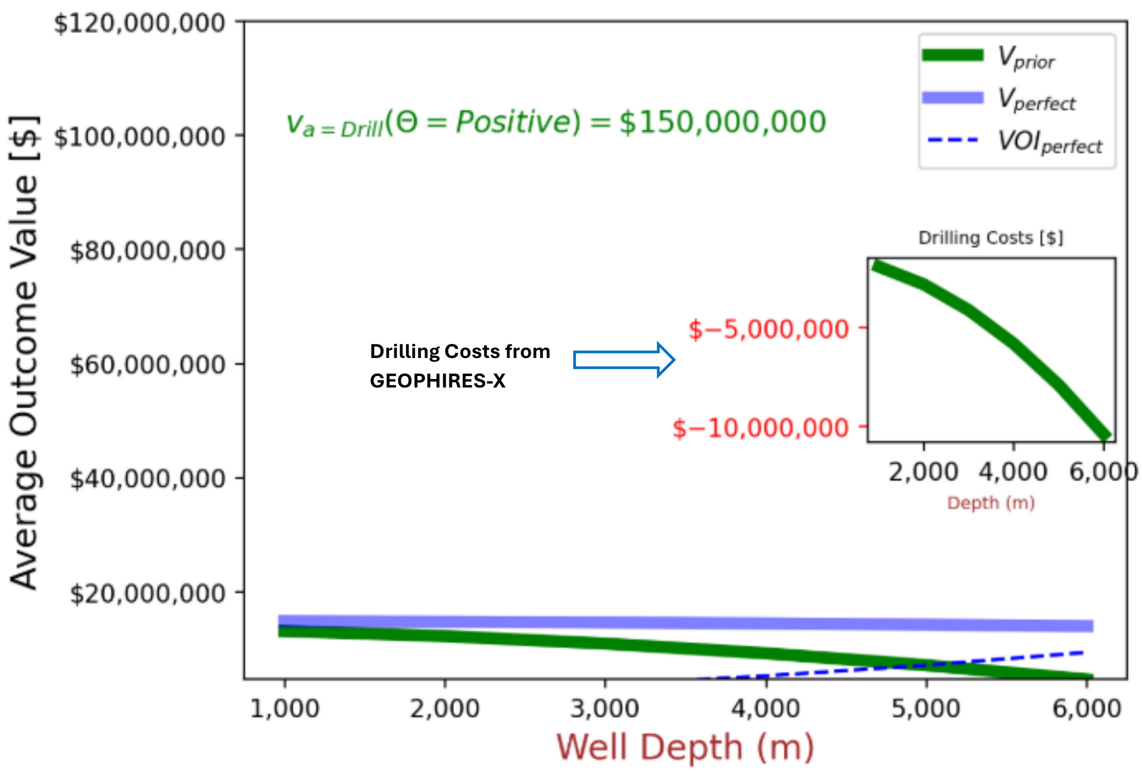
$V_{prior} = \max(V_{a=drill}, V_{a=walk\ away})$, where,

- $V_{a=drill} = Pr(Positive) * v_{drill}(Positive) + (1-Pr(Positive)) * v_{drill}(Negative)$
- $V_{a=walk\ away} = Pr(Positive) * v_{walk\ away}(Positive) + (1-Pr(Positive)) * v_{walk\ away}(Negative)$

In the demo example, when V_{prior} = $0, it says the best (highest) action is to walk away (don't drill). The risk (probability * outcome) of drilling is too high.

V_{prior} = best action given each weighted average

Show Vperfect



When you "know" when either subsurface condition occurs, you can pick the best ($\max_a$) drilling altervative first (v_a).

$$V_{perfect} = \Sigma_{i=1}^{2} Pr(\Theta = \theta_i) \max_a v_a(\theta_i) \;\; \forall a$$

$$VOI_{perfect}(Value\ of\ Information) = V_{perfect} - V_{prior} = 14810000.0 - 13100000.0$$

Figure 4: Visualization of Demo Problem in the app as a function of drilling costs (x-axis). The plot shows prior value (V_{prior} green), value with perfect information ($V_{perfect}$, solid blue) and value of perfect information ($VOI_{perfect}$, dashed blue).

The $V_{perfect}$ (the value with perfect information) represents an upper bound on the maximum value that the information can provide, assuming a hypothetical information source could perfectly reveal whether a geothermal resource exists or not. Its value depends on the prior probability chosen. In our demo problem, we assume the positive economic outcome is constant across all depths, while drilling costs increase with depth. Compared to V_{prior}, $V_{perfect}$ reverses the order of the maximum and weighted average operations. This mimics how the perfect information source allows for selecting the optimal (max) action for each scenario (positive and negative) before averaging outcomes.

$V_{perfect} = Pr(Positive) * V(Positive) + \big(1 - Pr(Positive)\big) * V(Negative)$, where,

- V(Positive) = max(v_{drill}(Positive), $v_{walk\ away}$(Positive))
- V(Negative) = max(v_{drill}(Negative), $v_{walk\ away}$(Negative))

The value *of* perfect information ($VOI_{perfect}$) compares how much of $V_{imperfect}$ is higher than V_{prior}(i.e. $VOI_{perfect}$ = $V_{perfect}$ - V_{prior}). As shown in Figure 4, $VOI_{perfect}$ has the most value when the drilling costs are highest and V_{prior} is lowest. Intuitively, this should make sense: new information will have more value when the economic consequences are large (e.g. high costs), and our perceived chances of a successful outcome are low.

Dataset Format & Uploading:

The next part of the VOI App allows users to upload their own labeled datasets. The App ***does not store*** the data, so users can upload proprietary data without jeopardizing any business sensitivities. First let's conceptualize what this labeled data could be.

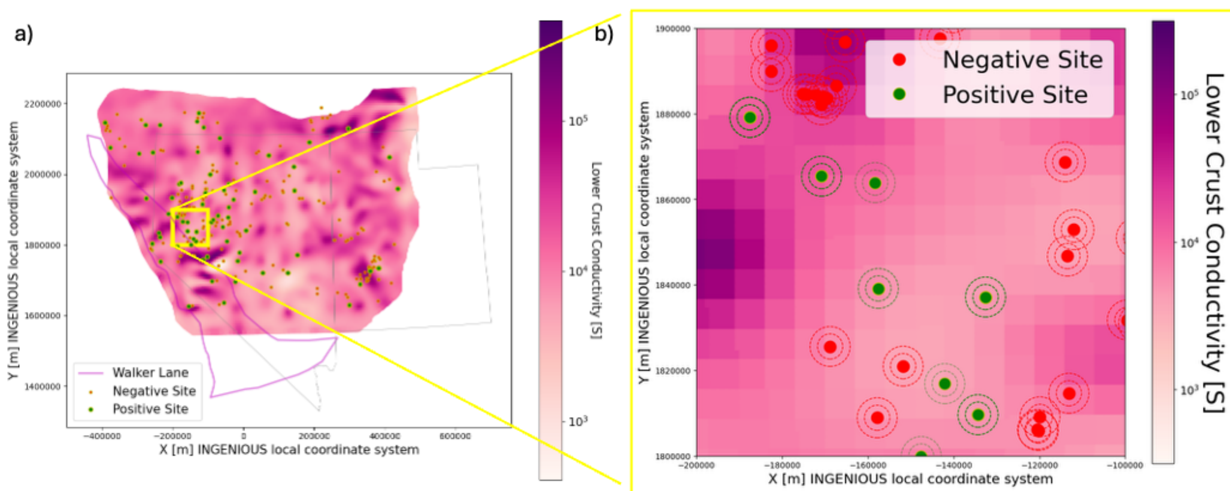


Figure 5: a) INGENIOUS area with Lower Crust Electrical Conductivity and positive (green) and negative (red) geothermal sites/labels. Walker Lane outline denotes area of subset of labels and data available in example files. Yellow box is locations of zoomed in area. b) Zoomed in area, showing two different possible radii around the labels. This radius distance around labels is controlled in the VOI App and determines which attribute values (e.g. conductivity) are assigned to the positive and negative labels.

First, we provide of visual of the example labeled datasets that are linked within the App. Figure 5a maps a version of geothermal labels within the INGENIOUS study area along with the lower crust electrical conductivity. (Ayling et al. 2022; Trainor-Guitton et al. 2025). Figure 5b is a zoomed in view of some of these labels with two possible depictions of radii distances around them. The labeled dataset is identifying the attribute values (e.g. lower crust electrical conductivity) that are collocated with a label. The user of the app can control the distance that defines this "collocation." As this distance increases, more attributes are pulled farther from the center location of the label.

Therefore, when users create and use their own data sets, they must identify positive and negative labels and then calculate their respective distances to any data attribute of interest. Figure 6 provides a visual of the bare minimum file format needed using the example of Figure 5: two files (with prefixes of "POS_" and "NEG_"), and with two columns containing the data attribute and it's distance to each respective label.

a) POS_*filename.csv*

	A	B
1	cond_mt_lowcrust_conus117_250m	PosSite_Distance
2	4.127771	707.1068
3	4.127771	559.01697
4	4.127771	500
5	4.127771	559.01697
6	4.055595	707.1068
7	4.127771	559.01697
8	4.127771	353.5534
9	4.127771	250
10	4.127771	353.5534
11	4.055595	559.01697
12	4.127771	500

b) NEG_*filename.csv*

	A	B
1	cond_mt_lowcrust_conus117_250m	NegSite_Distance
2	4.1377172	707.1068
3	4.1377172	587.67523
4	4.1377172	500
5	4.1377172	587.67523
6	4.1377172	707.1068
7	4.1377172	587.67523
8	4.1377172	353.5534
9	4.1377172	250
10	4.1377172	353.5534
11	4.1377172	587.67523
12	4.1377172	500

Figure 6: Example file format for example in Figure 5 with filename prefix highlighted. Data attribute and distance to positive labels (a) and data attribute and distance to negative labels (b).

To ensure that the user data can be read in, please follow the three steps:

1. Two files must be created: one containing data that is associated with the 'positive' label(s) and has the prefix 'POS' (POS_*filename*), and another that is from a 'negative' label with prefix 'NEG' (NEG_*filename*). These labels can be defined by the user and need not coincide with actual geothermal sites. For example, they could be generated synthetic data from forward modeling.
2. The positive and negative files must contain at a minimum two columns, representing feature of interest and distance. The columns for both files must have appropriate headers : 'PosSite_Distance' and 'NegSite_Distance, respectively. See Figure 6.
3. When these files are ready, a csv version must be uploaded onto the app (Figure 7 below for reference).

Figure 7 demonstrates how files are uploaded on the left side of the App. The example files are from the Walker Lane section of the INGENIOUS project[1] (located in the GitHub repository[2]). They display the correct naming terminology as mentioned above. The files

[1] https://gdr.openei.org/submissions/1391
[2] https://github.com/NREL/Value_of_Information_App/tree/main/File Template

must contain a distance column for the app to work (see the file template for more details), and the slider in the figure can be adjusted to include/exclude more points. The default value for the slider is to use the 25th percentile (1st quartile) taken from the negative label file.

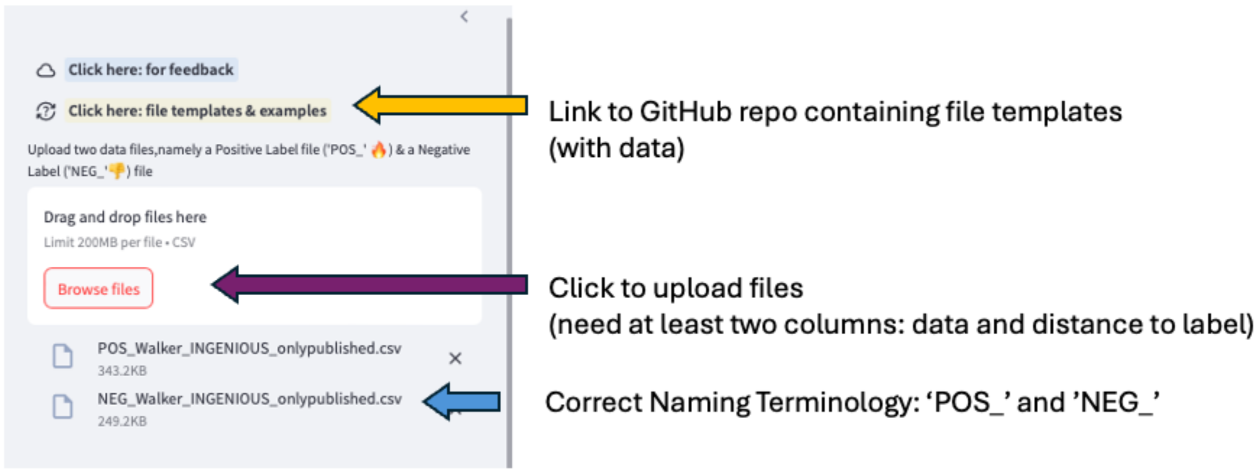


Figure 7: Upload function of where to labeled data files are loaded in VOI App

Once the files are uploaded, users can select built-in project economics options based on the open-source tool GEOPHIRES-X: *electricity* or *direct use applications* (top yellow arrow of Figure 8)(Beckers and McCabe, 2019). The default values for electricity costs are taken from the recent Fervo Power Purchase Agreements (Fervo PPA) and general default values are available on the GEOPHIRES-X GitHub page (GEOPHIRES-X defaults).

If the uploaded files contain multiple data features (e.g. the INGENIOUS example files on the VOI App Github), users can select among them using the field indicated by the bottom purple arrow of Figure 8 . The resulting plot is the likelihood of the raw data values: the green bars from the positive file and the red bars from the negative file. The user can also screen data values far from the label. The slider shown with the blue arrow chooses the radius distance (Figure 5b): a larger screening distance includes more points in the likelihood (e.g., data features farther away from labels of a positive or negative geothermal resource).

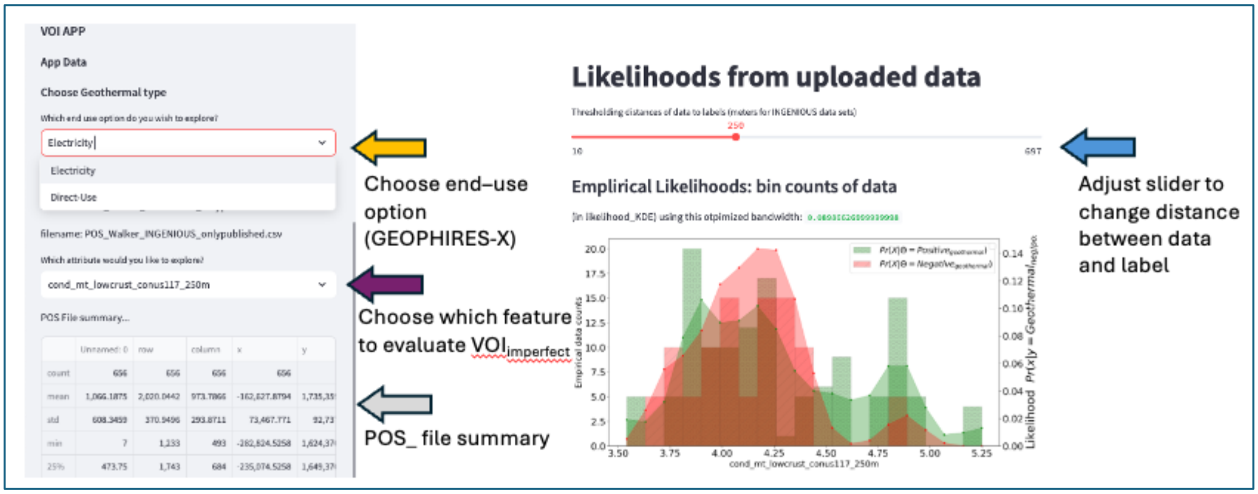


Figure 8: VOI App parameter selections: 1) Top-yellow arrow indicates end use selection (electricity or direct use), 2) Middle-purple arrow indicates data feature selection, and 3) right-blue arrow indicates distance threshold between data and the positive or negative label.

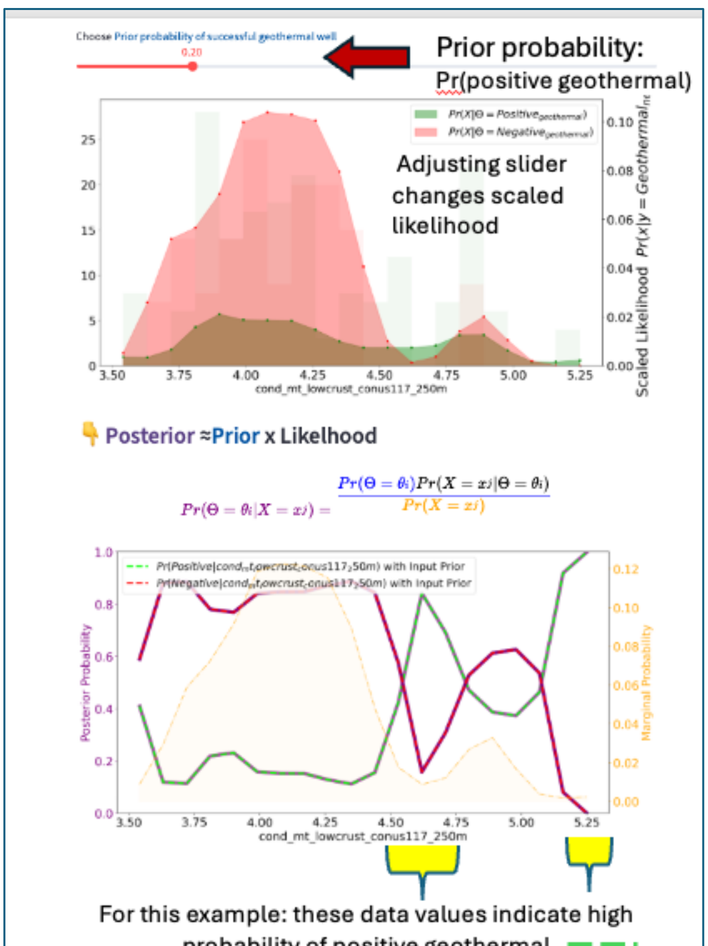


Figure 9: Scaled-likelihood and posterior plots for a given feature (example is from Walker Lane, lower crust conductivity).

Next, the user chooses the prior probability of success, which will be used to scale the likelihood of Figure 8. The top of Figure 9 illustrates a likelihood scaled by ***Pr(Geothermal = Positive) = 20%.*** Therefore, the green areas represent 20% of the total area covered by red and green in the histograms. The likelihood can be expressed as $\Pr\left(X = x_j \middle| \Theta = \theta_i\right)$, where the labels ($\theta_i$) are known (Trainor-Guitton and Rosado, 2023).

The bottom panel of Figure 9 displays the posterior, which looks at each data bin value (the x-axis) of the scaled likelihood. The posterior chronologically reverses the likelihood: *if I observe data bin x_j, there is a y_j chances of being a positive geothermal site (and $1-y_i$ being negative), where* $y_i = \Pr\left(\Theta = \theta_i \middle| X = x_j\right)$.

For the example shown, if the lower crust conductivity is 10**3.75, there's roughly a 10% probability a geothermal resource exists. However, if the lower crust conductivity is > 10**5.2, there's a very high probability for a positive geothermal resource.

The marginal probability is plotted in orange and indicates the data density of the uploaded data file for each given histogram bin. For lower crust conductivity, most of the values are between 10**3.75 to 10**4.5. This is also considered in $VOI_{imperfect}$.Figure 10

VOI Metrics:

Once an end use option has been selected, the user can input the temperature gradient and depth to the reservoir as shown in Figure 10 (blue and yellow arrows, respectively). GEOPHIRES-X is then run, and the corresponding project NPV is populated in the '*Geothermal Resource (positive)'* column, and the drilling cost per well is populated in the '*No Geothermal Resource (negative)'* column. The first column refers to the outcome if the well is successful, and the second is if the well is unsuccessful. The NPV changes based on the end-use option selected. The drilling cost remains the same and only changes if the depth parameter is altered.

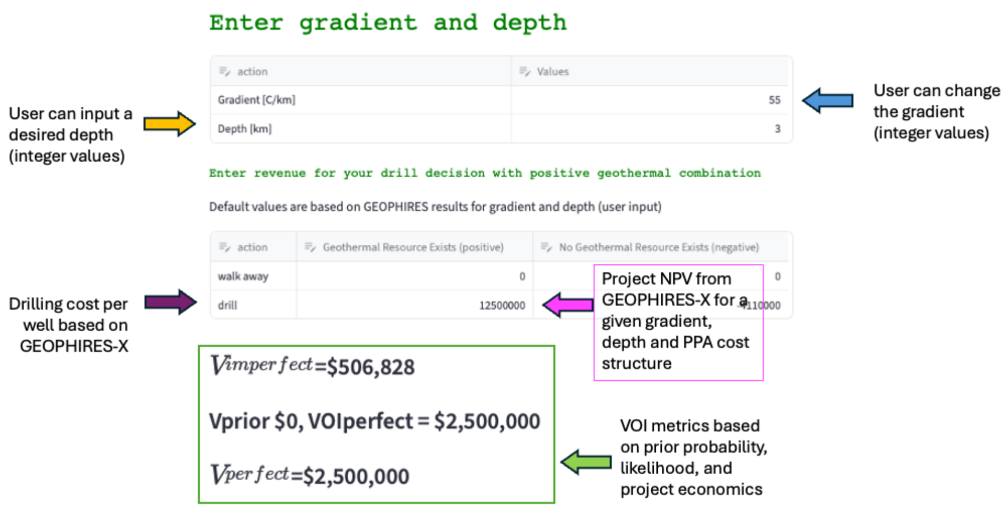


Figure 10: VOI app economics input parameters and output metrics.

The VOI App currently uses the same default parameters as would be used in a generic run of GEOPHIRES, with the exception of the following parameters:

"Starting Electricity Sale Price": "0.15",

"Ending Electricity Sale Price": "1.00",

The VOI App defaults to "Power Plant Type": "1", a subcritical ORC, with two injector ("Number of Production Wells") and two producer wells ("Number of Production Wells").

The parameters can be viewed in the file GEOPHIRES_X.py[3]. In theory, this is where parameters could be overwritten. However, this would have to be done on a local version of the code repository, then a local instance of the Streamlit app would need to be run. The fastest way to implement different parameters (e.g. porosity, fractures characterization, etc.) is to run GEOPHIRES separately, retrieve the output metrics (i.e. costs and NPV), and enter those into the value matrix (purple and pink arrows in Figure 10). The user is always able to input their own choices for project economics there.

The final metric is the $V_{imperfect}$ and comparing that to V_{prior}: can new data bring about higher expected economic outcomes? The posterior is the backbone of $V_{imperfect}$: determining how often correct interpretations of the geothermal site are made from each of the data bins. The posterior can be thought of the reliability of a data type to identify a positive or negative geothermal resource, and importantly, it is used in the value with imperfect information ($V_{imperfect}$):

$$V_{imperfect}(\boldsymbol{u}) = \sum_{j=1}^{N} \Pr(X = x_j) \left[\max_{a} \left[\sum_{i=1}^{2} \Pr(\Theta = \theta_i | X = x_j)\, v_a(\Theta = \theta_i) \right] \right]$$

$$a = drill, don't\ drill$$

Within $V_{imperfect}$, is the marginal probability: $\Pr(X = x_j)$. This scales the result to how frequently any of the data magnitudes are observed in the field, as represented by the uploaded data (orange plot in Figure 9).

If $\mathbf{V_{imperfect}} > \mathbf{V_{prior}}$, then it is a sound decision to purchase that data attribute before making the decision.

2025 Enhancements

Since the VOI App's first public release, the default economics are now linked to GEOPHIRES-X.

Providing App Feedback

For any feedback related to the app, users can click on the feedback form button on the top-left hand side of the app which will re-direct them to a Microsoft form. See the upper left corner of Figure 2.

[3] https://github.com/NREL/Value_of_Information_App/blob/main/GEOPHIRES_X.py

Ayling, Bridget, James E Faulds, A Rivera, et al. 2022. "INGENIOUS - Great Basin Regional Dataset Comilation. Study Area Boundary." Geothermal Data Repository. https://doi.org/10.15121/1881483.

Beckers, Koenraad F., and Kevin McCabe. 2019. "GEOPHIRES v2.0: Updated Geothermal Techno-Economic Simulation Tool." *Geothermal Energy* 7 (1). https://doi.org/10.1186/s40517-019-0119-6.

Clemen, Robert T., and Terence Reilly. 2014. *Making Hard Decisions with DecisionTools*. 3rd ed. South-Western, Cengage Learning.

Trainor-Guitton, Whitney, Karthik Menon, Pavlo Pinchuk, et al. 2025. "'Hidden' Hydrothermal Technical Potential & Technoeconomics: Revealing Permeability & Fluids with More Data." *Geothermics* 133 (December): 103473. https://doi.org/10.1016/j.geothermics.2025.103473.

Trainor-Guitton, Whitney, and Sierra Rosado. 2023. "A VOI Web Application for Distinct Geothermal Domains: Statistical Evaluation of Different Data Types with the Great Basin." *Geothermal Resources Council Transactions* 47: 1836–51. https://www.geothermal-library.org/index.php?mode=pubs&action=view&record=1034920.