PROCEEDINGS, 42nd Workshop on Geothermal Reservoir Engineering
Stanford University, Stanford, California, February 13-15, 2017
SGP-TR-212

Iterative Complexity Addition: A Method to Estimate Uncertainty in Geophysical Inversions

D. Keith Smithson, Egill Júlíusson, Hlynur Stefánsson, Samuel Perkin

Menntavegur 1, 101 Reykjavik, Iceland

dksmithson@gmail.com

Keywords: Geophysics, inversion, uncertainty, optimization, geothermal, complexity, magnetotelluric

ABSTRACT

Iterative Complexity Addition (ICA) is a method to estimate the uncertainty obtained from the inversion of geophysical data. The method utilizes a grid search with gradient descent for each set of model parameters to invert 1D magnetotelluric (MT) data. The algorithm finds a set of minimally complex local optima that meet the error criteria of the measurement, implying a possible solution to the inversion. The set of solutions generated can be analyzed for variance in the model parameters to generate statistical estimates of their values. These statistics reflect the likelihood of existence of each set of model parameters that can be used to enhance decision-making for subsurface resource development. The method design is here described with a discussion on algorithm strengths, weaknesses, and potential improvements.

1. INTRODUCTION

Inverse theory is a set of mathematical techniques that are used to distill estimates of model parameters from measurement data. (Menke, 2012). In many cases, such as through the surface exploration for geothermal reservoir, the data collected cannot be uniquely correlated to a single set of subsurface conditions.

Inversion methods fit broadly into two categories: deterministic and stochastic (Yang, 1999). The varieties of deterministic inversion methods vary in how a single subsurface structure model is selected from a multitude of possibilities. Stochastic methods of inversion treat both the data and model parameters as probability distributions. Inversion methods typically employ optimization techniques to generate a good fit between a solution's response and the measured data. Inversions tend to be computationally expensive and much work has been done to make calculations more efficient (Trainor-Guitton and Hoversten, 2011).

Different inversion methods may be applicable to certain types of problems, so they be difficult to compare. Value of Information (VOI) analysis provides a way of quantifying the quality of data. Data reliability is defined as the probability that a measurement accurately describes the property being measured. This provides a way to compare different measurement methods as well as different inversion methods (Bratvold, 2009; Trainor-Guitton et al, 2011).

Ramirez et al (2005) discusses stochastic inversion of electrical resistivity using a Markov Chain Monte Carlo (MCMC) method. The paper identifies the advantage of stochastic inversions where multiple solutions are generated and can be used to estimate the posterior probability distribution via the Metropolis algorithm. Ramirez also suggests a process using dynamic k-means clustering as a method of comparing solutions to each other and identifying the statistical likelihood of each group or "mode" of models. This method of grouping solutions is a promising way of comparing solutions to determine uncertainty, but it doesn't account for the additional uncertainty associated with non-clustered models that may share some of the same parameters. Ramirez et al show an example that inverts the properties of 28,800 voxels from 128 transmitting electrodes that would require 45 days to compute on a work station with one CPU. Running the calculation in parallel significantly reduces the time required. Additionally, it is noted that the convergence of the algorithm can be significantly impacted by the prior distribution that is chosen.

Trainor-Guitton and Hoversten (2011) discuss many of the practical challenges of stochastic inversion as applied to controlled source electromagnetic (CSEM) and MT data. One of the biggest challenges they reference is in the ability to appropriately diagnose the convergence of the algorithm. Convergence for a MCMC stochastic algorithm occurs when the Markov chain covers the entire posterior distribution (all possible solutions) and the Markov chain is independent of the starting candidate model. Depending on the chosen criteria for divergence and the prior distribution, there is risk of excessive computational load or finding imprecise parameter variances.

An obvious pitfall to the inversion process is the requirement to make a decision from a single solution when it is difficult to determine if one is more correct than another (even before considering measurement noise and error). Determining the uncertainty of a solution to a complex inverse problem is not well understood (Mosegaard and Tarantola, 1995; Chen et al, 2012).

Iterative Complexity Addition is an inversion method that tries to combine the benefits of both deterministic and stochastic methods to provide information about the uncertainty associated with each of the model parameters in the solution set.

For initial testing of the ICA method, MT data for the Þeistreykir Geothermal Field in North Iceland was used. A 3D inversion of the measurements has been completed by Karlsdóttir et al (2012). The 3D inversion was completed using a deterministic inversion (Occam

inversion) that imposes a constraint on model parameters to vary smoothly. This process is very sensitive to the initial model that is chosen to start the calculation. To alleviate this risk, the inversion was completed using three separate initial models. Total computation time on a 32-core computer with 132 GB RAM was greater than 1000 hours. Comparison is made between the three inversions from different starting points as well as with a smoothed version of compiled 1D inversions. Some features are common between all models while other features are unique to one model or another.

2. METHOD

ICA tries to match the computational simplicity of a deterministic inversion with the insight that can be given by a stochastic inversion. While the goal of the algorithm is to be applicable to a wide range of inversion problems, this paper primarily focuses on its application to 1D MT inversion for development. Some of the assumptions and limitations associated with the algorithm are mentioned within this context, though the concepts are more general.

2.1 Background and Framework

For clarity, the entire inverse problem can be described by the flow chart shown in Figure 1. Each box represents a certain type of data that is input to and/or output from each named process. $\boldsymbol{\Theta}$ is the set of all possible properties of interest (rock type and depth for MT), $\boldsymbol{X}$ is the set of all indirect properties that can be measured (resistivity and depth for MT), d is the measurement data (impedance/apparent resistivity/phase for MT), and ε is the collection of all error that is incurred through the measurement. The superscript * is used to indicate the real underlying values, and subscript j is used to denote one of many possible model predictions. Treating the information flow as a stochastic process, each part can be modeled using Bayes' theorem.

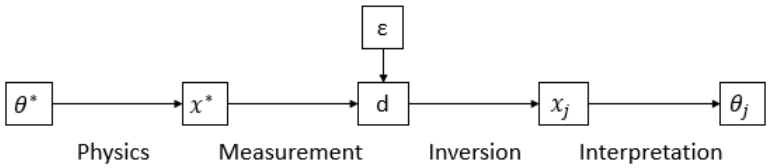

Figure 1. High-level flow chart showing the flow of information through the inverse problem. There are three basic data types that are passed between the different processes, each of which can be described using Bayes' theorem.

The step labeled Physics is described by:

$$P(x^* \mid \theta^*) = \frac{P(\theta^* \mid x^*)P(x^*)}{P(\theta^*)} \tag{1}$$

where θ^* is the property that exists and x^* is the indirect physical response of θ^*. Because the physical response is expected to be deterministic, it is assumed that $P(x^*) = P(x^* \mid \theta^*) = 1$.

The measurement process is described by:

$$P(d \mid x^*) = \frac{P(x^* \mid d)P(d)}{P(x^*)} \tag{2}$$

where d is the measured data. This term illustrates the uncertainty in the measurement, namely through the culmination of error in the measurement (as illustrated in Figure 1). Identification of all sources of error is important in understanding the limitations that the inversion will be subject to in all future steps. The likelihood of equation (2), $P(x^* \mid d)$, is related to how well the response data fits with the measurement given the probability distribution of the error. This is the likelihood to which most inversion methods refer when determining a solution's likelihood. This is only part of the story.

The inversion process itself can be described as:

$$P(x_j \mid d) = \frac{P(d \mid x_j)P(x_j)}{P(d)} \tag{3}$$

where x_j is the most likely set of parameters to describe the indirect property that was measured. These sets of parameters may also be referred to as "models" as they provide a representation of the physical state that is being approximated. The prior, $P(x_j)$, can be used to describe information that is already known about the system so that a certain set of parameters within the probability space can be discounted. Much of the variation in inversion methods rely on how the likelihood, $P(d \mid x_j)$, is calculated. As mentioned previously, this is where most inversion methods

The final step of the inverse problem involves interpreting the values of the parameters as the properties that are needed for the decision-making process. This interpretation process can be described as:

$$P(\theta_j \mid x_j) = \frac{P(x_j \mid \theta_j)P(\theta_j)}{P(x_j)} \tag{4}$$

where θ_j is the interpretation of x_j into the properties desired. The parameters found from the inversion may, themselves, be non-unique to the properties that they describe. Trainor-Guitton et al (2011) describe the need to understand the geophysical interpretation as a part of establishing a data reliability measure. While this is non-negligible, it not considered for the development of ICA.

Value of information (VOI) analyses require an understanding of the value of perfect information as well as a measure of how imperfect the data may be (i.e. data reliability). Data reliability can be used to determine what inversion methods are better than others as well as provide an objective way of comparing different measurement techniques. The data reliability, R_θ, can be described as:

$$R_\theta = P(\Theta = \theta_j \mid \Theta = \theta^*) \tag{5}$$

where Θ is the random variable representing the set of all possible properties. In simple terms, (5) determines the probability that the solution provided by the inverse process reflects the measured value. By inspection, it is evident that it is desirable to maximize the probability of any of these probabilities.

ICA, as well as most inversion methods, focus on a sub-problem that can be described as:

$$R_x = P(X = x_j \mid X = x^*) \tag{6}$$

where R_x is related to the overall data reliability, but focused only on the parameter space rather than the full property space. This differs from the full inverse problem in that it does not include the effect of the physical transformation from property to parameter nor the interpretation of properties from the resulting parameters. Figure 2 below shows a detailed view of the information flow for this portion of the inverse problem.

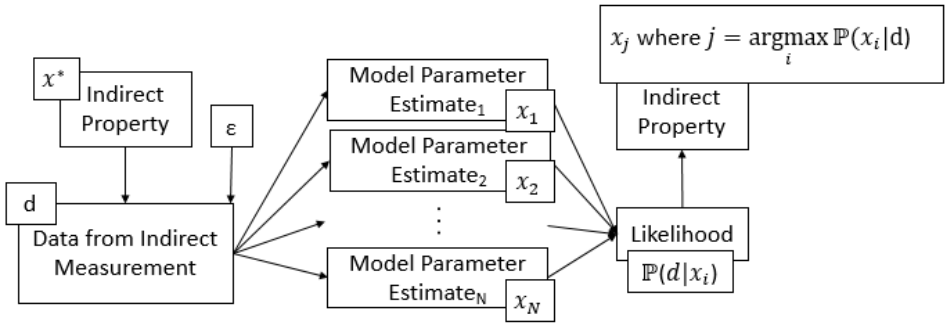

Figure 2: Chart showing the flow of information in the measurement and inversion processes.

Any inversion must deal with the selection of a parameter estimate that has the highest likelihood of occurring. Figure 2 shows that N parameter estimates are evaluated by the inversion, through which a likelihood is determined for each parameter estimate. An x_j is selected that has the highest product of likelihood and prior probability $P(x_j)$ as can be seen from equation (3) since the evidence is independent of the parameter estimate $P(d)$. For a set of parameter estimates whose prior probability is identical, then the likelihood is the sole parameter determining which parameter estimate is selected.

ICA considers a parameter estimate's prior probability as a binomial variable based on the error bounds of the data. If a set of parameter estimates falls inside the confidence interval of the data's error, then $P(x_j)=1$. Otherwise, $P(x_j)=0$. Due to the nature of the algorithm (described in the following section), all parameter estimates that have $P(x_j)=1$ are considered to be equally likely to be the correct representation of x^* and are considered a part of parameter estimate sub-set $\boldsymbol{X}_M$ resulting in a total of M possible parameter estimates.

Each parameter estimate within a parameter estimate sub-set, $x_m \in \boldsymbol{X}_M$,, can be compared to the others to determine its likelihood provided M is sufficiently large. Each dimension k of the parameter space ($\{s_1, s_2 \ldots s_K\} \in \boldsymbol{S}$) can be discretized into r discrete intervals to achieve a desired accuracy for the dimensions of $\boldsymbol{S}$. The discretized parameter space is composed of $c = r^k$ discrete and mutually exclusive subsets of $\boldsymbol{S}$ denoted by $\{D_1, D_2, \ldots, D_c\} \in \boldsymbol{D}$. D^* can be defined as the subset of $\boldsymbol{D}$ through which x^* passes. This likelihood

$P(d|x_i)$ can be defined by analyzing the frequency that each parameter estimate within $\boldsymbol{X}_M$ that passes through each element of $\boldsymbol{D}$ using the indicator function:

$$P(D_a \in D^* \mid \mathbf{X}_{\mathrm{M}}) = \frac{1}{\mathrm{M}} \sum_{b=1}^{M} \mathbf{1}_{D_a}(x_b) \tag{7}$$

Because $\boldsymbol{D}$ is a discrete variable but x is continuous, it is evident that the probability in (7) will be dependent on the chosen discrete interval size r. This sensitivity is not analyzed in depth here and will require further research.

Equipped with the probability from (7), the relative likelihood of each set of parameter estimates can be defined by the average probability of the discretized bins through which the parameter estimates pass, similar to the processes described by Mosgaard and Tarantola (1995) and Ramirez et al (2005). Defining E as the set of $\boldsymbol{D}$ bins through which x_m passes and r_m as the number of bins within E:

$$P(d \mid x_i) = \frac{1}{r_m} \sum_{u=1}^{r_m} P(E_u \in D^* \mid \mathbf{X}_{\mathrm{M}}) \tag{8}$$

This is the likelihood term that can be used within (3). Using this, the most likely parameter estimates can be chosen using:

$$x_j \text{where} j = \arg\max_i P(x_i|d) \tag{9}$$

2.2 Algorithm Description

A flow chart showing the procession of the algorithm is shown in Figure 3.

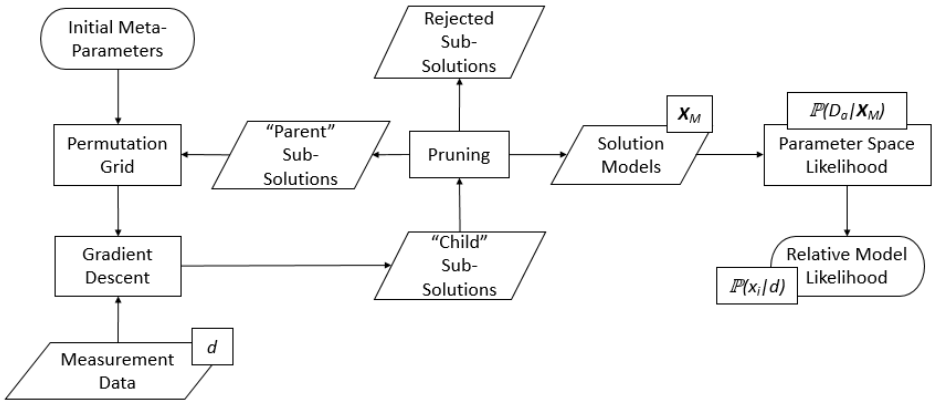

Figure 3. Flow chart of the ICA algorithm. The Permutation Grid-Gradient Descent-Pruning loop continues for as many iterations as are identified in the meta-parameters. The decision process for the Pruning of the "Child" Sub-Solutions is described in detail in Section 2.2.4.

2.2.1 Initial Meta-Parameters

The initial parameters define the extent of the parameter space and the starting points for the initial coarse models. The coarseness (resolution) of the first set of models is defined by how many layers will be used in the first calculations. For these layers, resistivity and depths can be divided into predetermined subdivisions. For this analysis, the model is initialized with three layers where each layer is described by a resistivity and a depth (six variables in total per model).

Maximum and minimum values for resistivity are selected as well as the total depth the models are expected to reach. These values are not hard limits, but set the initial boundary of the search space. If sub-solutions approach these boundaries, the boundaries are relaxed for the following iteration. Sub-solutions are sets of parameter estimates that meet the pruning criteria (detailed later in Section 2.2.4) to become the parents of the next more-complex iteration while not satisfying the constraints of being a possible solution with the set $\boldsymbol{X}_m$.

The data provided for this analysis reported a one standard deviation error for each measurement point, though the sources of error in the calculation were not made clear. The data also contained some sharp corners in the data that were not considered to be physically feasible. It is expected that the reported error does not fully account for the total error within the experiment. An additional

measurement allowance was added directly to the error bounds of the resistivity response (in Ohm-m) and phase (in degrees) in order to account for this additional error.

2.2.2 Permutation Grid

For the first iteration of the algorithm, the initial parameters are used to generate a matrix containing all permutations of models for the number of layers given, whose properties are set as the limits of the parameters and subdivided into the number of values decided in the meta-parameters. This sets up a multidimensional grid of evenly-spaced parameters throughout the whole parameter space.

For each subsequent iteration, the previous sub-solutions are used to seed a new permutation grid. For each incumbent model, each layer of the 1D model is subdivided into a number of layers. For this proof of concept each layer is divided into three layers. The boundaries of the new layers' parameters are set to the appropriate value. Resistivity boundaries are set to the same as the initial conditions while the thickness boundary is set by the depth of the incumbent model being separated. The boundaries for the incumbent values are reduced as a form of annealing around the solution. These boundaries are used to compose a grid of values for the new layers, while the values for all other parameters are the same as the incumbent parent model. The grid is composed of every permutation of parameters for the new layers at the defined boundaries and meta-parameter-decided discretization. Because each layer of each incumbent model is broken up this way, the number of models to be calculated grows exponentially.

This process of breaking a simple layer into component layers is the process of adding complexity to the model. When the new, more complex layers maintain identical resistivity values as the parent model, it can be shown that the MT response of both child and parent will be identical. By adding the new layers, we are expanding the degrees of freedom of the problem, which guarantee that at least one child will have an improved fitness value beyond the parent. By creating a grid of values, a near-exhaustive search is generated. The succession of starting with a simple model and iteratively adding complexity allow for large portions of the search space to be discounted with relatively low computational demand, reserving computation time for later iterations.

2.2.3 Optimization Process

All child models generated by the permutation process are optimized using a gradient descent with the least sum of squares (LSS) between the model's response and the measurement data as the objective function. Only the parameters generated from the permutation matrix are allowed to be perturbed by the gradient descent function. The fixed parameters are assumed to have been optimized already in a previous iteration; and from the exhaustive nature of the algorithm, it is guaranteed that each layer is broken into more layers to be analyzed.

The gradient descent algorithm used to optimize this problem is an unconstrained interior-point method with penalties assigned to values that exceed the predetermined boundaries. The step size can be determined in the initial meta-parameters and can change the computation time considerably. Different step sizes for each parameter type can be accommodated to help normalize the gradient descent as disparate ranges will affect the slope of the basins of attraction.

Because this method utilizes gradient descent, each grid point is pulled towards the lowest point of its local basin of attraction. Because the search space is evenly gridded, comparing the starting and ending point of each model with respect to the search space yields information about the shape of the objective function. Namely, if a grid point is attracted to a hyper-volume that is not adjacent to its starting point, it is implied that the hyper-volumes that were skipped do not contain a solution, regardless of future perturbation. Discarding these parts of the search space reduces the computational expense of the ICA algorithm.

Initially, simple models result in a relatively smooth objective function that does not dip very deeply towards an adequate fit. With each successive iteration (and resulting increase in complexity), the objective function becomes more textured and dips deeper towards an adequate fit, with the possibility for multiple basins with an adequate fit. The depths of these basins – the fitness value – can be used to determine what to do with these optimized models.

One particular difficulty of the gradient descent for the 1D MT inversion is that the fitness value is a function of two different types of responses: the apparent resistivity and phase. Because the range of these values is variable with the different iterations, a coefficient is added to the phase part of the LSS equation. This multiplier is intended to adjust the weight that a mismatch in phase will have on the overall fitness such that the gradient descent will try to fit both the phase and apparent resistivity equally. This value is set as a meta-parameter.

2.2.4 Sub-Solution Pruning

The optimized models that result can be subjected to a number of criteria to determine whether they should be parents of more complex models in subsequent iterations. The pruning methods implemented for this proof-of-concept can be divided into two subsets: computational and absolute criteria.

Computational criteria trim away models to attempt to preserve only the best models and conserve computational resources. The severity of these criteria can be adjusted within the initial meta-parameters. For this proof of concept a fitness limit is imposed that decreases exponentially with each increasing iteration as shown:

$$Fit_t < (C1 * Fit_{\max})^{-K1*t} + Fit_{\max} \tag{10}$$

where Fit_t is the fitness that will not be pruned, $C1$ is a coefficient to set the initial limit, Fit_{max} is the fitness obtained by a set of parameter estimates that match the outermost error boundaries, $K1$ is a decay coefficient, and t is the iteration number that is being computed. Because the first iterations are expected to be too coarse to closely match the measurement, a larger error is allowed, while later iterations require a better fit. If a model is not below this fitness threshold, the model is removed.

When generating the permutation matrix, only a set number of models are allowed to have children. For this proof of concept, 300 models are passed to the permutation matrix generator; the top 150 models are used while the balance of the incumbent sub-solutions are divided into 150 equal parts, and one set of parameter estimates from each part is used.

A clustering algorithm is used as an absolute criteria (discussed below), but requires large computational resources to calculate. In order to calculate this efficiently, a limited number of models are included in the calculation. These constraints can be adjusted to match the available computational power.

Absolute criteria are applied to remove those solutions that cannot possibly be a solution. All duplicate models are removed. All children models whose fitness has not improved from their parents are removed. Models are clustered together wherein only one model from the cluster is preserved. This clustering is done in relative terms to accommodate the variety of parameters that must be compared. Rather than using the absolute Euclidian distance between points as in k-means clustering, relative changes between points are used. For this if the sum of the relative changes between all points is less than some value set as a meta-parameter, the models are considered clustered together and only one is kept. No sensitivity analysis has been performed on this parameter.

Models whose MT response is within the error bounds of the measurement data are set aside as possible solutions and do not proliferate into subsequent iterations. Along with the ability of adjusting the confidence in the measurement data, the confidence also defines the number of data points (with error) the response function must fit within in order to be considered a solution (i.e. a 95% confidence will require that the response data fall within 95% of the error margins).

The order that the pruning methods are applied may have an impact on the results. They should be applied such that the maximum number of sub-solutions are created with the available computing capacity in the time allowable.

2.2.5 Iteration – Increasing Model Complexity

All models that were not pruned are passed back to the permutation matrix generation stage so that each sub-solution can be used to create the next set of models that will provide a better fit to the measurement data. These that are passed on are considered to be sub-solutions and have limited potential value in understanding the uncertainty in the overall inversion.

2.3 Uncertainty Assessment

The uncertainty assessment follows a very similar process to that of Ramirez et al (2005) where the parameter space is discretized and similar models are grouped together. Where Ramirez et al use a k-means clustering algorithm to group similar models, ICA uses all solution models for a likelihood comparison at each discretized element. Each possible solution is then analyzed to compare its overall likelihood across the entire parameter space to determine the uncertainty.

Once a predetermined number of iterations have been run or a minimum threshold of solutions have been found, the solutions can then be compared to provide an assessment of the uncertainty associated with the parameters of the inversion. Each model in the solution set represents the least-complex version of possible models within its discrete solution space, which provides insight to the resolution that a given measurement is capable of providing.

Rather than treating the likelihood of any individual solution as being proportional to its fitness, each model is assumed to represent the body being measured with equal likelihood. Because a solution is removed as soon as it meets the minimum error requirements, the total solution set will not become saturated with related children, leaving a population of relatively unique solutions. These solutions can all be compared to each other through plotting them in the parameter space (resistivity vs depth for this example). These plots can show where all solutions are in agreement and disagreement with each other.

By discretizing the parameter space into an acceptable resolution, each block in the parameter space can be assigned a probability of occurrence through comparison with the frequency the parameter appears within the solution set. If every solution shows the same parameter estimates, the method suggests that there is a relative certainty that the condition exists. The likelihood of a parameter's existence is proportional to the frequency that the solution set agrees with the parameter space's likelihood.

Using this parameter space likelihood, the relative likelihood of each solution model can be computed. This relative likelihood is determined by the average likelihood of each element of the model within the parameter space. The model with the highest relative likelihood best describes the nature of the measured structure.

The likelihood of these models, however, is sensitive to the resolution of the discretized parameter space. The model with the highest certainty has been observed to change as the resolution of the parameter space is adjusted for the calculation of the likelihood. This exact sensitivity has yet to be explored.

2.4 Assumptions and Limitations

The MT model shown here does not require many interrelated properties. The 1D layer half-space Earth forward model relies only on two related properties: layer depth and resistivity (Pethick, 2013). It is expected that the parameters being estimated by the inversion have known possible boundaries. The algorithm does allow for the ability for those boundaries to be somewhat flexible.

It is assumed that no precise prior information is known. Prior knowledge can be added as a means of limiting the search space of the algorithm, though the intent of the algorithm is to assume that very little is known about the system as would be the case when analyzing a system from measurements alone.

While most inversion methods rely on finding the global optima for the fitness, it is assumed here that local optima contain important information about the texture of the objective function that allow for narrowing of the solution space.

While it is the goal of any inversion algorithm to describe a system in full, ICA also attempts to identify the resolution limits of the measurement system used. As simple models progress in complexity with each successive iteration, the sub-solution error approaches the measurement values. The rule of superposition implies that acceptable children models will yield better fitness than their parent models. Once a solution meets the minimum threshold for error, the model cannot yield any more information since all children will still fit within the error threshold.

MT data was retrieved from the Iceland Geosurvey (ÍSOR) that was used for the Karlsdóttir et al (2012) inversion that was given in the .EDI standard format. An error measurement is given in the report as a standard deviation from the reported mean. It is assumed that this error is related specifically to noise in the data and does not account for other measurement error similar to the measurement errors discussed by Gamble (1979).

The ICA method computes a sub-solution set within every iteration from which the sub-solutions are accepted for perturbation with more complexity or rejected. The rejection process, or pruning, is accomplished through a candidate solution meeting several criteria that are set by meta-parameters. Aggressive pruning results in taking out too many models and not adequately sampling the search space.

The determination of the discretized space $\boldsymbol{D}$ can potentially cause an issue. Two near-identical sets of parameter estimates that are parallel to each other, but divided by the boundary of the discretization would potentially appear to be very different. Determining the likelihood through usage of a clustering algorithm, although computationally expensive, will likely yield more consistent results.

3. RESULTS AND DISCUSSION

Results are shown for the analysis of a single MT measurement of the Þeistareykir geothermal field at THR031. While the analysis of the solution set $\boldsymbol{X}_m$ is the object of the algorithm, the results shown here include all sub-solutions in the analysis.

3.1 Results

3.1.1 Initial Meta-Parameters

The initial meta-parameters used for this analysis described a three-layer earth model with resistivity and depth values for each layer. The initial maximum resistivity was set to 1000 Ω-m while the initial minimum resistivity was set to 0.1 Ω-m to aim below the 0.3 Ω-m of seawater cited by Karlsdóttir et al (2012). The initial basement depth was set to 25,000 m while the algorithm allowed the gradient descent to minimize any depth to 1 m.

From the reported measurement error, a confidence interval of 95% was used to determine the length of the error bars from the reported data. An additional +/- 2 Ω-m and +/- 2° of error was added to the apparent resistivity and phase response of the measurement, respectively, to account for additional measurement errors. This results in a possible 9490 solutions after 6 iterations. Two similar runs have been analyzed using no additional error as well as +/- 1 Ω-m and +/- 1° of error. The former resulted in finding no possible solutions while the latter resulted in a possible 12 solutions being found, both likewise after 6 iterations. Understanding how the additional error impacts the confidence in the results is beyond the scope of this paper.

The nominal apparent resistivity ranged from 11.19 to 107.07 and the nominal phase shift ranged from 24.16° to 72.01° (0.4216 to 1.257 rad). The measurement contained 69 apparent resistivity and 69 phase shift data points, each with respect to a frequency ranging from 0.0014 to 265 Hz.

The phase coefficient used in the LSS calculations (see Section 2.2.3) is set to 10,000. The sensitivity of results to this figure has not yet been explored.

For these results, $C1$ is 50, and $K1$ is 0.9. The resulting maximum fitness function value, Fit_{max}, is computed to be 3976.

When selecting from incumbent sub-solutions to generate the next iteration's permutation matrix (as described in Section 2.2.4), only 1000 models are used. To calculate the clustering of candidate models, the best-fit 10,000 are used. The total relative distance threshold used for these clusters is 1.

The figures on the following pages show the solutions after 6 iterations. Figure 4 shows the responses of both apparent resistivity and phase shift as compared to the measurement and its error. Figure 5 shows the solutions in the parameter space with resistivity as a

function of depth. The variety of models shown reflect the variety of parameter estimates that produce responses (x_i) within the margin of error of the measured response (d). Figure 6 shows the likelihood of each block of the discretized parameter space based on Equation 7. Figure 7 shows the distribution of the fitness values for all solutions as well as the relative likelihood of each model to the others given the likelihood of the parameter space. The parameter estimates are organized in order of fitness, with the best fitness ranked as 1 on the left. Note that the most likely model is not the best-fitting model. Figure 8 shows a pedigree chart for the solutions that tracks the relative positions of each solution by their relative ranking by fitness as sub-solutions in non-final iterations. When a large number of solutions can be traced back to earlier sub-solutions, this suggests that the pruning method has not been too aggressive and that the full parameter space has been searched.

Additional descriptions of each figure are provided in each figures' caption.

3.1.2 Results of 6 Iterations – Up to 13 Layers

The most coarse parameter space is evaluated in the first iteration. Due to the meta-parameters chosen, 243 models were evaluated. 33 sub-solutions passed the described pruning algorithms. Figures 3 through 6 show the results of the first iteration.

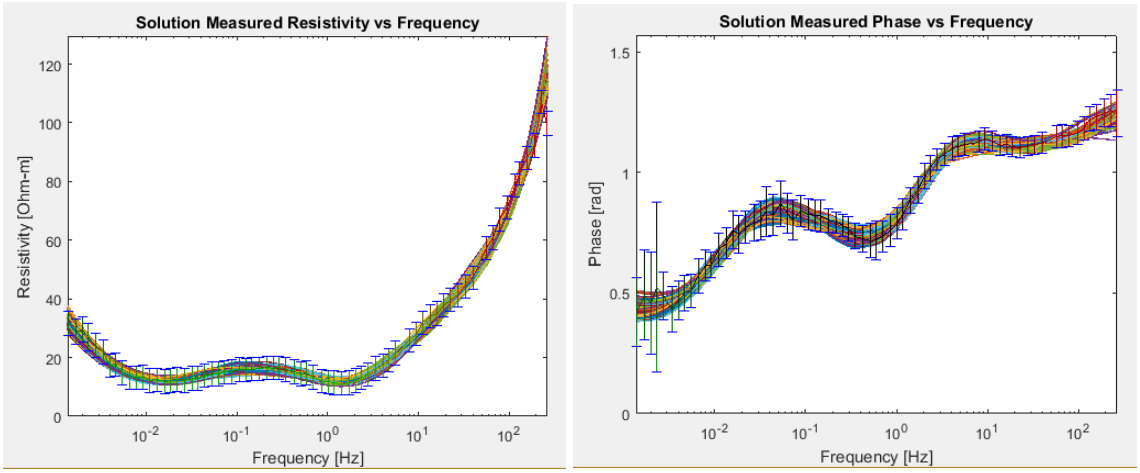

Figure 4. All 9490 solution responses are shown as compared to the measurement data with error bars. The graph on the left shows the apparent resistivity in Ω-m while the chart on the right shows the phase shift in radians. The 95% condfidence bound and the added error are responsible for the error bar widths. The 95% confidence also allowed up to 5% of the error bars to be exceeded. Notice how wide the bands of the solutions are in this case. It also worth noting the highest-frequency apparent resistivity – and how it aligns with so few models; and that no arrangement of parameters can accompish that sharp a bend.

Figure 5. The parameter space of the 1D MT is shown with the 9490 resulting solutions. The bolded black model indicates the parameters of the most likely model. The dashed line indicate the approximate solution given by the Karlsdóttir et al (2012) 3D inversion. The "most-likely" model is dependent on the discretization of the space (shown in the next Figure). If the resolution of the discretized space is changed, the most-likely model will also change.

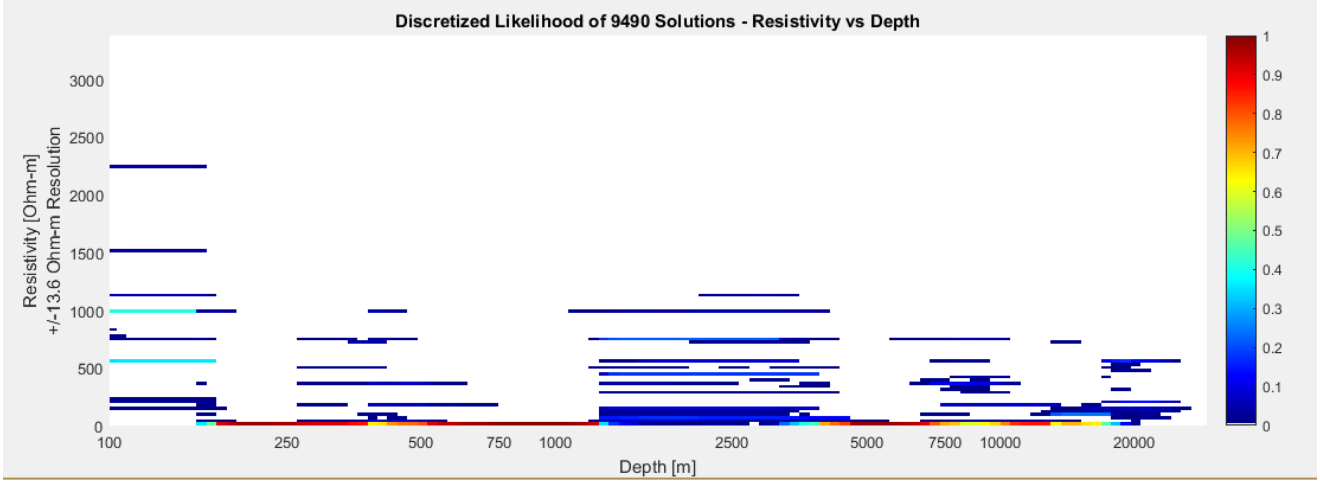

Figure 6. The parameter space likelihood described by Equation 7 can be seen. The color scale shows the likelihood of finding the shown resistance value at the given depth for the 9490 sub-solutions. For this particular measurement, the certainty of the resistivity with depth alternates. While the exact value for the high-resistivity regions is unknown, it is interesting to note the certainties associated with the fact that there is a high-resistivity body present between 1250 and 4500 m.

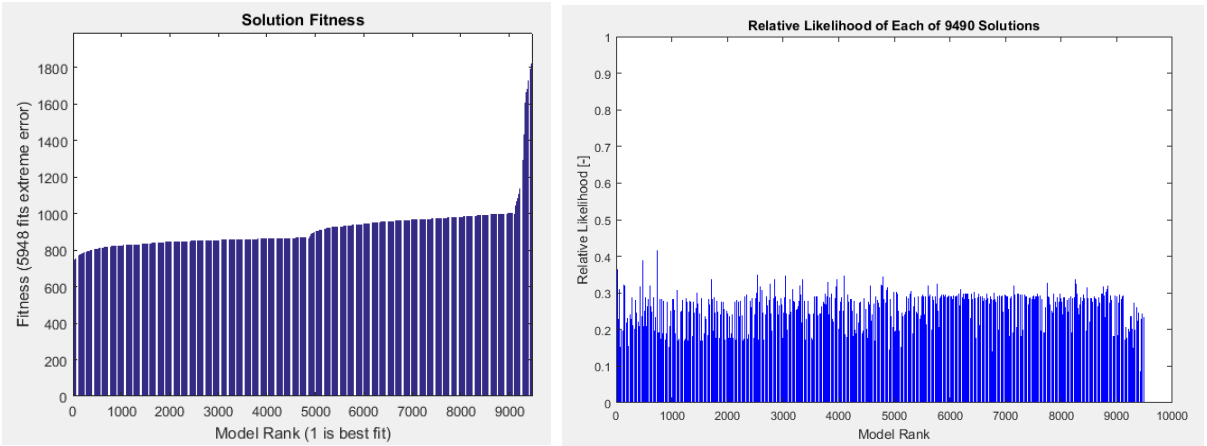

Figure 7. On the left, the fitness value of each of the 9490 solutions is shown. The fitness associated with following the error bars is 5948, so even the worst-fit solution has a fitness quite far from that maximum threshold. The fitness also evidently breaks into 3 sections (roughly 1 – 4800, 4801 – 9000, and 9001- 9490) that is separated by which iteration (ie parameter resolution) the solution was found in. The first set correlates to the last iteration, the middle segment is the 5th iteration, and the last segment is fromt the 4th iteration. It must be noted that a single solution was found in the 3rd iteration. On the right, the relative likelihood of each model is shown. It is interesting to note how the best-fit models of the group (far left) share the least in common with the other models suggesting they have a low relative likelihood. Model 739 has the higest relative likelihood with 41.6%. The identity of the model with highest relative likelihood will change as the size of the parameter space discretizations are changed.

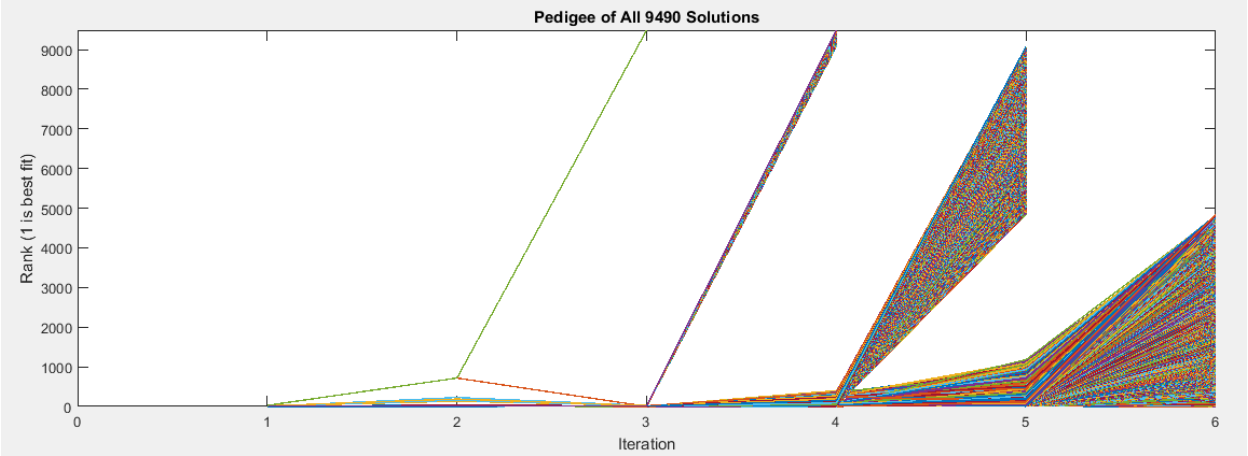

Figure 8. The pedigree chart shows how the different solutions are related. With the reduction of scale in print, it is difficult to see some of the details that come to light from this chart. It is easy to see how the earlier solutions that are found have the least fitness (highest rank). It is also possible to see how many solutions are retrieved from each iteration. Following a group of solutions backwards to their convergence shows how diverse each iteration was at selecting a solution. Notice how the solution found in Iteration 3 came from the worst-fit parent of all solutions in Iteration 2. It is common to see that the best-fit models produce children that are not as fit as some other models.

3.1.3 Comparison to previous inversion

Comparing the results from the ICA method to that of the 3D suggests that both approaches are reasonable. Figure 9 shows the North-South 2D cross section of the Þeistareykir field that intersects the measurement location of THR031 copied from Karlsdóttir et al (2012). The red arrows in the figure show the location of THR031.

Because the Karlsdóttir et al data was performed at a 3D inversion, it took into account the surrounding rock physics while the 1D inversion in this paper did not. Even without the influence of the surroundings, ICA is able to pick out similar structures as the 3D inversion, though with discrepancies in the values of resistivity. Ideally ICA could be expanded to solving the 3D inversion problem, though much more development would be needed.

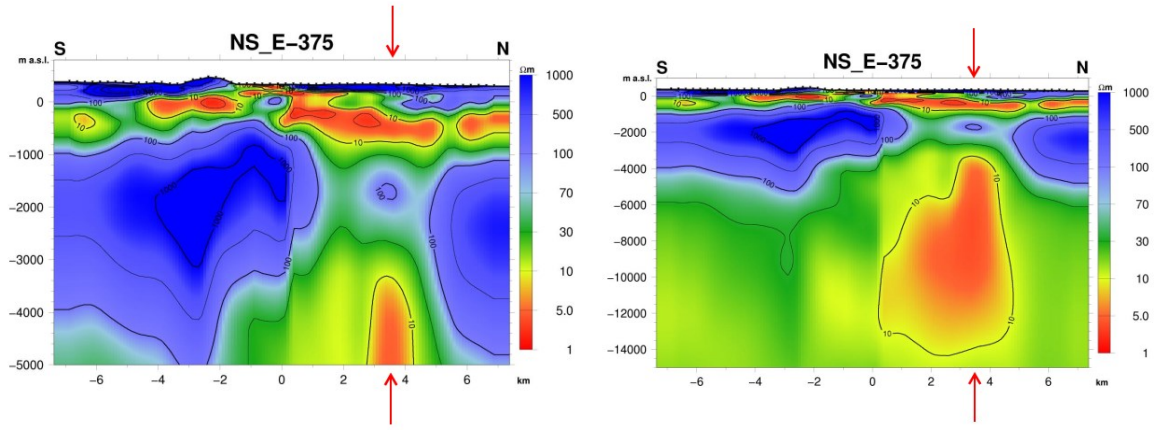

Figure 9. North-South 2D cross section of Þeistareykir field that intersects THR031 measurement location. Red arrows indicate the location of the measurement. The left picture is focused on depths from sea level to 5 km. The figure on the right shows the same area, but with a depth to 14 km. Note that the surface here does not start at 0 m, though figures 4 through 24 indicate the surface as 0 m.

3.1.4 Computational Requirements

The results shown here were processed using Google Cloud Platform (GCP) running MatLab R2016b on Ubuntu 14.04. The GCP machine had 32 vCPUs (16 real cores) with 120 GB or RAM and 25 GB of storage capacity. The algorithm ran all six iterations in 23.7 hours. The total size of $\boldsymbol{X}$ was 8,725,862 models.

3.2 Discussion and Next Steps

Throughout this report, a number of assumptions and limitations have been described. It is understood that the method is not complete in its current state and will require further development in order to be shown as a viable tool for geophysicists and other inversion analysts. Here are outlined the development plans that are in progress.

This method will be applied to multiple measurements in the Þeistareykir geothermal field to allow for the development of some 2D cross-sections of the field that coincide with existing wells. These cross-sections will be compared to the existing inversion created by Karlsdóttir et al as well as to drill cuttings that coincide with areas that have been analyzed. This will provide the framework to begin understanding how reliable MT measurements are and provide a metric with which to compare future developments in inversion techniques, especially where they apply to geophysics.

This paper has used a problem assuming a lack of prior information about the system being measured. Prior information can be added in the permutation matrix generation step. Parameters can be added such that the prior information is allowed to stay within a locked boundary. The exact implementation of this has yet to be completed.

Several variables have been mentioned throughout this paper that have been assumed without any insight on sensitivity. These will need to be explored in greater detail to determine the best operating scenario for the process. It is uncertain whether the meta-parameters used for the problem shown here will be applicable to a broader scope of inverse problems.

4.1 CONCLUSION

Iterative Complexity Addition provides a different approach to solving inverse problems. Rather than predetermining the parameter resolution, the complexity addition steps allows the algorithm to find minimally complex solutions that would result in a response within the error margins of the measured data. This provides some information about the resolution that a given measurement can provide about a system. The plurality of solutions produced by the ICA algorithm allows for a statistical comparison of the credible solutions to identify the most likely set of parameter estimates with which to build a conceptual model of a system. Compared to a MCMC process, ICA is still considerably computationally expensive. The treatment of error, however, takes away the need to repeat the calculation with multiple prior distributions to try to determine likelihood. The algorithm is applied to a single 1D MT measurement and shown to have feasible results that agree with prior studies on the same dataset. There is a need for further development of the process to quantify the data reliability, to develop its application to 2D and 3D problems, and to validate the approach.

REFERENCES

Bratvold, R., Bickel, J., & Lohne, H. (2009). Value of Information in the Oil and Gas Industry: Past, Present, and Future. *Society of Petroleum Engineers Reservoir Evaluation and Engineering*, 630 - 638.

Chen, J., Hoversten, G., Key, K., Nordquist, G., & Cumming, W. (2012). Stochastic inversion of magnetotelluric data using a sharp boundary parameterization and application to a geothermal site. *Geophysics*, E265 - E279.

Gamble, T., Goubau, W., & Clarke, J. (1979). Magnetotellurics with a remote magnetic reference. *Geophysics*, 53 - 68.

Karlsdóttir, R., Vilhjálmsson, A. M., Árnason, K., & Beyene, A. T. (2012). *Þeistareykir (Theistareykir) Geothermal Area, Northern Iceland: 3D Inversion of MT and TEM Data.* Reykjavík: ÍSOR - Iceland Geosurvey.

Menke, W. (2012). *Geophysical Data Analysis: Discrete Inverse Theory, 3rd Ed.* Elsevier Inc.

Mosegaard, K., & Tarantola, A. (1995). Monte Carlo sampling of solutions to inverse problems. *Journal of Geophysical Research, 100*, 12431 - 12447.

Pethick, A. (2013, Dec 17). *Tutorial - 1D Forward Modelling (Magnetotelluric)*. Retrieved July 2016, from Digital Earth Lab: http://www.digitalearthlab.com/tutorial/tutorial-1d-mt-forward/

Ramirez, A., Nitao, J., Hanley, W., Aines, R., Glaser, R., Sengupta, S., . . . Daily, W. (2005). Stochastic inversion of electrical resistivity changes using a Markov Chain Monte Carlo approach. *Journal of Geophysical Research, 110*.

Trainor-Guitton, W., & Hoversten, G. (2011). Stochastic inversion for electromagnetic geophysics: practical challenges and improving convergence efficiency. *Geophysics*, F373 - F386.

Trainor-Guitton, W., Caers, J., & Mukerji, T. (2011). A methodology for Establishing a Data Reliability Measure for Value of Spatial Information Problems. *Mathematical Geosciences*, 929-949.

Yang, X. (1999). *Stochastic Inversion of 3D ERT Data.* PhD Dissertation, University of Arizona, Department of Mining and Geological Engineering.