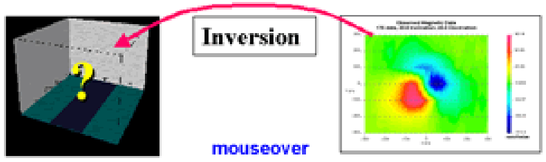
Inversion concepts:
Introducing geophysical inversion

This chapter deals with basic concepts underlying geophysical inversion. Four sections provide an overview of essential ideas without mentioning mathematical details. Extensive use of flowcharts and figures should quickly introduce what inversion does, what is needed, and what can be expected for outcomes.

Note that the current chapter (Inversion Concepts) is conceptual in nature while Chapter 4 (Inversion Theory) covers the same ground in a more rigorous manner. It is recommended that the conceptual chapter be studied first before tackling the details provided in the theory chapter.



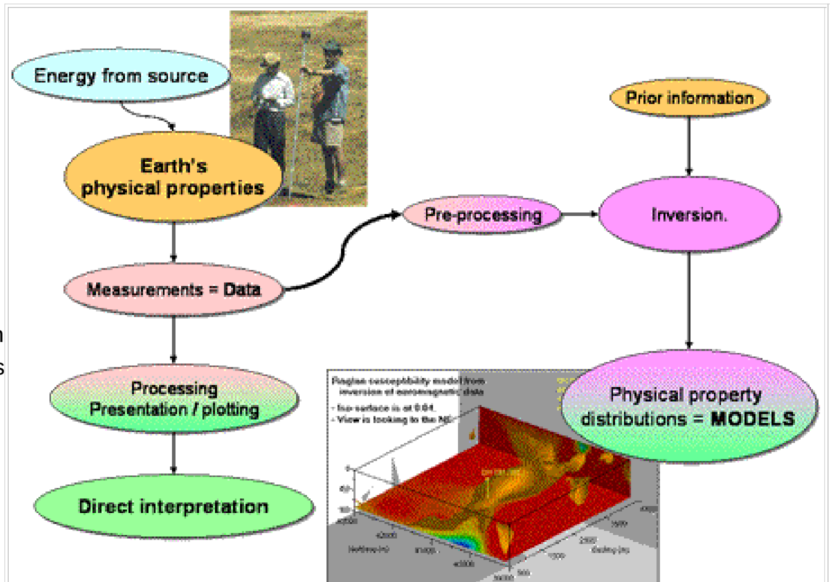
Inversion: Estimating models of physical property distributions based on geophysical survey data.

Overview

Geophysical remote sensing data can be used to help solve practical environmental, engineering or exploration problems. In some cases, when only limited knowledge about the subsurface is required, inferences drawn directly from the data can be sufficient, as Figure 1a illustrates. This is a familiar linear framework in which data are gathered, then processed, plotted and interpreted.

However, when more detailed information about the subsurface is needed, quantitative models of the earth need to be estimated. This is geophysical inversion, as Figure 1b illustrates. Here the framework is not so simple. Data are used to constrain possible **models** of the earth, which have been estimated using some procedure suitable to the particular problem.

This short package does not explain details, but it should provide a conceptual understanding of what is involved when trying to solve the geophysical inverse problem.



1a. Geophysical information can come directly from data.
1b. If models are needed, inversion must be applied.

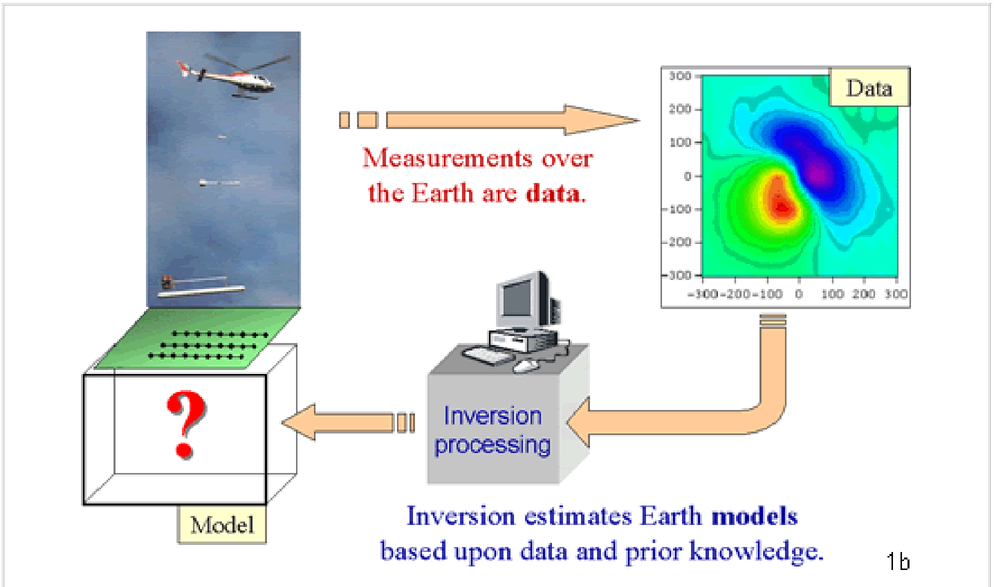
NOTES about the last page: Section **f. Inverting 2D DC data** is challenging at this level. It assumes some understanding of DC resistivity surveying, and some experience with geophysical data. However, it does illustrate what's involved in carrying out inversion work. Readers with no more background than a good understanding of this "Inversion Concepts" chapter will gain a clearer impression of geophysical inversion by reading through the page carefully. There is a slightly more advanced version of this page at the end of **Chapter 4 "Inversion Theory".** If Chapter 4 is to be covered, you might gain more by deferring this page until then.

Inversion concepts:
Outline of inversion

On this page:

This page uses interactive figures to provide a pictorial introduction to geophysical inversion. You will be able to characterize just what the geophysical inverse problem is, and what processes and decisions are involved. Key concepts are: observed versus predicted data; prior information (or knowledge); and that a model is acceptable if it can predict data and if it is consistent with prior information.

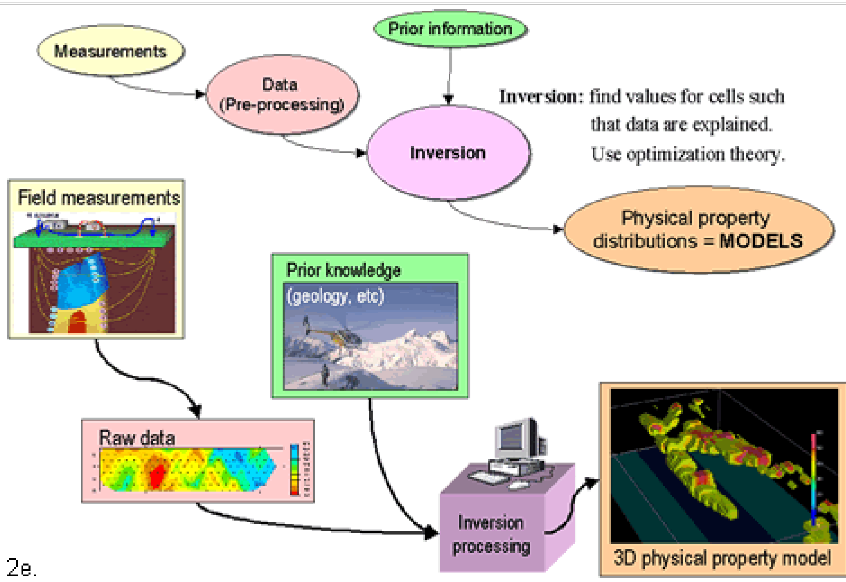
In a typical geophysical survey, we put energy into the ground and record a response, which we refer to as data or observations (Figure 1a). The values of the data depend upon the distribution of physical properties in the subsurface. This is illustrated in Figure 1 which is a generic slide for all geophysical surveys. The goal of the inverse problem is to determine the distribution of the physical property or properties that gave rise to the data (Figure 1b). Unfortunately, this is not rigorously possible in practical surveys because only a limited number of data can ever be recorded and the data are also inaccurate. Nevertheless, approximate solutions can be found, and the methodology is designed to include other information about the problem so that the calculated solution is more likely to represent the true earth structure.



1a.	Measurements of physical processes are gathered over the earth and a data set is produced.
1b.	The purpose is to learn something about what is beneath the surface. Answering this using the data is inversion.

Creating models based upon data

Inversion is a mathematical procedure that can take on several forms. In order to generate understanding about the subsurface without digging or drilling, measurements must be gathered, data must be generated from these measurements, and some degree of understanding about what is being investigated ("prior knowledge") should exist. Then inversion processing can be carried out, using the data and prior knowledge as input. The result will be a set of "models" characterizing how the relevant physical property is distributed in the ground. These models will have characteristics determined by the inversion method used, by the data and by prior knowledge. Figures 2a-2e illustrate.



2e.

2a. ○	Measurements in the field over the region of interest must first be made.
2b. ○	A data set is made from these measurements. More or less data processing may be involved, and plots or graphs must be made to help with interpretation and data quality assurance.
2c. ○	Prior information is what you already know about the problem.
2d. ○	Inversion is the processing step that uses data and prior knowledge as input to estimate models of how physical properties are distributed under the ground.
2e. ◉	Models are estimates of a physical property distribution.

Requirements of inversion

Extracting from the flow chart of Figure 2 we have the basic ingredients of inversion summarized in Figure 3:

- measured data
- prior information
- the inversion algorithm
- a physical property distribution within the volume being studied

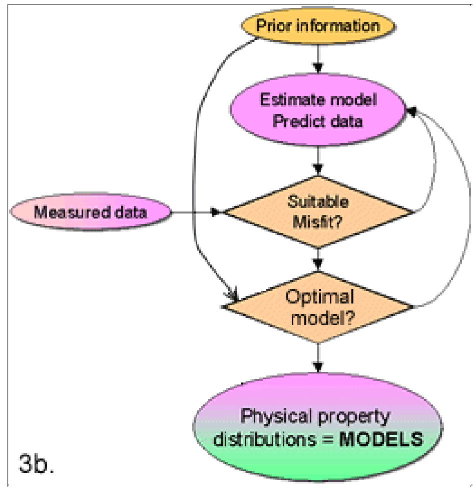
The Inversion section of Figure 3a consists of a model estimation algorithm and two decisions which must be made before the model is considered satisfactory. These are shown in Figure 3b, and they can be summarized as follows:

1. **Model estimation and predicting data:** The inversion algorithm must estimate the values of parameters that define the model. The first estimate will be based upon a prior understanding of the situation. Predicting data means calculating the measurements that would be made by a survey over the model.
2. **Decision 1; Suitable misfit:** One decision is whether the predicted data (caused by the current model) are sufficiently similar to the real survey measurements. In other words, is the misfit between *observed* and *predicted* data sufficiently small?
3. **Decision 2; Optimal model:** The second decision is whether the the model itself is "optimal" in a mathematical sense. Successful inversion requires careful design of this optimization problem.

In order to make these decisions, both **data** and **prior information** are necessary. They both serve to limit the possibilities for a recovered model, therefore we say that measurements and prior knowledge (about likely properties and structures) act as constraints upon the optimization problem.

3a. ◯ A very generic flow chart with essential components of inversion

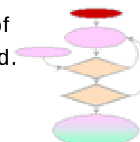
3b. ◉ Principle requirements of inversion:
1: A model must be estimated and resulting predicted data must be calculated.
2: Predicted and measured data must compare favourably.
3: The model must optimal in some well-defined sense.



A solution to the inverse problem is obtained by estimating a model, testing it using misfit and acceptibility criteria, and then iteratively perturbing (i.e. adjusting) the model until the two testing criteria are satisfied. The arrows in Figure 3b show how information is used and where iteration (feedback) occurs.

Of course, once a mathematically acceptible model has been obtained, the geoscience professional must asks whether the result is geologically and geophysically plausible. This generally involves some degree of interaction with people who understand both the fundamentals of inversion and the specific geologic and geophysical problem.

The next section explains some of the implications of the two decision boxes in Figure 3b. Small icons in the shape of this flow chart (like the one shown to the right) will be used to highlight which section of this chart is being discussed. Mouseover on the icons displays the name of the highlighted segment.



Inversion concepts:
The inversion process

Objectives

This page outlines how inversion works without going into details. After absorbing the concepts, you should be able to:

1. Explain the difference between observed and predicted data;
2. Understand why both are needed;
3. Give three reasons why we say that the inversion problem is "non-unique";
4. Understand that there will still be significant "prior knowledge," even though geophysical work may be done at a site where very little previous work was done.
5. Have a preliminary appreciation for various ways of examining 3D models that are built using a discrete set of rectangular cells.

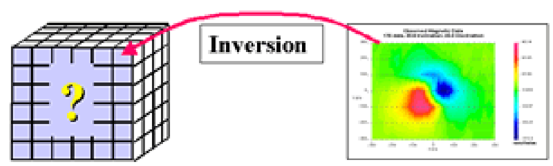
Introduction

In this section we step through the major ideas involved in carrying out practical inversions and we develop a generic flow chart for the process that will be used throughout this CD. The two types of constraints on the model - data and prior knowledge - will be emphasized. We must understand how they can be used to decide whether the computed model is a good candidate for the earth structure. We also talk about the importance of visualizing and presenting resulting models in a meaningful way. Finally, we sum up with a new, more complete flow chart, which will serve as the basis for discussions about inversion throughout the CD-ROM.

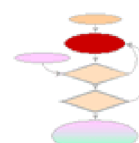
Inversion and forward modelling

To reiterate, the *goal of the inversion is to produce a reasonable model that generated data, given field survey observations, some knowledge about their errors, a decision about how to represent the earth and an ability to forward model*. Also, the distinction between forward modeling and inversion must be clear, as summarized in the following interactive figure.

1a. Calculating data when the model of the Earth is known is called *forward modeling*.
1b. Iinversion is the opposite procedure - estimating possible models when data are available.
1c. For most problems the earth will be represented with discrete rectangular cells.

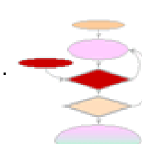


Getting the process started

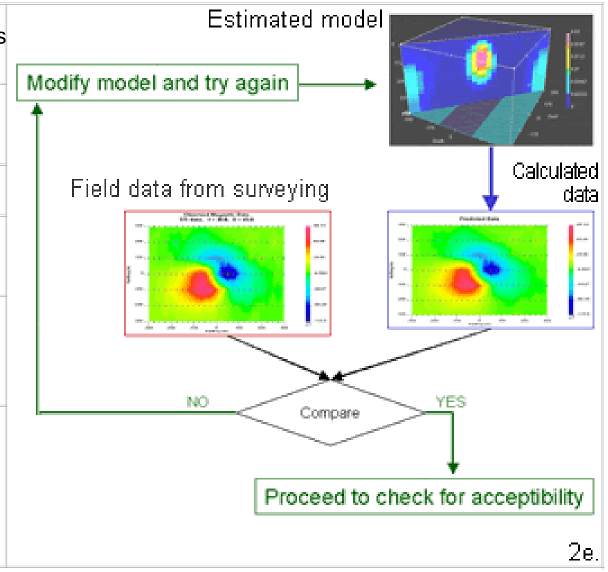


Unlike other geophysical "data processing" procedures such as filtering or data reduction, inversion does not take a raw data set, manipulate it directly and output some sort of answer. That is why the flow chart is not a simple linear path with data at one end and resulting model at the other. Instead, inversion processing starts with an initial estimate for the earth model, then proceeds with a forward calculation on that model to predict what measurements would be if a survey were carried out over that initial model. The resulting data set is called the predicted data. Only at this stage is the actual field data, or "observed data," brought in. These survey results are compared to predictions generated for the initial model. This is the first of two decisions in the inversion algorithm, and both are described in more detail next.

Decision number 1: Fitting the data



Once there is a preliminary model, a predicted data set for that model, and an observed data set collected in the field, the inversion algorithm can go to work on the two decisions that have to be made within the inversion process. (The important business of how to *estimate a model* will be discussed in detail in a subsequent chapter). The decisions are not necessarily made "first" and "second", but we will start by outlining the *misfit* decision using the following figure:

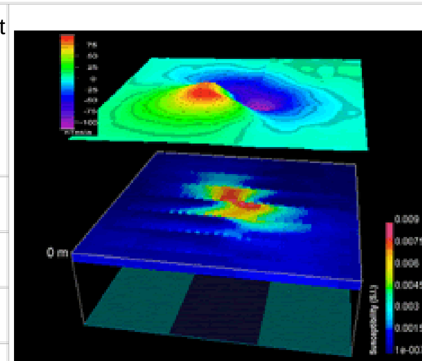
2a. Once a model has been estimated (initially it could be as simple as a halfspace) ...	
2b. ... a data set must be calculated by forward modelling. This data set is referred to as *predicted data*.	
2c. Data from field surveys are now brought in, ...	
2d. ... and compared to the calculated data obtained by forward modelling from the estimated model.	
2e. If the two data sets are not sufficiently similar, a new model is estimated and the process repeated. If the two data sets are sufficietly similar, we can proceed to the next step in the process.	
This set of figures does not explain how each step is done, it simply outlines what is done when testing to see if the model is capable of explaining the observations (field data). Also, we have not explained what the criteria are for deciding whether two data sets are "sufficiently similar." That will be covered in the next chapter.	

A crucial concept: non-uniqueness

The other decision involves a more subtle issue. For most inversion problems that will be discussed in this work, we will have digitized the Earth with more cells than we have data values. Therefore, it will be impossible to find a*unique* value for each cell without additional information. This issue is referred to as the problem of non-uniqueness. All problems like this are called "under-determined," and it is common to say that they have non-unique solutions. That is, all problems with more unknowns than data will have an infinite number of possible solutions. Furthermore, there will be errors associated with every data point, making the problem even more difficult to solve.

Parametric problems: Situations when there are *fewer* parameters describing your model than there are measurements are called "over-determined" problems. They are usually dealt with in a completely different manner, but interpretation of over-determined problems is also non unique. For example, if a particular geometric shape is assumed in the parametric solution (a buried cylinder perhaps) one unique solution will be recoverable. However a different parameterization could be invoked - perhaps a buried multi-sided object - so the interpretation is in fact non-unique. See Section 4.6 for more discussion of parametric problems. The main point here is that, because we want the flexibility of explaining as much of the true complexity of the 3D Earth as possible, non-uniqueness is a pervasive characteristic of solving the geophysical inverse problem.

The next figures illustrate non-uniqueness for under-determined problems, when the Earth is divided into many more cells than there are data values. The data set on top is a synthetic magnetics data set generated by measuring the fields on the surface above a cube of magnetic material buried between the depths of 50m and 75m. Applying inversion with this data set means trying to determine what distribution of magnetic material caused the data.

If the region where magnetic material can occur is constrained to only a thin 25m layer, it is possibile to obtain acceptable solutions (models), regardless of the depth assigned to that constraining layer. The figure shows the inversion solution with the magnetic material constrained to occur only within ...	
3a. the top 25m;	
3b. between 50 and 75m;	
3c. between 125 and 150m;	
3d. between 175 and 200m.	

Two aspects of non-uniqueness have been mentioned (that this is an under-determined problem, and that data are noisy), and there is also sometimes a third aspect. Some physical measurements do not contain adequate information. For example, potential fields data (including magnetic and gravity surveys) do not contain information about the distance between the cause and the effect. In other words, as illustrated in Figure 3 above, it is possible to generate any data set with a thin layer of material if the particles can take on unlimited values of the relevant physical property. The problem is further illustrated with a hypothetical gravity survey shown in Figure 3e below.

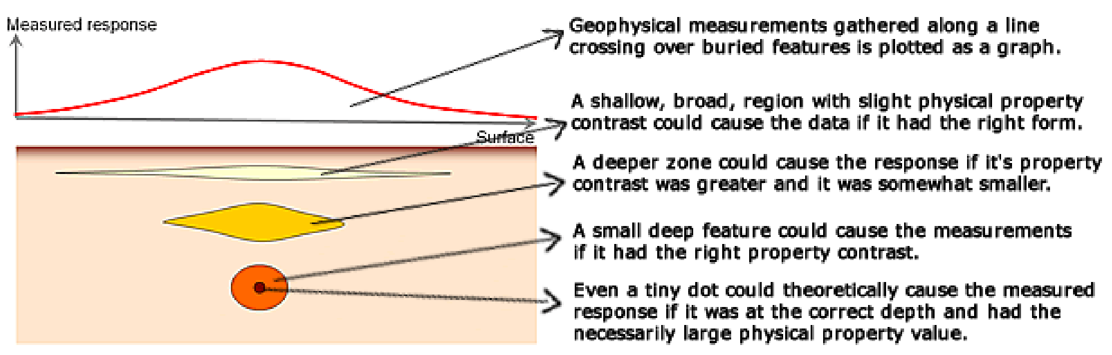
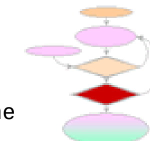


Figure 3e: Measured response of a gravity survey along a survey line (in red) over four different models. All four models are equally capable of causing the data, and the data contain no information about which of the four models is the true model.

Resolving this aspect of non-uniqueness requires modifications to the general solution, as explained in the section entitled "Mods for mag/grav" in Chapter 4 "Inversion theory".

Decision number 2: Optimal model

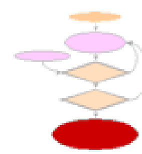


How does the issue of non-uniqueness relate to decision number two, which relates to whether the model is an optimum one. Since an infinite number of models are possible, alternative information must be used to constrain the models so that one can be chosen in favour of all the rest. This is where prior information (what we already know about the problem) comes in. What do we know? In other words, what constitutes prior knowledge? Regardless of how much or how little work has been done at the field site, a significant amount may be known about the geoscience problem. Here are examples:

- There will be **geological information**. For instance, it may be known that background rocks hosting the target are non-magnetic (limestone for example). It may be known that there is non-magnetic overburden that is at least a few metres thick. The basic structural situation may be reasonably well understood from surface mapping and other work.
- There will be **geophysical information**. For example, we know that magnetic susceptibility must take on positive values. Density contrasts, on the other hand, may be either positive or negative depending upon whether "target" materials are more or less dense than host materials.
- There will be **logical information**. For example, it is sensible to look for a "simple" solution. It will be possible to find an arbitrarily complicated model to explain the data but it will NOT be possible to find an arbitrarily simple solution if there is any pattern to the data set. Therefore it makes sense to look for the simplest model that can explain the data. "Simple" may mean as little structure as possible, or as little departure from an initial structure as possible, or a structure that is as smooth as possible.

Selection of an "optimum" model means constraining the infinite number of solutions to those that are consistent with what we already know or assume about the situation. Exactly how this prior information influences the inversion process is discussed in detail in the next chapter.

Usable results



Displaying and using 3D results of inversion often takes care. It is rather easy to give the wrong impression about the result if the model cannot be examined visually in a convenient and versatile manner.

The next figures show the model obtained by inverting the data set shown in Figure 3 above, using a method that allows susceptibility to occur with equal likelihood anywhere in the volume.

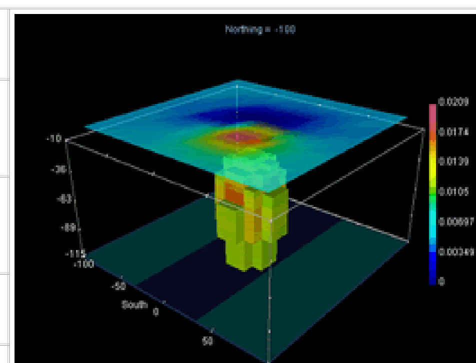
4a. The data set is shown above an isosurface image of the susceptible zone recovered.

4b. An isosurface is an image with all cells with values less than some chosen cutoff value hidden from view. The choice of cutoff significantly affects the *appearance* of the model.

4c. Alternatively, a slice through the volume can be displayed with the cell value coded by colour. This gives a correct impression of the range, and distribution pattern, of values in the mode.

4d. , 4e. , 4f. It can be useful to display slices through the volume in several directions.

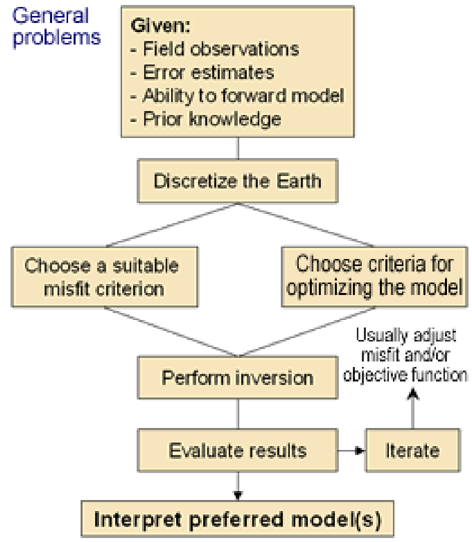
4g. A combination may be useful, and the direction used to view the model may contribute to understanding the model that has been recovered.



Summary - the fundamental flow chart

General problems

What is needed to invert a data set? This flow chart summarizes the requirements for proceeding with inversion of geophysical data for general problems. The flowchart for parametric problems is shown by rolling the mouse over the chart, and is described later in this section. The introductory page for the next chapter is an interactive version of this chart which provides details for each portion. For the present, it is sufficient to introduce the chart, and to summarize the procedure for implementing inversion as follows:



- **Given**: In order to obtain models of the subsurface by inversion, it is necessary to start with field data, estimates of errors and noise on those data, a forward modelling calculation procedure, and well described prior information or assumptions about the situation.
- **Discretize**: Describe the earth by dividing it into cells, each with fixed size and unknown but constant value of the relevant physical property. (There are other ways of describing your model of the earth, but this CD-ROM focusses on the use of rectangular cells.)
- **Choose decision criteria**: The choices made for how data predictions and survey measurements are compared, and how an optimal model is chosen (based upon prior information) are crucial for obtaining useful results.
- **Inversion**: Find values for cells which are consistent with both the measured data and the prior information. The inversion process is implemented using mathematical optimization theory.
- **Evaluate the inversion result**: Use a comparison between predicted data and field measurements, and take into account what is known or expected about the earth's properties and structures.
- **Iteration is inevitable**: No initial outcome should be used without exploring a range of equally possible models.
- **Interpret the result**: models of physical properties must be interpreted in terms of useful geologic or geotechnical parameters. The models must be easy to understand and convenient to manipulate.

Parametric problems

As mentioned above, when there are fewer parameters than data values, the problem does not involve choosing from an infinite variety of potential models - adjusting parameters until the model can reproduce the data as closely as possible is sufficient. Mouse over on the chart to see how the procedure changes in this case. The procedures for parametretic problems are described in greater detail in subsequent chapters.

As the theory of inversion is developed in subsequent sections, it will become clear that the primary difficulties are that the solution is non-unique, and that the necessary mathematical procedures are computationally demanding.

Inversion concepts:
Flow chart for geophysical inversion

Before inversion can be carried out, the requirements of this flow chart must be addressed. Click each link on this chart to see a summary of that item.

Given:

1.
1. Field observations
2. Error estimates
3. Ability to forward model
4. Prior knowledge

2. Discretize the Earth

3. Choose a Suitable Misfit Criterion *and* Design Model Objective Function

4. Perform Inversion

Usually adjust misfit and or the objective function

5. Evaluate Results → Iterate

Interpret preferred models

1.1. Field observations must include the measurements, as well as accurate details about sources, receivers, locations, and all data processing; that is, everything that affects the ability to reproduce the data.

For specifics about how surveys work please see "Geophysical surveys" in the "Foundations" chapter.

close

Notes on iteration

Geophysical inversion is an iterative process. Owing to the nonuniqueness of the problem, several equally valid solutions should be obtained. Generally a model obtained from a first successful inversion should be refined by exploring the importance of misfit and by adjusting the model objective function. These adjustments should be made within the context of as much understanding of the problem as possible. Then the preferred model of the earth can be chosen based upon the range of acceptable models, and what is already known about the problem and the geology. In other words, the person doing inversion must work as a member of a team with professionals who have geologic, geotechnical, geochemical, and/or other relevant expertise.

The use of geophysics in general is also usually part of an iterative process. Geophysical information can build on geologic information already obtained, and it can help guide further investigations as the project proceeds from preliminary reconnaissance, through follow-up of anomalies in the field, delineation of subsurface details, and further project development.

Inversion concepts:
Suitable problems for geophysical inversion

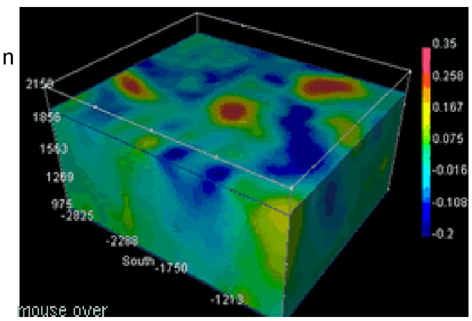
For those making use of geophysical inversion for the first time, it is natural to ask "*will inversion contribute towards my problem?*" This page provides a ten-point outline of criteria to consider when answering this question.

The range of problems requiring inversion

If your geoscience question can be answered **without** knowing the ***values*** and ***distributions*** of physical properties within the ground, then rigorous inversion may not be necessary. One example is an object search question (such as locating underground storage tanks) for which a simple map of a geophysical anomaly might provide a clear indication of where the desired object is located.

Inversion is essentially a processing step that attempts to find the cause for a set of measurements. Therefore inversion can contribute to geoscience problems at any scale. See the **sidebar** for examples of inversion being used at all scales of problems, from studying the structure of a whole planet, down to characterizing features at the scale of only a few cubic meters.

Ten aspects affecting suitability of problems for inversion



As the needs of exploration, engineering, environmental, and other industries become more sophisticated, so too do the requirements for inexpensive, non-invasive acquisition of detailed quantitative information about subsurface materials. In the image to the right, the value of density throughout the volume of interest has been estimated by inversion of ground-based gravity data set, in order to characterize an ore deposit as quantitatively as possible.

The question now is, "*what aspects of a problem affect its suitability for inversion*?" The following ten points below should be considered - click numbers to jump to corresponding details below.

1. Physical property contrast	**2.** Illumination energy	**3.** Problem size	**4.** Consistent data & model type	**5.** Topography
6. Permissible locations of buried features	**7.** Consistency with prior knowledge	**8.** Accurate, clearly understood data	**9.** Well-characterized data errors	**10.** Consistency between discretization & data

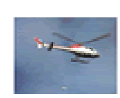
1. **Physical property contrast**: There must be a physical property contrast corresponding to the geological problem. This is true for all geophysical work, and it is true for inversion. If the data contain no response related to the target, inversion will recover nothing.

- **Example**: In the Century Deposit case history (in Chapter 9), the model of electrical conductivity obtained by inversion of DC resistivity data did not show where the ore body was, although other structural information was obtained. However, the chargeability model did include zones of chargeable material corresponding with economic ore. A table of physical property values obtained by drilling confirms that the ore body's electrical conductivity is similar to host rocks, while it's chargeability is significantly different from surrounding geologic materials.

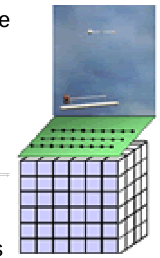
2. Illumination energy: Data should be gathered with source energy interacting with the target in as many different ways as possible. When the source energy cannot be moved, some prior knowledge about how material is likely to be distributed can be incorporated into the inversion. This is done for potential fields data - see chapters on inverting magnetic and gravity data.

- **Example**: One variety of DC resistivity survey (a so-called 'gradient array' survey) involves using only a single location for source electrodes. This type of data is hard to invert successfully and techniques similar to inverting magnetic or gravity data may be necessary. More discussion can be found in the San Nicolas case history of Chapter 9, in section 3,"*Regional scale geophysics*", under "*Chargeability*".

3. Problem size: What is meant by problem size? This issue is covered in detail throughout the CD-ROM, but there are two essential aspects: the number of cells used to discretize the Earth (referred to as *N*), and the number of data values (referred to as *M*). The numerical implementation of inversion schemes will involve working with matrix calculations that are as big as *N x M*.



- **Example**: How serious is this? Imagine a normal airborne survey covering an area 4km by 4km, involving survey lines spaced 100m apart and measurement spacing along the lines of 5m (represented by the lines with dots in the cartoon to the right). For this survey, N = 32,000. If we want the subsurface model to include cells that are 10x10x5m down to a depth of 4km, then our volume includes 400 x 400 x 400 = 64,000,00 cells (represented by the volume of cubes under the survey area in our cartoon). Even for this seemingly reasonable situation, $N \times M$ is too large for normally available computing tools. A compromise will be necessary. The size of each cell must be increased (reducing spacial resolution), and the number of data values can be reduced so there are only a few data points for each cell at the model's surface.



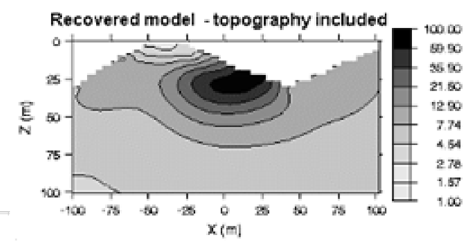
4. Consistent data and model type: Inversion for 1D or 2D models (see the model types summary page in the "Foundations" chapter) can only produce sensible results if the measurements are unaffected by geologic conditions that change in the "missing" direction. In addition, if the data do not contain information about variations in all 3 dimesions, then full 3D inversions are not likely to be successful. In other words, the data set and the choice of inversion methodology must be consistent.

- **Example**: DC resistivity and IP surveys are commonly gathered along survey lines. To cover large areas, several lines may be used. These lines are usually rather far apart compared to the measurement spacing along the lines. Individual 2D inversions are recommended for each data set gathered along a line. Fully 3D inversions using many lines will likely be successfull only if the survey *line spacing* is less than the maximum electrode spacing *along* the lines. An example of the latter situation is given in the Cluny, Mt. Isa case history of chapter 9.

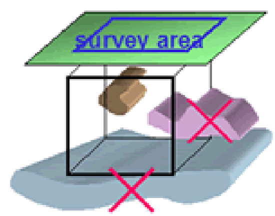
5.Topography: 2D and 3D inversions must have good topography data available. Many types of data are affected by topography, so these affects will be explained by erroneous structrues if topography is not correct in the inversion model.

- **Example**: The 2D synthetic model shown to the right has some variations in electrical conductivity under a mountain and a valley. Synthetic DC resistivity data generated over this model have been inverted with and without proper topography. Click the following buttons to see the two resulting models of the earth's distribution of electrical conductivity:

  ○ True model; ○ first inversion result; ◉ second inversion result.

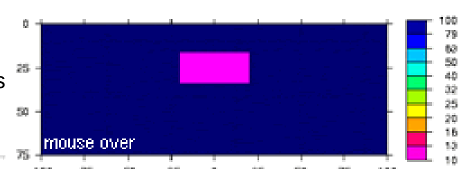


6.Permissible locations of buried features: If there are geologic features affecting the data which do NOT lie within the volume encompassed by the model, then the inversion will be forced to place erroneous features within the model in order to account for those components of the data. ***This is very important, especially for magnetic and gravity surveys.***



7. Consistency with prior knowledge: If a process is designed to generate "smooth" models, you should expect to interpret the recovered models in terms of smooth variations of the physical property. This is not necessarily a problem if the "smooth" models can be interpreted in terms of structures expected. The point here is that interpretations can be effective only when the inversion process being used is properly understood.

- **Example**: The figures to the right show how a discrete block of conductive material may be revealed by inversion using a process that returns smooth models. Move your mouse over the figure to see the inversion result. This issue is much clearer with a good understanding of why inversion procedures do what they do.



8. Accurate, clearly understood data: It should be obvious that inversion results can be only as good as the input data. In addition to having accurate data, it is necessary to know exactly what the physical measurements were, and how the input data were generated from those measurements. Predicted data cannot be produced properly, and therefore a successfull inversion outcome cannot be expected if all the relevant details are consistent with the forward modelling procedure used in the inversion.

- **More details**: For the outline of this point, see Decision number 1: fitting the data, in Chapter 3, " Inversion Concepts".

9. Well characterized data errors: In addition to understanding exactly where the original data come from, there must be an estimate of the errors that are associated with each data value. It is not common for quantitative statistics to be available, so assumptions often must be made about the errors associated with data.

10. Consistency between discretization and data: The size of cells used in the model should be smaller than the size of all features that will affect the measurements. In other words, data must NOT contain information caused by features that are smaller than the cell size.

- **More details**: To meet this demand with magnetic or gravity data, it may be necessary to calculate an equivalent data set that would have been gathered some distance from the surface. The relevant data processing step is called upward continuation.

- **More details**: For other types of data it may be necessary to increase the errors assigned to data if geologic features exist which are very small and/or close to the measurement location.

Inversion concepts:
Inverting 2D DC resistivity data

Contents and objectives

In this section, fundamental concepts are implemented for a synthetic 2D DC resistivity data set. Treatment is not rigorous because the purpose is to illustrate how inversion is applied, and what the effects of various decisions will be on the model recovered by inversion.

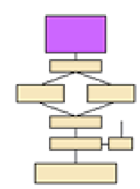
This page is challenging at this level. It assumes some understanding of DC resistivity surveying, and some experience with geophysical data. However, it does illustrate what's involved in carrying out inversion work. If you are reading this from the IAG cd-rom, there is a slightly more advanced version of this page at the end of **Chapter 4 "Inversion Theory".** If Chapter 4 is to be covered, you might gain more by deferring this page until then.

1. Introducing the scenario

To illustrate the work involved in carrying out inversion, we will use a synthetic DC resistivity survey and the UBC-GIF's forward modeling & inversion program library called DCIP2D. The artificial 2D electrical conductivity structure of the Earth that is used to generate the data will be revealed later, but for now we will start with the data set and work through an inversion sequence as if we were doing a normal job.

Generic flow chart icons are included to help remind you of where in the inversion methodology we are currently working. There are occasional sidebars to provide clarifaction without interrupting the flow. Questions to guide your thinking about inversion are included via the questions icon. Click here to see all the questions on one page.

2. DC resistivity data, errors and prior information



First all four pre-requisites in the top box of the flow chart must be considered. They are: 1. Field observations, 2. Error estimates, 3. Ability to forward model and 4. Prior knowledge. Each pre-requisite is described next.

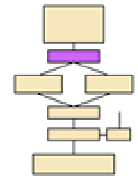
The measurements include source current, resulting potentials, and the geometry of electrodes, but recorded data usually are apparent resistivities. The conventional plotting procedure called pseudosections is not self explanatory, therefore, a sidebar is provided to briefly describe them. There is no other processing of data prior to inversion, except if voltages were not saved during the survey - then they must be derived by working the apparent resistivity formula backwards, assuming a source current of 1 Ampere and the known geometry of electrodes. Also, of course, the input data must be formatted into a text file with the proper format. The input file used in this exercise can be seen here. In this file, columns are not labelled - they are (left to right) transmitter electrode positions along the line (2 columns), receiver electrode positions (2 columns), normalized potential measurement, and an estimate of the standard deviation of that measurement. This last column can be added using the inversion program, as explained in part 5 below.

Errors in the data arise for various reasons including inaccurate locations of electrodes and electrical noise (signals can be in the range of low milli-volts or on the order of microvolts). Also, if the real Earth is very different from being two dimensional near any electrodes, then "geologic errors" occur because the features causing data cannot be recovered perfectly with a strictly two-dimensional (2D) method.

Forward modelling of DC resistivity data based upon a 2D earth (discretized as shown below) is carried out using equations that are explained in section 3 (Principles) of "DC resistivity" in the "Geophysical Surveys" section of Chapter 2 "Foundations".

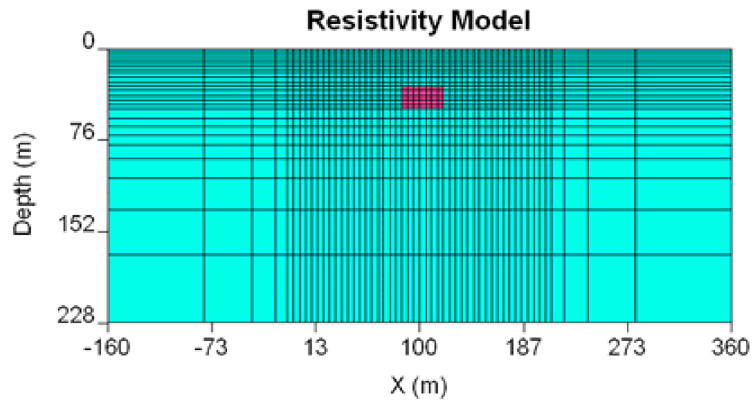
Prior knowledge includes field geology, "educated guesses" about physical property values, expected structures, etc. These will be used in the interpretation. In addition, the methodology makes use of *implied* prior knowledge. For example, the program looks for smooth models that are close to a reference model that can be defined by the user. This reference model is often simply a uniform Earth with fixed electrical conductivity.

3. Discretizing the earth



2D models of the Earth are cross-sectional views of structure and materials directly under the survey line. A very important assumption is that there are no variations perpendicular to the cross-section. For inversions using the UBC-GIF methodology, the cross-section is discretized using rectangular cells.

Smaller cells are used to fill the zone where current has been injected. These determine the spacial resolution of the model. More cells may seem to be better, but with only limited data, the finest feasible resolution will be obtained using cells with widths equal to half the electrode spacing, and depths equal to a quarter of the electrode spacing.



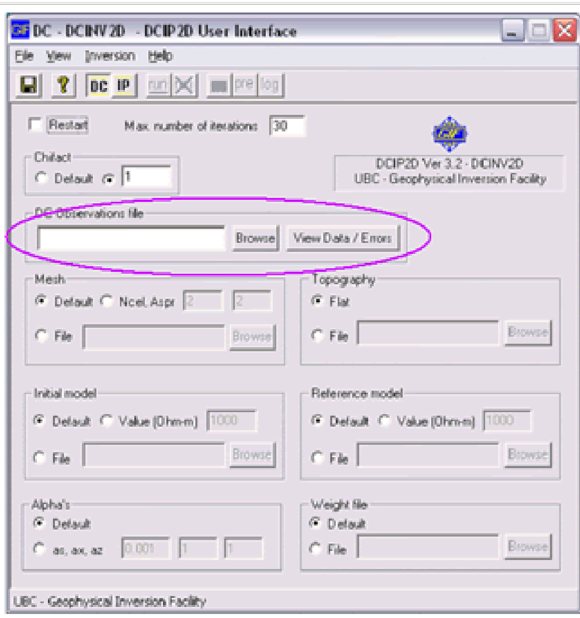
The mesh used to define a discretized 2D model under the survey line. The survey extended from 0m to 200m. Zones outside this region are "Padding", explained in the text. **Mouseover** illustrates.

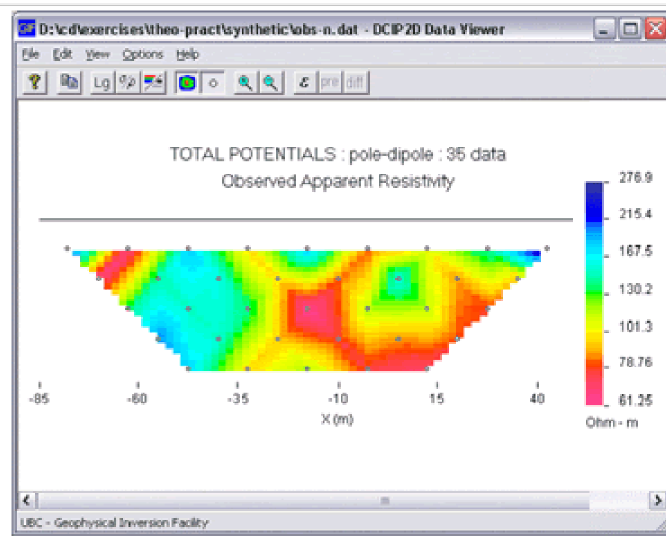
Larger cells are used around the main region of interest to allow for mathematically smooth transitions towards the edges of the domain. Without this buffer (or "padding") zone, it would be impossible to obtain sensible values for the region of interest. Larger cells are acceptible since this zone of the model will not be interpreted.

Topography must be approximated using rectangular cells. When a "default" mesh is used with UBC-GIF codes, the program builds a discretized 2D Earth with cells in the zone under the electrodes that are 0.5*S* in width and 0.25*S* in depth (where *S* is minimum electrode spacing). Outside this region, cells are added with increasing size towards the periphery.

4a. Getting started with DCIP2D - the data

The first task is to inspect the data set itself in order to gain first impressions and to ensure that sensible errors are applied. For some hints on familiarization with the data, click the questions button to the right. A





Above: Data pseudosection for our synthetic data set.

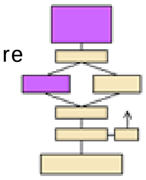
< = = Left: The graphical user interface for DCIP2D.

B

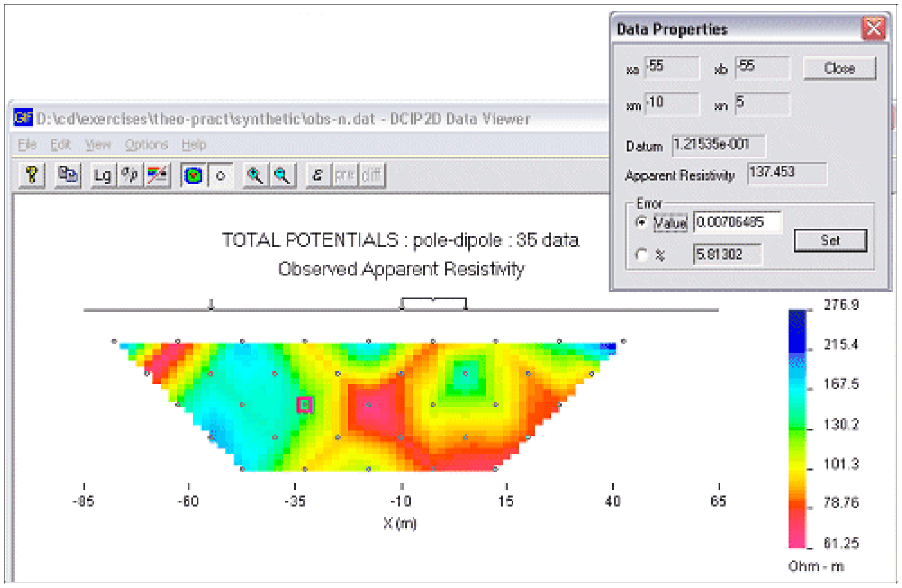
When first examining a DC resistivity data several questions should be considered. The two main objectives are to gain some appreciation for what is likely to be in the ground prior to carrying out the inversion. Click the question button for some suggestions of applicable things to think about.

4b. Misfit criterion: specifying errors

We need to judge the reliability of the data set, and to specify standard deviations for the data accordingly. Making this judgement requires some knowlege about the acquisition of data in the field, and for the likelihood that there were problems or external sources of noise.



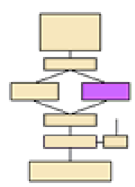
True statistical estimates of error from the field can only be obtained if many versions of the data set were obtain. Since this is almost never done, it is common to assume there are random Gaussian errors, and standard deviations for each datum can be applied using both a percentage of the datum and a minimum value, or offest. For the data set used in our synthetic example, the added noise is a random Gaussian value based upon 5% of each datum plus a 0.001 Volt minimum.

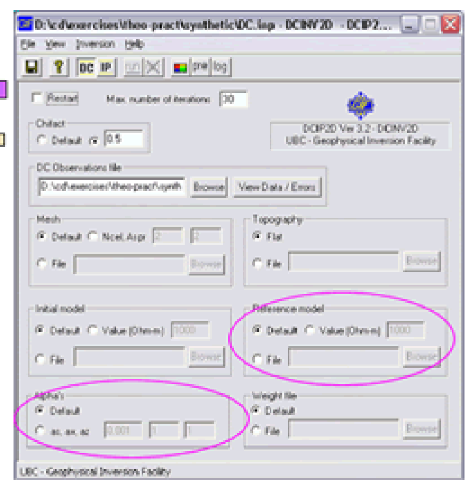


Pseudosection display in DCIP2D, with the properties dialogue for a single datum.

If you have some field experience, some questions you might like to consider regarding errors are in Question Set C. Errors can also be adjusted for individual data points if you suspect any datum is particularly noisy. For example, it is not uncommon for all data values recorded at one electrode location to have additional noise, due for example to a poor electrical contact, a nearby metallic fence, or other reasons. Specifics for every datum can be examinied in the data display program by clicking on any data point (see the previous figure).

5. Specifying the desired optimal model





A so-called *model objective function* is used to define the type of "optimum" model the inversion algorithm is looking for. This function is a way of quantifying desirable features of a physical property model. The inversion chooses an optimal model by searching for a model which will minimize this function *subject to the constraint that the chosen model can generate predicted data that satisfy the misfit criteria*. The model objective function is

$$\phi_m = \alpha_s \int (m - m_0)^2 dx + \alpha_x \int \left(\frac{d}{dx}(m - m_0) \right)^2 dx + \alpha_z \int \left(\frac{d}{dz}(m - m_0) \right)^2 dz,$$

and it is defined using two components coloured blue and pink above:

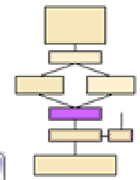
1. The algorithm will try to find a model that is as close as possible to a reference model defined either as a half space (by default a halfspace with a resistivity equal to a weighted average of measured apparent resistivities), or as some other, more complicated model defined by the user (if there is enough prior knowledge).
2. The model will be as smooth as possible in the X and Z directions.

In fact, the significance of each component is controlled using the "Alpha" coefficients α_s, α_x, and α_z in the equation above. Therefore the user can request a model that emphasizes either component 1 or component 2.

Default values of these coefficients are determined by the program based upon the length scales of the survey and mesh. The inversion's task is to find a model that minimizes this relation; the result will be an optimal model.

For the program DCIP2D, the default specifications for these "Alpha" parameters have been found to work well as a first attempt, but experimentation and adjustment of the parameters defining the desired model type is expected during the course of inversion processing. This will be discussed in the "Alternative models" section below.

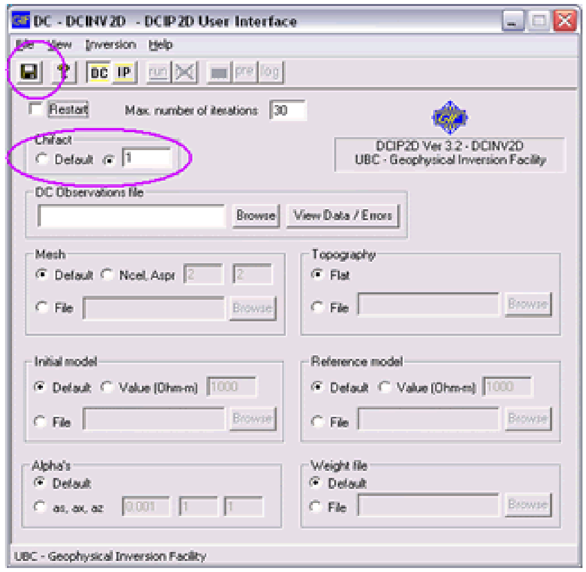
6a. First inversion



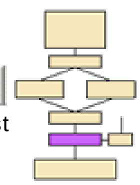
The first inversion should be run only after learning as much as possible from the raw data, including how to set errors properly. The DCIP2D user interface has defaults for all parameters except the input file name.

If our synthetic data set has errors assigned using 5% + 0.001V minimum, then running the first inversion with all default parameters will produce a reasonable initial solution.

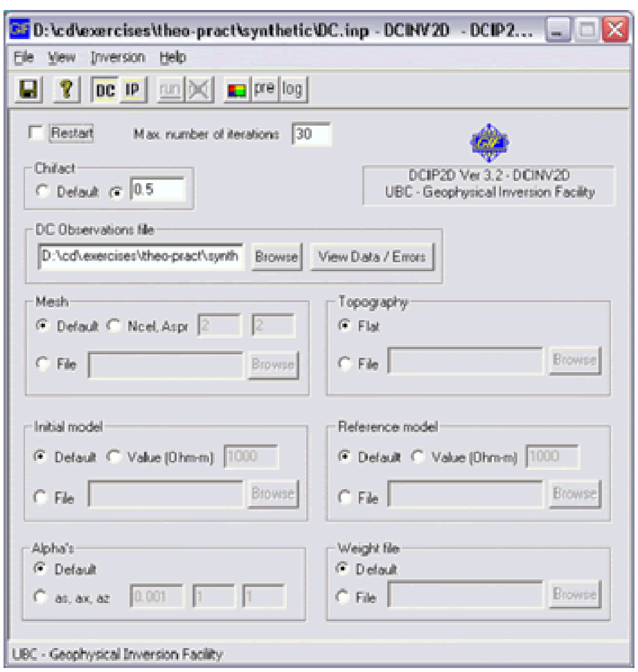
A sidebar shows predicted data and models as the inversion progresses.

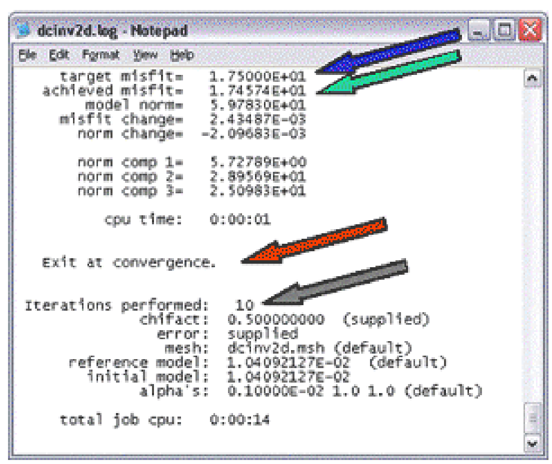


6b. Evaluating results



The GUI after completing the inversion is shown in the next figure. Point your mouse to either model, pre, or log buttons to display an image of the corresponding user-interface window for the inversion that was just completed.





See text for notes on images obtained by pointing to either of the model, pre, log buttons on the image to the right.

Reset to blank

How should you examine these results to determine if the inversion was successful at returning a reasonable model? There are **five** aspects to observe and consider:

1. **A log file** is produced containing information about how the inversion progressed. For a quick assessment of whether your inversion run proceeded normally, point to the log button and look at the last few lines of the resulting text file. Here are some questions to consider when using the log file. Arrows in the figure correspond to colours in this list:
   - Did the inversion end with “convergence?”
   - What value of target misfit was specified ("target misfit" is the "desired" value of a comparison between predicted and measured data. It is specified via the *Chifact* parameter - explained further below.)?

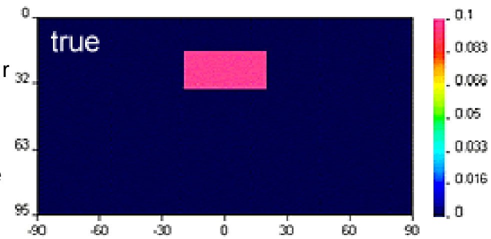
- Was this target misfit achieved?
- How many iterations were performed ("Iterations" means cycles through the inversion flowchart.)?

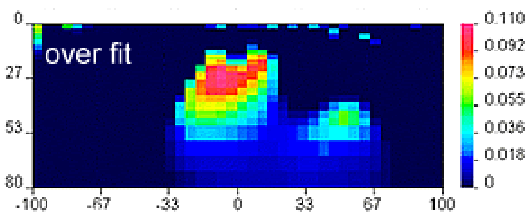
2. Next, **predicted data** should be compared to observations. Pointing to the  button displays pseudosections of the observations and the predicted data. They should look nearly identical. To see variations more closely, click the "diff" button in the data viewing window. This changes the second pseudosection to a "misfit map," which shows the differences between the two data sets. This is what is shown above. The misfit map should look random, with maximum values of some small percentage of the measured data (based upon noise specifications).

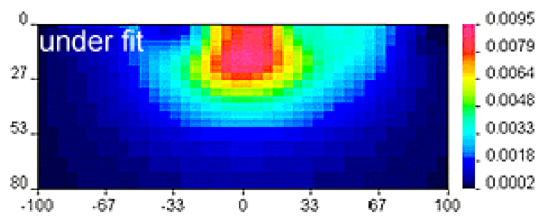
3. Now look at the **resulting model**. The  button brings up the model viewing window with the complete model displayed, including all the padding cells. Select "Padding cells" in the "Options" menu in this window to specify how many padding cells to drop from the display. The image above has 3 cells dropped from either side, and 4 cells dropped from the bottom. You can also adjust the minimum / maximum values for the colour scale - necessary for comparing various models.

4. The **progress of the inversion** during it's iterations should also be checked. In the model viewing window, the algorithm's progress can be displayed graphically by selecting the "Curves" toolbar button in the "View" menu. The resulting graph shows how the values of misfit and model norm varied at each iteration. ("Model norm" is the value of the *model objective function* - this is what we are trying to "minimize". Further explanation is below.) The algorithm is programmed to add structure gradually in order to find a model that explains the data - i.e. it works on *reducing* the misfit value (**blue** curve) until the target misfit is reached. Then it must try to minimize the model norm without changing misfit. Thus, you should see a slight drop in the model norm value (**red** curve) until no more adjustments can be made to improve the situation.

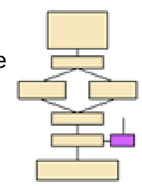
5. **Geologically reasonable?** It is important to decide whether the resulting model is geologically reasonable. This final consideration is more subjective. A simple example is shown here, in which data produced by calculating data over the "true" 2D model (top right) are inverted twice to produce two inversion results which are both inadequate. The image labelled "underfit" is a model recovered when the target misfit was too large. The program has stopped looking for details when predictions look only somewhat like observations. The image labelled "overfit" is a model recovered when the program has tried too hard to find details that explain every nuance in the observations.







7. Second inversion



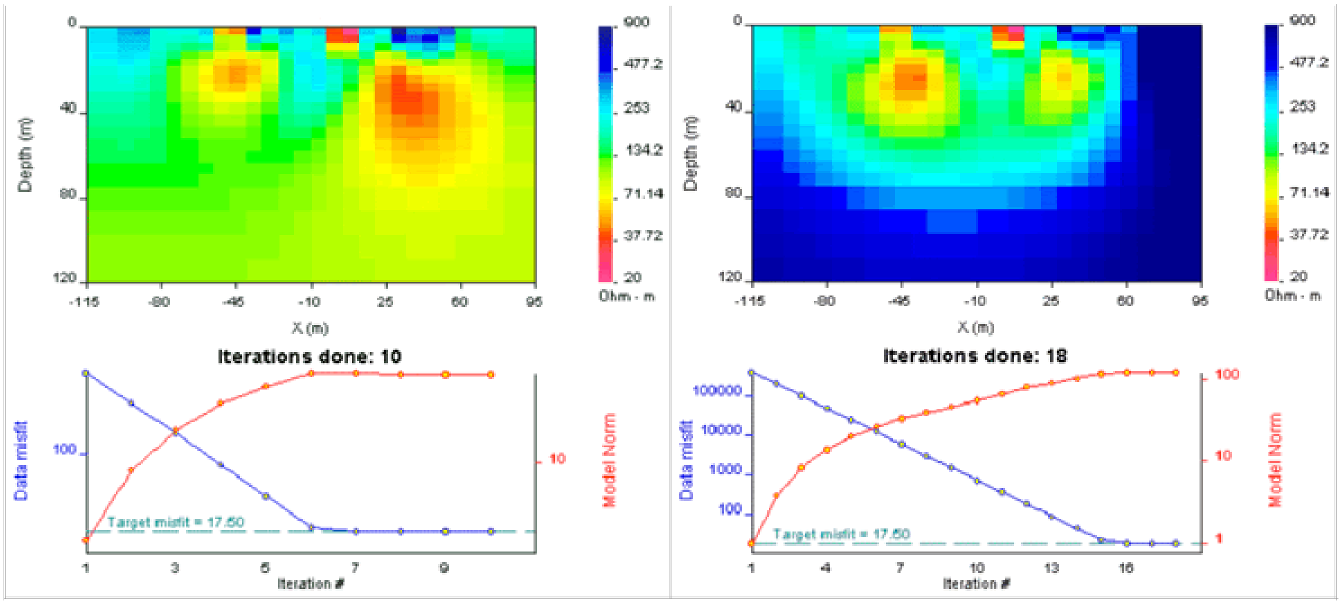
Your first inversion rarely produces an optimal result. More inversions must be run to obtain alternate models. Alternate models are obtained by changing either the model norm or the target misfit (both terms were defined in the previous section). Changing the model norm can be done in two ways: a "reference model" can be adjusted, or the degree of smoothness can be adjusted. See the last part of "Feasible model norms" in the "Norms and misfit" section of this chapter.

First you should identify what values the program set for its default run (see question set D above), and then you should adjust one of these to obtain a second result.

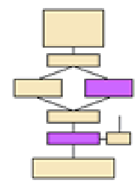
Here we will start by specifying a different reference model. One with a value similar to the first model's lowest conductivities is a good choice.

Upon completion there will be a second model can be compared with the first, default model. These are both shown in the figure below. Typical questions to consider are given in question set E.

Resistivity model using default reference model.	Resistivity model using reference model of 1000 Ohm-m.



8. Depth of investigation (special for 2D DC/IP inversions)

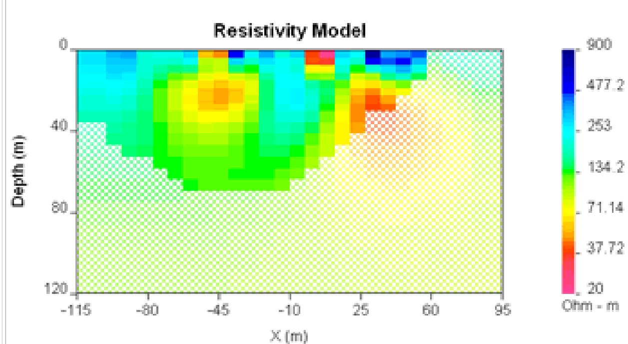


Models produced by inversion of DC resistivity data appear to fade with depth, but how can we gain a more quantitative understanding about which regions of the model are reliably constrained by our measurements? If there are at least two reasonable models obtained using different reference models, the two models can be compared to identify which regions of the model most significantly affect the measurements. Results of doing this are explained next.

Using the DCIP2D program, the method is applied within the model viewing window, using "Depth of investigation" option in the "Options" menu. There must be a second model that was recovered using the same mesh as the one being observed. Results for our synthetic model can be seen as soon as the two inversions described above have been done. Any two different inversions results can be used. Four versions are shown here, and typical questions that should be considered with such results are given in question set F.

- Default model (chifact = 0.5), depth of investigation (DOI) via a model that used mref=0.001S/m.
- Same model, DOI via a model that used mref=0.02S/m.
- Model using mref=0.001S/m, DOI via a model that used mref=0.02S/m.
- Model using mref=0.02S/m, DOI via a model that used mref=0.001S/m.

Evidently, all versions will be interpreted similarly. The left buried conductor appears to be small with more resistive material beneath approximately 40m depth. However the stronger buried conductor to the right does not have a bottom that has been imaged by this survey. DC current tends to accumulate in conducting regions so investigation depths in conductive ground tend to be shallower.



9. Alternative models

When geophysical models created by this inversion methodology are used to make geological interpretations, it is a crucial that more than one model should be used. The reason is that with any single model, many questions can be posed. Only by comparing several results can you begin to gain reliable answers to these questions. For example:

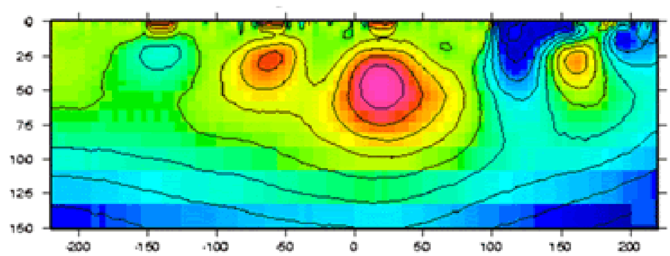
- Which of the features in the model actualy emulate the earth?
- To which regions of the earth are data most sensitive?
- How deep are we seeing?
- Are anomalous regions in the model open or closed at depth?
- Which features are in fact due to the parameters in the inversion process rather than to the measurements?

There are three main approaches to finding alternate models, and these are outlined next.

9a. Adjusting the reference model

Our second inversion done above involved specifying an alternative reference model. This is a powerful tool for exploring which regions of the model are reliably defined by the data. This point is emphasized in the interactive figure below, which shows five models recovered for a new, synthetic model, which is somewhat more complex than the small example we have been using. All recovered models (2. through 6.) are equally valid because they can each reproduce the data to within the specified misfit, but they are different because of the specific reference model used for each.

1. ○ True model
2. ○ m_{ref} = 40 Ohm-m
3. ○ m_{ref} = 400 Ohm-m
4. ○ m_{ref} = 4,000 Ohm-m
5. ◉ m_{ref} = gradient from 400-400,000 Ohm-m
6. ○ m_{ref} = gradient from 400-0.04 Ohm-m



9b. Adjusting the model norm

Recall that the way in which a model is "measured" in order to distinguish between different models involves an equation with two parts. As noted above, this equation has one part which measures how close the resulting model (m) is to a reference (m_0) (blue part), and a second part that measures how smooth the model is (pink part). In fact, the smoothness in the horizontal and vertical directions are measured separately so the second part of our norm contains two components.

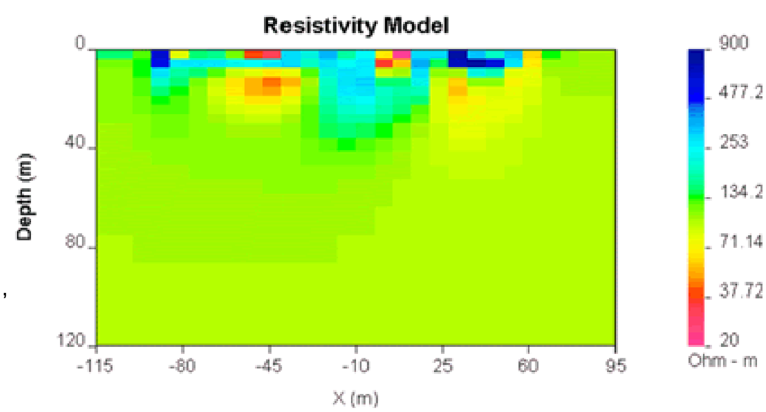
$$\phi_m = \alpha_s \int (m - m_0)^2 dx + \alpha_x \int \left(\frac{d}{dx}(m - m_0) \right)^2 dx + \alpha_z \int \left(\frac{d}{dz}(m - m_0) \right)^2 dz$$

The UBC-GIF inversion routines make the reference model, m_0, and the three α's available for adjustment. What is the effect of adjusting these α coefficients? Recall that this function will be minimized. If one of the α's is very small, then the corresponding term will contribute little to this minimization. So, reducing the size of αs will result in smoothness in the X- and Z-directions in favour of closeness to the reference. Reducing the size of αx or αz will result in models that are preferentially smooth in the X- or Z-directions respectively. Images below illustrate the effect of different choices for values of the three α's .

Click buttons to see models resulting when the three α parameters are modified as follows:

○ (αs, αx, αz) = (.001, 1, 1)
◉ (αs, αx, αz) = (1, .001, .001)
○ (αs, αx, αz) = (.001, 1, .001)
○ (αs, αx, αz) = (.001, .001, 1)

Are all models equally capable of reproducing the data? Yes, predicted data for all results are very similar. You can try this yourself, and consider the questions in question set G.



What about the actual values to specify for the α parameters? This is beyond the scope of this page. G

9c. Adjusting misfit

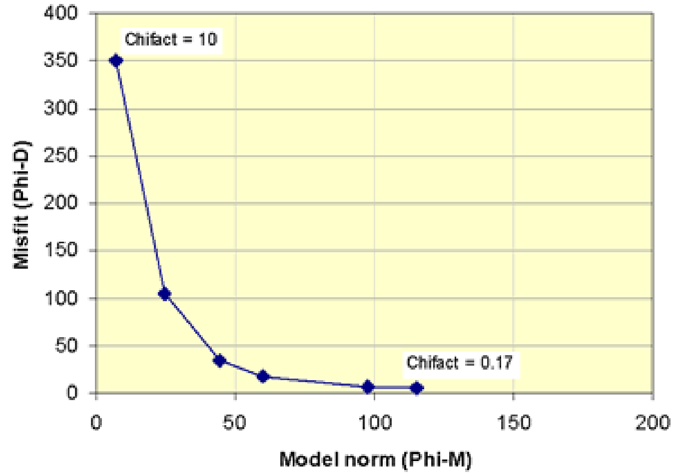
Finally, what is the effect of adjusting the target misfit value to numbers other than N, the number of data? An initial comment was made above, but you can explore the consequences of adjusting the misfit by changing the *Chifact* parameter.

In fact, it is instructive to generate models for a range of values, and use the resulting values of model norm and misfit to generate curve which plots these values for several trials.

The results of inverting our example using several different values for *Chifact* are summarized in the following table and corresponding curve. Some questions to consider are in question set H. H

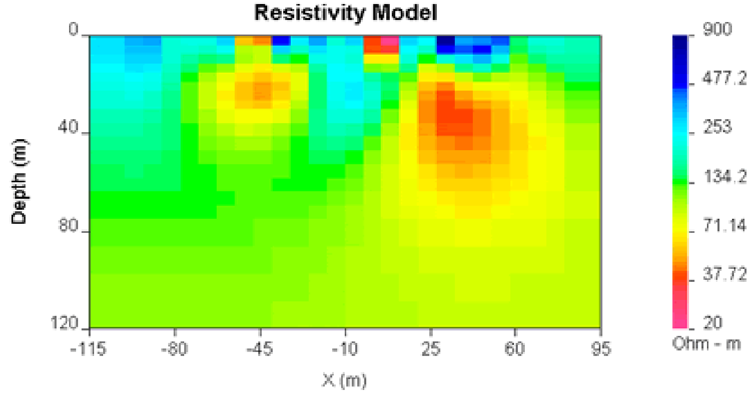
Table of values resulting from inversion using five values of *Chifact*.
phiM and phiD are defined in the graph's axis.

chifact	phiM	phiD	iterations
0.17	115.1	5.9	28
0.2	97.6	7.0	19
0.5	59.8	17.5	10
1	44.5	34.9	8
3	24.4	105.0	6
10	7.3	350.3	4



Click to see models recovered using each value of *Chifact*.

- Chifact =0.17
- Chifact =0.2
- Chifact = 0.5 (Model recovered earlier)
- Chifact =1
- Chifact =3
- Chifact =10



10. Conclusions

After working through this page, some relevant closing remarks are:

- Non-uniqueness can be quite severe. By this we mean that we have seen many models that can be interpreted, and all are equally capable of generating the measured data set. We know that no single answer is possible because we can never sample the earth well enough (i.e. collect enough data) to render the problem over-determined, unless the earth is modeled with only a few parameters (i.e. cells).
- Some models are not geologically reasonable. This is subjective, but over-fit and under-fit results should be detectible by assessing the smoothness or roughness of the result. And, not surprisingly, experience makes this easier.
- Finding alternate models is facilitated by altering either
  - the reference model, or
  - the alphas, or
  - the target misfit value.
- Interpreting geologic or geotechnical properties should never proceed without several models to work from. It is crucial that you be able to defend a model in terms of (i) zones that are reliable because their features depend upon measurements, and (ii) zones that should not be used in the interpretation because they depend upon the methodology rather than on field measurements.