Application of sparse layer inversion and compressive sensing for thin reservoir delineation using seismic data

Manisha Panda

Guide: Prof. Aurobinda Routray,

Professor,

Electrical Engineering (P.I.)

Mentor: Sanket Smarak Panda



Indian Institute of Technology, Kharagpur

**CERTIFICATE**

This is to certify that Mr. Manisha Panda has worked with me during the period 15-5-2019 to 30-06-2019. The student has worked in a project/research topic entitled “Application of sparse layer inversion and compressive sensing for thin reservoir delineation using seismic data” during the aforesaid period.

I am favorably impressed with the student's diligence, obedience and sincerity towards work assigned to the said student.

The undersigned may be contacted for further information about the intern.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

(Signature of Faculty)

Name of faculty: Aurobinda Routray

Designation: Professor

Department: Electrical Engineering

**Abstract**

The present project aims to develop an intuitive & GUI based software for sparse layer inversion (SLI) both in MATLAB for seismic exploration. Novel techniques based on SLI and compressed sensing will be developed. Thus, the proposed software will assist geoscientists to delineate thin seismic layers having width below the tuning thickness which are otherwise not visible in conventional seismic volumes thereby improving detectability and resolution.

This project addresses research issues viz. information retrieval, sampling theory, and optimization techniques. Our research team has the expertise in developing information theoretic framework for reservoir characterization this expert team with new researchers will be associated in the proposed project to develop and test the software with real field data in collaboration with GEOPIC.

Several SLI algorithms will be implemented and compared visually and through appropriate metrics. This will immensely help the user to choose the optimal one with better performance for a given dataset.

**Keywords:** Spectral inversion, sparse layer inversion (SLI), sparse spike inversion (SSI), tuning thickness, matching pursuit, basis pursuit, cost function, optimization, software packages, compressive (compressed) sensing (CS), signal to noise ratio (SNR)

**CONTENTS**

* Introduction
* Literature
* Methodology and results
* Conclusion
* References

INTRODUCTION

The basic aim of this project is to apply several algorithms in order to look at the levels of earth layers in a way bearing algorithms that ultimately lead to higher resolution of the surfaces. . The prime objective using different wedge dictionaries is to filter out the major layer responses and reconstruct the Seismic data using only thin layer responses only. Our objective is to construct the Seismic signal by elimination of the thick layer. This a new approach to the thin layer detection. In this approach, Basis Pursuit is first applied on an overall wedge dictionary to detect every layer response. A software based on GUI model is made for sparse layer Inversion technique both in MATLAB and PYTHON.

The following report deals with the MATLAB part only.

LITERATURE

* Layers of Earth

Geophysical inversion involves mapping the physical structure and properties of the subsurface of the earth using measurements made on the surface of the earth. We can also say that the process of inversion is the way of recreating the earth model using seismic data as input. It can be said as invert process of the Forward model.

* What is Thin Layer?

The layers which have thickness less than one-fourth of the wavelength of the seismic wavelet can be said as thin layers. The Seismic Layer Inversion technique mainly aims to detect these thin layers and so increase the resolution of the earth’s layers.

* Algorithms and geologic review:

The SLI aims to improve detectability by improving the resolution of earth’s layers.

The convolutional model-

The most basic and commonly used one-Dimensional model for the seismic trace is referred to as the convolutional model, which states that the seismic trace is simply the convolution of the earth's reflectivity with a seismic source function with the addition of a noise component. In equation form,

Where \* implies convolution,

**s(t) : w(t) \* r(t) + n(t)s**

Where

s(t) = the seismic trace,

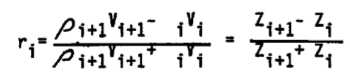
w(t) : a seismic wavelet,

r(t) : earth reflectivity,

n(t) : additive noise.

In Seismic processing we deal with sampled data. If we consider the reflection coefficients to be present at each time sample and the wavelet to be a smooth function of time then the reflectivity can be said to be superposition of scaled wavelets according to respective reflection coefficients.

Each reflection coefficient can be thought as response to an acoustic impedance change in the earth surfaces. This acoustic impedance is basically defined as the product of compressional velocity and density. So the formula for reflection coefficient is the ratio of difference of the acoustic impedance to sum of the acoustic impedance of two adjacent layers of earth. It therefore deals with the boundary between two layers.

The formula goes as such -

r - reflection coefficient,

V - compressional velocity,

Z -- acoustic impedance,

Layer i overlies Layer i+1.

We use to see the seismic impedance and thus reflectivity by deriving from well log. So we create an impedance curve by multiplying the sonic log with density log. If the density log is not available then the impedance can be approximated to the sonic log.

We consider the reflectivity to be a truly random sequence but we can see that if we consider such consideration, apart from having an autocorrelation of 1 at zero lag we do also have some noise components for the time stretch. So we cannot approximate the reflectivity to a random sequence. We can approximate it as a Bernoulli- Gaussian sequence. The Bernoulli part deals with the sparseness in the spike positions and the Gaussian part deals with the amplitude part of the spikes. So for this sequence a lambda term is introduced which controls the sparsity as if lambda is zero then there are no spikes and if lambda is one then the system is perfectly Gaussian.

Zero phase Ricker Wavelets are the perfectly needed characteristic for a wavelet here. This is because the entire energy is concentrated at the center of the wavelet and convolution of this wavelet with the reflection coefficient will better resolve that reflection.

Nguyen and Castagna used Matching pursuit decomposition to decompose a seismic trace into a superposition of reflectivity patterns observed from well data. Matching pursuit decomposition involves the following steps: i) correlates a wavelet dictionary against a seismogram and finds the location and amplitude for the best fit wavelet, ii) subtracts the best fit wavelet and renews the table, iii) repeats the process on the residual trace.

Basis Pursuit Decomposition has many advantages over Matching Pursuit Decomposition. BPD finds a global solution whereas MPD is path dependent solution. MPD iteratively solves the problem but a single change in the order the wavelets are subtracted can result in an entirely different solution. To avoid these problems we consider the use of Basis Pursuit Inversion (BPI) to invert a seismic trace for a reflectivity series which can be superposed or added to give the seismic impedance model. So here we the reflection coefficient dictionary which is wedge shaped convolved with the seismic wavelet. Any pair of reflection coefficient can be said to be sum of odd and even reflection coefficient pairs. The use of this dictionary is important for the sparse layer inversion, and can be said as such inversion involves resolving thin layers.

Sparsity, what and why?

Sparse approximation deals with the sparse solutions of system of linear equations. Let us assume a system of linear equation x = Dα where D is the dictionary matrix of dimensions specified as m x p. x is the signal of interest of dimension m x 1, so the main problem is to find the sparsest possible α satisfying the above equation. But as D has underdetermined nature, the above system of equation has infinitely many solutions so among these solutions we want to get the ones with fewest non zero elements. Mathematically putting it:

min|| α ||0 in the condition satisfying the system of equation mentioned

So it is a pseudo-norm, which counts the number of nonzero elements. But this problem is a NP Hard problem. (P-NP problems) so currently being unsolved we just approximate for the solution of this. { NP hard problems are those subsets of the P-NP problems which specifies that if a problem A is NP hard when any problem B in the NP set can be reduced in the polynomial time of A.}

**Important Algorithms:**

* **Matching Pursuit:**

The basic idea behind MP algorithm is to approximate a signal as a linear expansion of finite number of waveforms chosen from over-complete redundant dictionary. The waveforms are so selected that a best match is found for the signal structure. This sparse approximation algorithm finds “best matching” projections of multidimensional data onto the span of the dictionary.

Suppose is a signal from Hilbert space. Dictionary contains many functions (called atoms). Signal can be approximated using atoms as

(t)

where is the weight factor for the atom . MP does not choose every atom in. Instead, only one atom is chosen at a time for which the approximation error is minimum. The objective is achieved using a series of steps shown in the table below:

|  |
| --- |
| **Matching Pursuit algorithm** |
| Suppose the initial residual is   1. Initialize the residual. 2. Select the atom  from the dictionary that maximizes. 3. Find the residual as. 4. Repeat steps 2 and 3 as |

It is desirable that the residual quickly converges to zero so that only a few atoms sufficiently approximate. The sparsity problem approximately solved by MP can be defined as

|  |
| --- |
| s.t. |

* **Orthogonal Matching pursuit:**

The basic idea behind MP algorithm is to approximate a signal as a linear expansion of finite number of waveforms chosen from over-complete redundant dictionary. The waveforms are so selected that a best match is found for the signal structure. This sparse approximation algorithm finds “best matching” projections of multidimensional data onto the span of the dictionary.

While the residual in MP algorithm may not be orthogonal to the span of the first atom selected, the OMP algorithm ensures that the residual is always orthogonal to the span of the atoms already selected from the dictionary. Each iteration in OMP gives rise to only one selected atom and orthogonalization step ensures that one atom is selected only once. OMP is expected to improve performance over basic MP since OMP reduces error using an orthogonal projection. The steps followed by OMP is shown in Table 1.

|  |
| --- |
| **OMP algorithm** |
| 1. Initialize the residual. 2. Select the atom  from the dictionary that maximizes. 3. Find the residual as. 4. Apply the orthogonal projection operator to the residual. 5. Repeat steps 2 and 3 as |

OMP Algorithm

*G* is the wedge dictionary matrix, that is used in the above table, of size q e, *m* of size e 1, *y* be of size *q* 1 and *z* be of size *e* 1.

* **Basis Pursuit Inversion:**

For such a sparse decomposition we urge to find the minimum number of columns that can be used from D, known as atoms, which can be combined to produce α. So by suitable relaxation (or can be said as approximation) we can solve the above problem by using a L1 norm instead using the L0 norm. So this is what the **Basis Pursuit algorithm** does.

Thus the inverse problem for having the reflectivity series do have mainly two tasks to do – Location of the reflection coefficients and estimation of their amplitudes. So again looking the convolutional model we see **s = Wr + n,** where s is the column vector representing the seismogram, r is the reflectivity series column vector, W represents diagonal wavelet kernel matrix, and finally n is the noise vector.. So the convolution operation has the form: **d = Gm + n,** where d is the data vector, G is the kernel, m is the modeled parameters and n is the noise.

The Basis Pursuit algorithm solves for parameters by simultaneously minimizing both the L2 norm and L1 norm of the solution:  **min [|| d – Gm ||2 + λ|| m ||1]**

Basis Pursuit also uses dipole decomposition to represent the reflectivity series. The dipole refers to a sum of even and odd pair multiplied by a scalar.

Suppose the impulse functions and represent the top and base of a layer. The terms and are two reflection coefficients, is sample rate and, is time thickness of the thin bed. The reflector pair and can be decomposed into an even pair and an odd pair as mentioned earlier,

|  |  |
| --- | --- |
|  |  |

Where and. The coefficients and vary from +1 and -1 times a scale factor. As the layer thickness is usually unknown, is varied from to to include all possible layer thicknesses. Then will represent maximum possible layer thickness. This forms the *wedge dictionary* comprising the odd and even wedges in the form of collection of dipole reflectors with an increasing time separation.

Suppose is the number of samples in the seismic trace. The reflector kernel matrix for the even wedge reflectivity pair is constructed by shifting the reflectivity pair along the time axis with. So each even wedge reflectivity can be written as

|  |  |
| --- | --- |
|  |  |

Similarly, the odd wedge reflectivity can be written as

|  |  |
| --- | --- |
|  |  |

A reflectivity series can be expressed in terms of even and odd wedge reflectivity patterns as shown in Eq. (10)

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |

BPI is used to calculate coefficients and to obtain the inverted reflectivity series.

**Data Description**

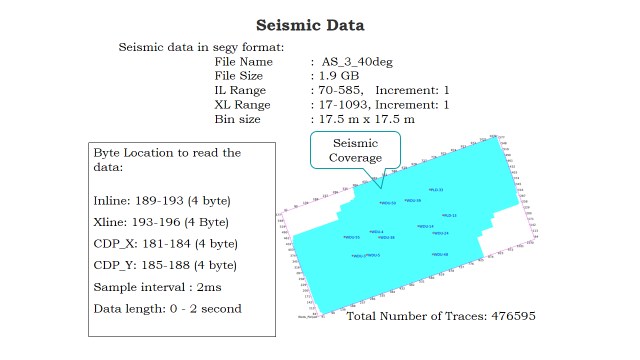
We implemented OMP algorithms on Synthetic and real data to assess their relative ability to resolve thin beds and reveal fine stratigraphic features. The two types of data are described below:

* **Synthetic Data**

To test the algorithms, a set of artificial data is synthesized by choosing a random reflectivity series of arbitrary length (501 in our method). It is designed to be sparse with non-zero values at only a few random locations (40 in the algorithm). The non-zero values are limited within -1 to +1. Basically the choice of random reflectivity series is 92% sparse. A Ricker wavelet of 40 Hz frequency is defined and convolved with the reflectivity series to obtain the synthesized seismic data. We choose an upper limit and lower limit (generally +1 and -1 respectively) randomly to create the synthetic data. We create a matrix of zeroes of size*.* Then 40 locations are randomly chosen and random values were assigned in the matrix of zeroes to form the Synthetic Data

* **Real Data**

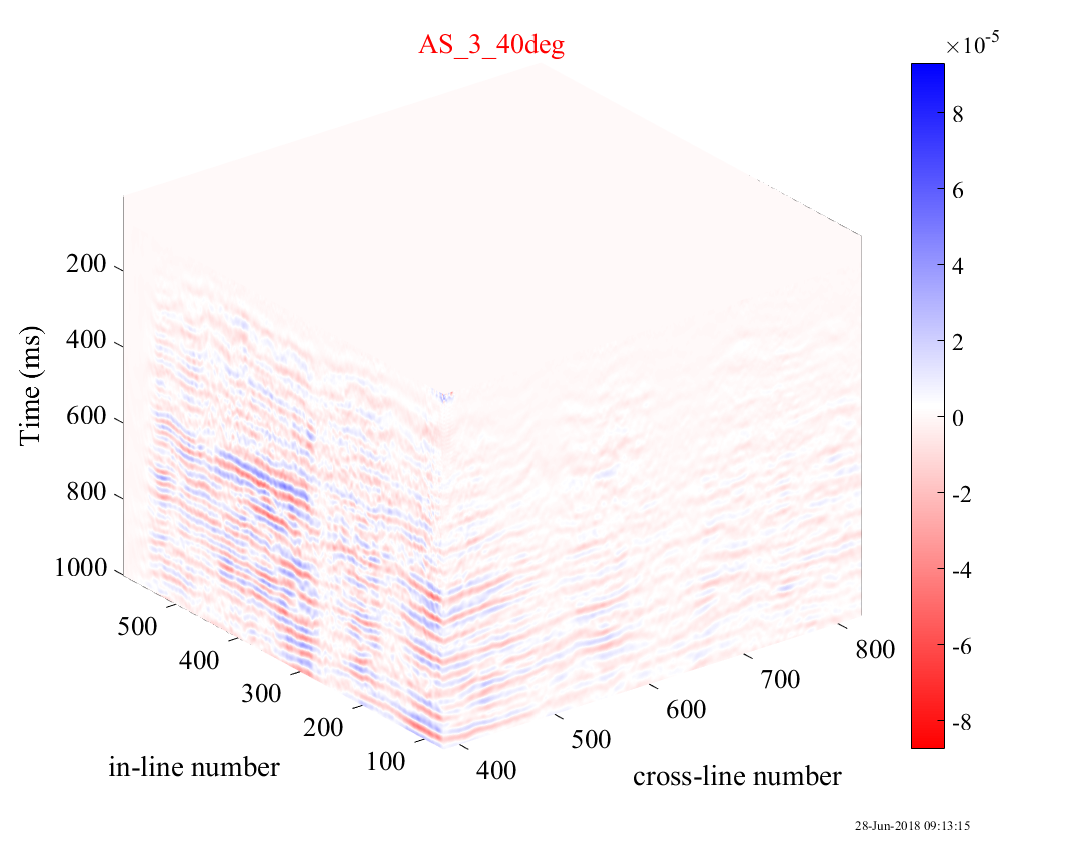
The seismic data *AS\_3\_40deg.sgy* and the well data are provided by GEOPIC, ONGC. The dataset covers a surface are of 150 km2 surface area with 12 wells. Fig. 1 shows the seismic coverage with in-line ranging from 70-585 and cross-line ranging from 17-1093. The sampling interval of the seismic data is 2ms and the time values vary between 0 to 2seconds yielding 1001 samples.



.

Area of interest and seismic data format

The dataset ‘AS\_3\_40deg.sgy’ is a 3D volume data and hence can be shown as a cuboid consisting of inline 70-585 and crossline 384-824 where no trace is missing. The time interval is from 0 to 2000ms with sampling time of 2ms. The volume plot is as follows:



3-D view of seismic cube

Methodology and Results

**Implementation of Matching Pursuit on synthetic data**

**Single Trace:**

The results of SLI implementation using MP algorithm are shown in figure below. The atoms in the dictionary are generated using the 40 Hz Ricker wavelet. Synthetic signal is decomposed into these atoms in accordance with the maximum correlation factor.



Application of matching pursuit on synthetic data of length 501 with wedge dictionary convolved with 40Hz ricker wavelet

**Implementation of Matching Pursuit on real data**

Results on real data are shown for a slice through cross-line 712 covering in-lines shown in Figure (a) and (b) a cube comprising in-lines 70 to 585 and cross-lines 384 to 825 shown figure below; for the said range of in-lines and cross-lines there are no missing traces



SLI results on seismic section at cross-line 712 of real data using MP algorithms (a) Original slice (b) Basic MP

**Implementation of OMP on synthetic data**

**Single Trace:**

The results of SLI implementation using OMP algorithms are shown in Fig. 3. The atoms in the dictionary are generated using the 40 Hz Ricker wavelet. Fig 3(a) shows generated synthetic sparse signal of length 500 and the recovered sparse signal from the algorithm. Fig 3(b) shows the seismic signal obtained by convolving the sparse reflectivity series with 40Hz ricker wavelet along with the signal reconstructed from the wedge dictionary obtained by the 40Hz ricker wavelet.



(a) Implementation of OMP on single trace synthetic data (sparse reflectivity)



(b) Implementation of OMP on single trace synthetic data (Seismic Traces)

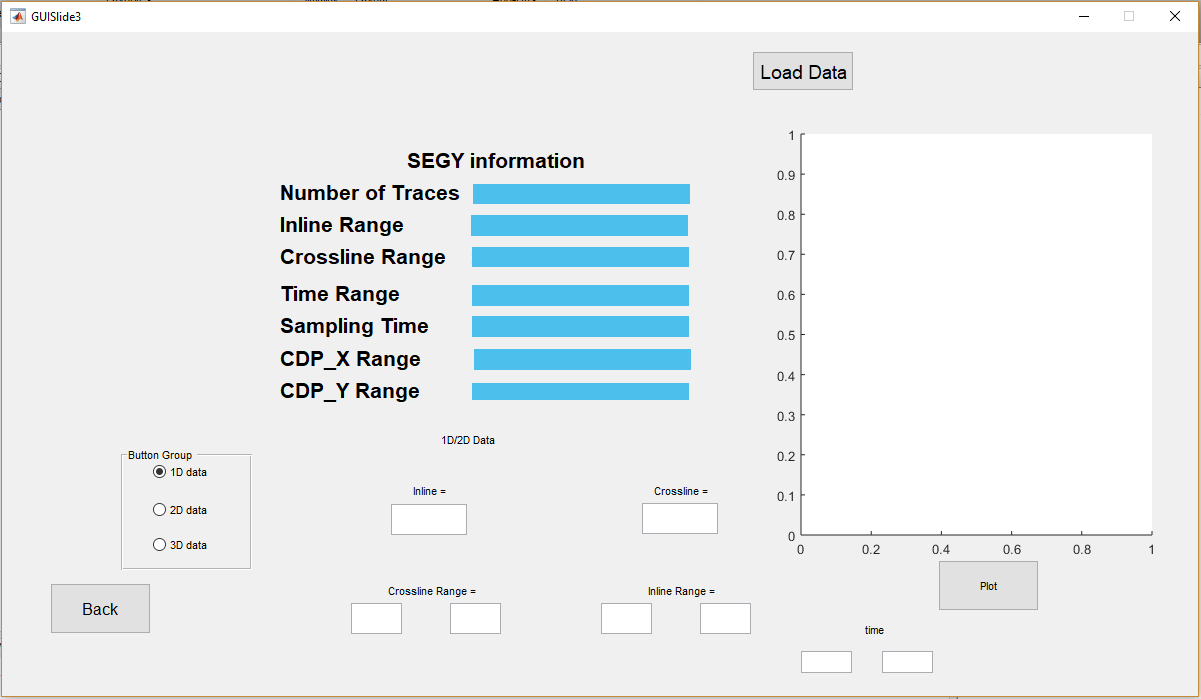
**Implementation of OMP on real data**

The seismic data, time-information, in-lines, and cross-lines are extracted from the segy file *AS\_3\_40deg.sgy* provided to us by GEOPIC and saved using SeisLab, a MATLAB Package for the analysis of Seismic data. Results on real data are shown for a slice through inline 321 covering crosslines 200-400 as shown in figure.



Implementation of OMP on real data (inline 321 and crossline 200-400)

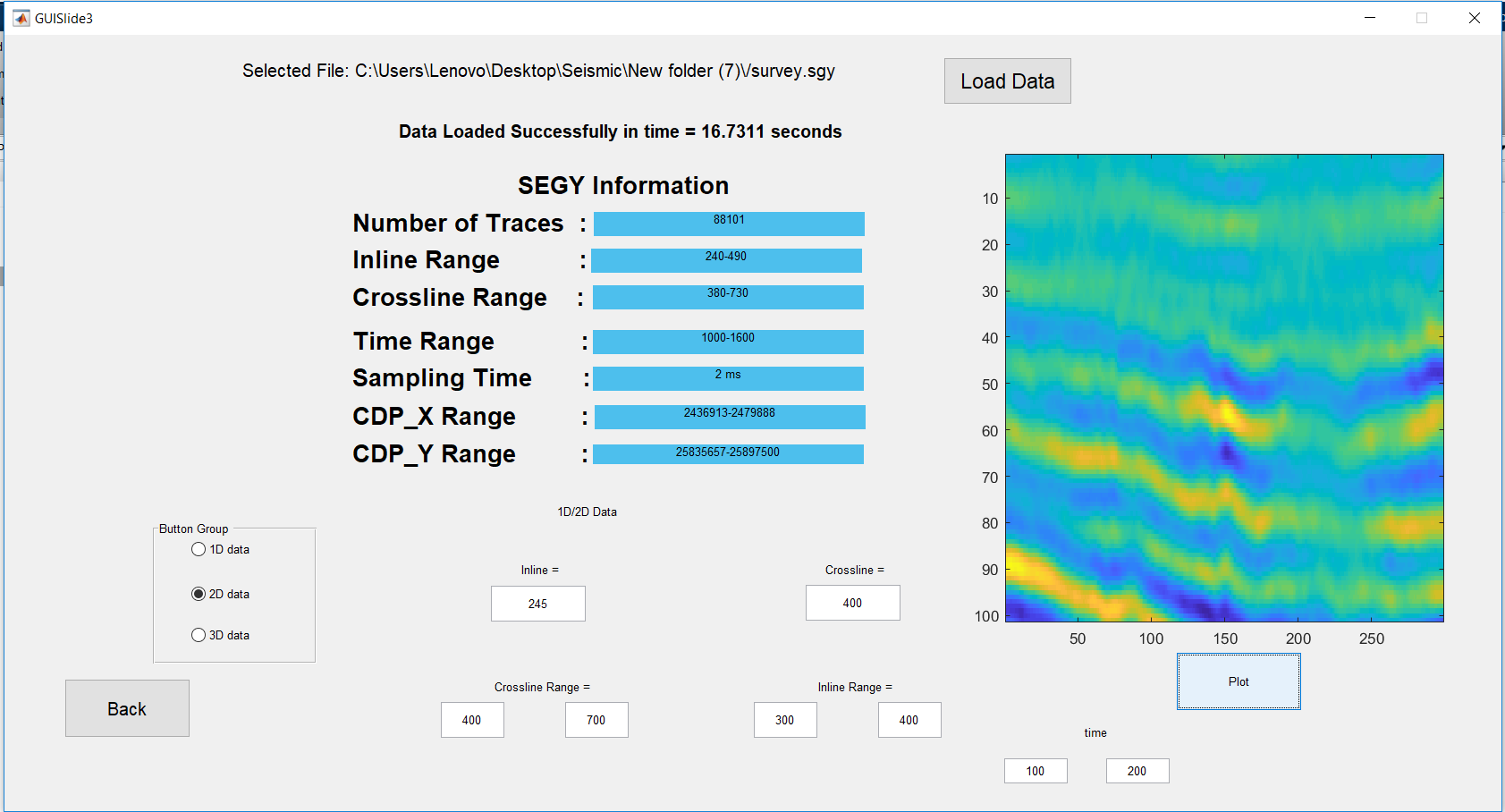
We used to prepare the GUI slides for the software for the sparse layer inversion in MATLAB. One of the example of the GUI slides prepared is:-

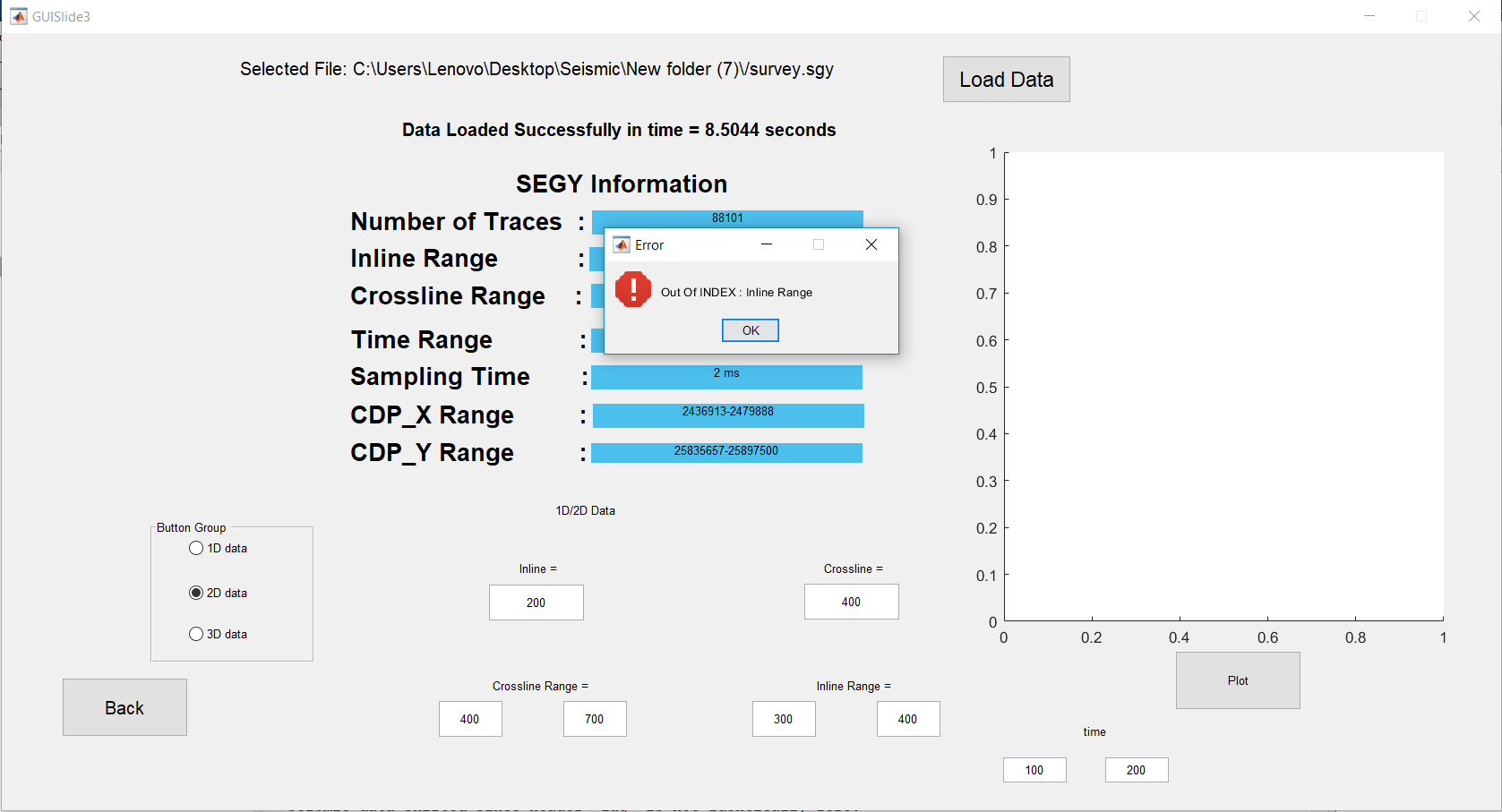


This slide shows representing the trace in either one, two or three dimensions by loading it from any well data. By selecting the respective crossline and inline values for the one dimensional representation or respective crossline and inline ranges for two and three dimensions we are able to plot the data in the respective forms. If we select wrong crossline ,inline values then error handling is also taken care of , it is as follows:-

Here first a correct representation is shown for a 2D trace by selecting correct values for the respective inline, crossline and the ranges for them, then intentionally selecting wrong values for them the error handling will be shown.

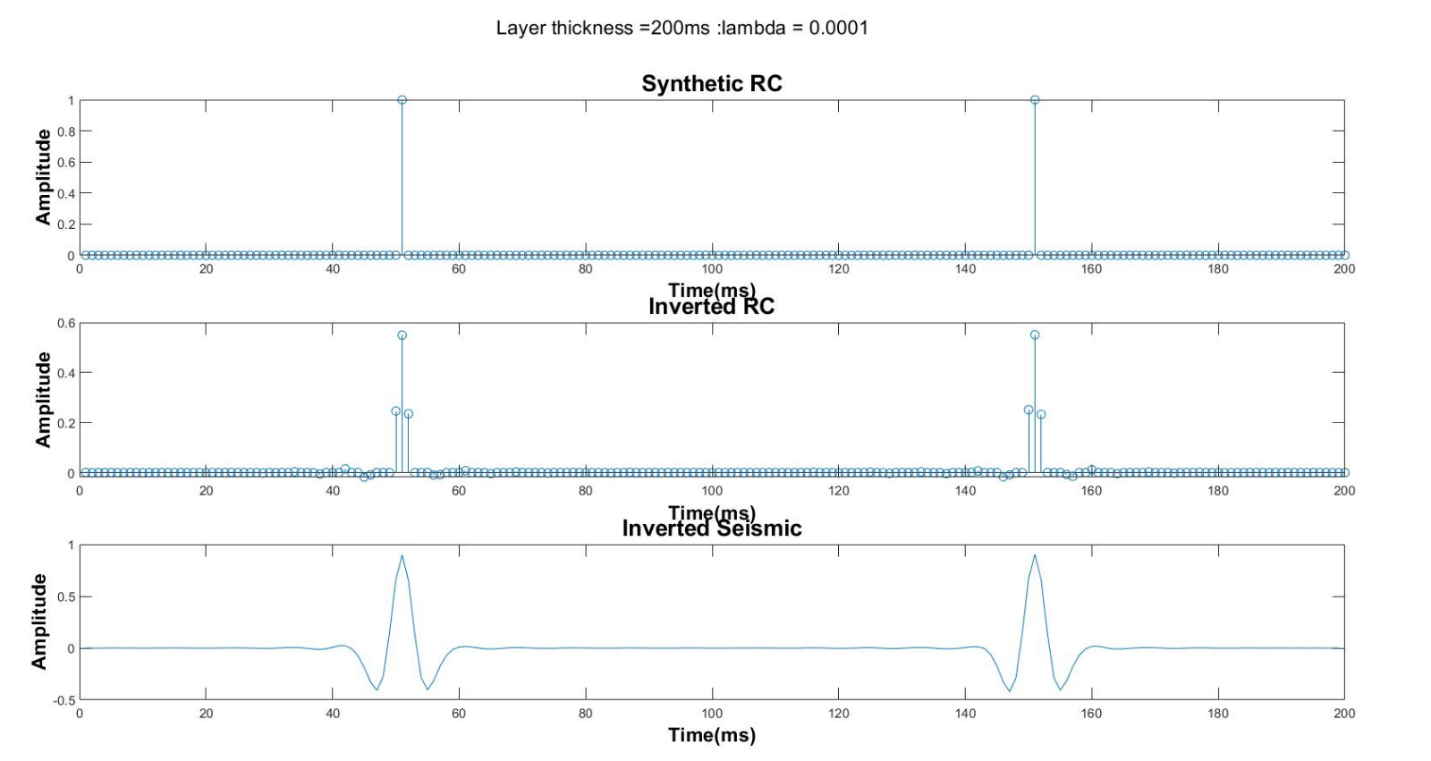
(all these loaded data are only for a particular well)

This is the picture for the correct values input

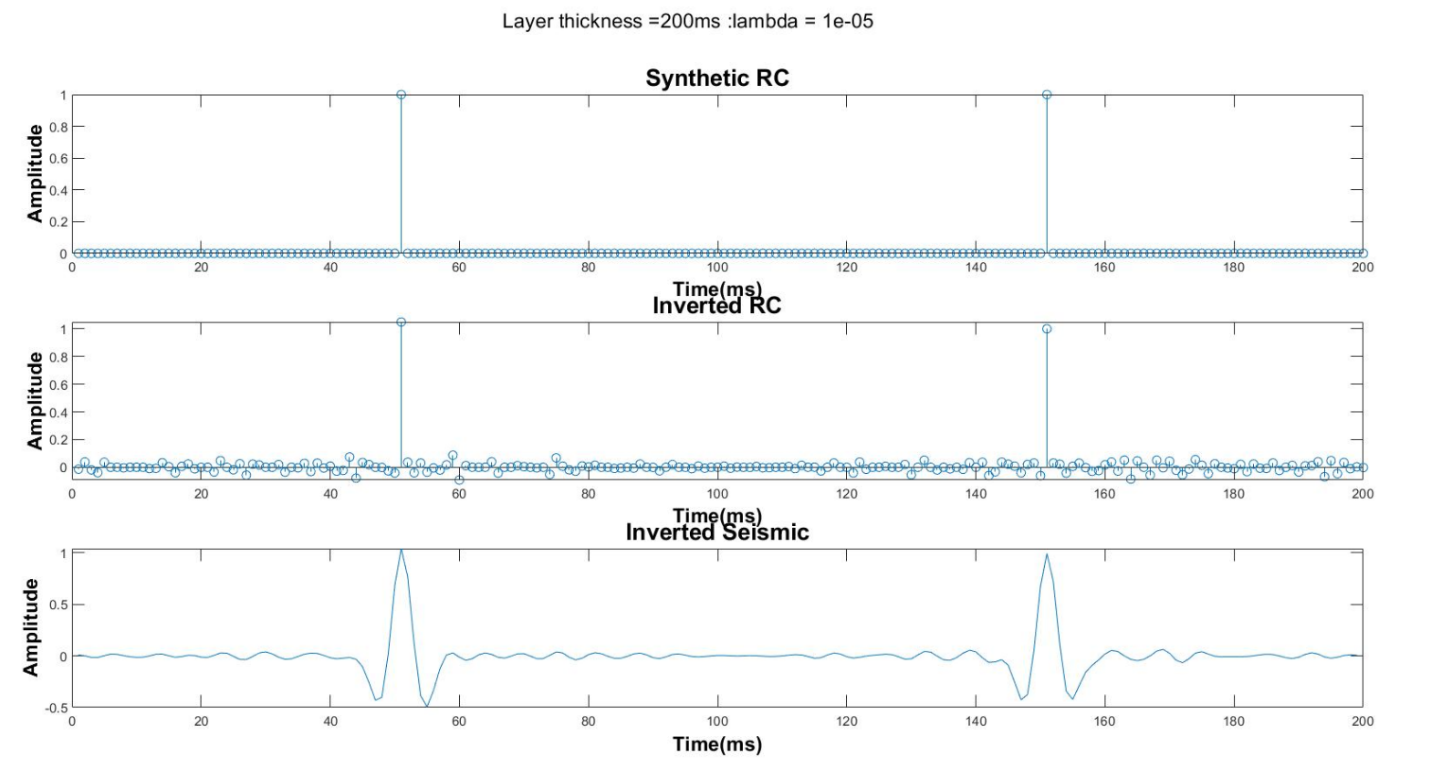


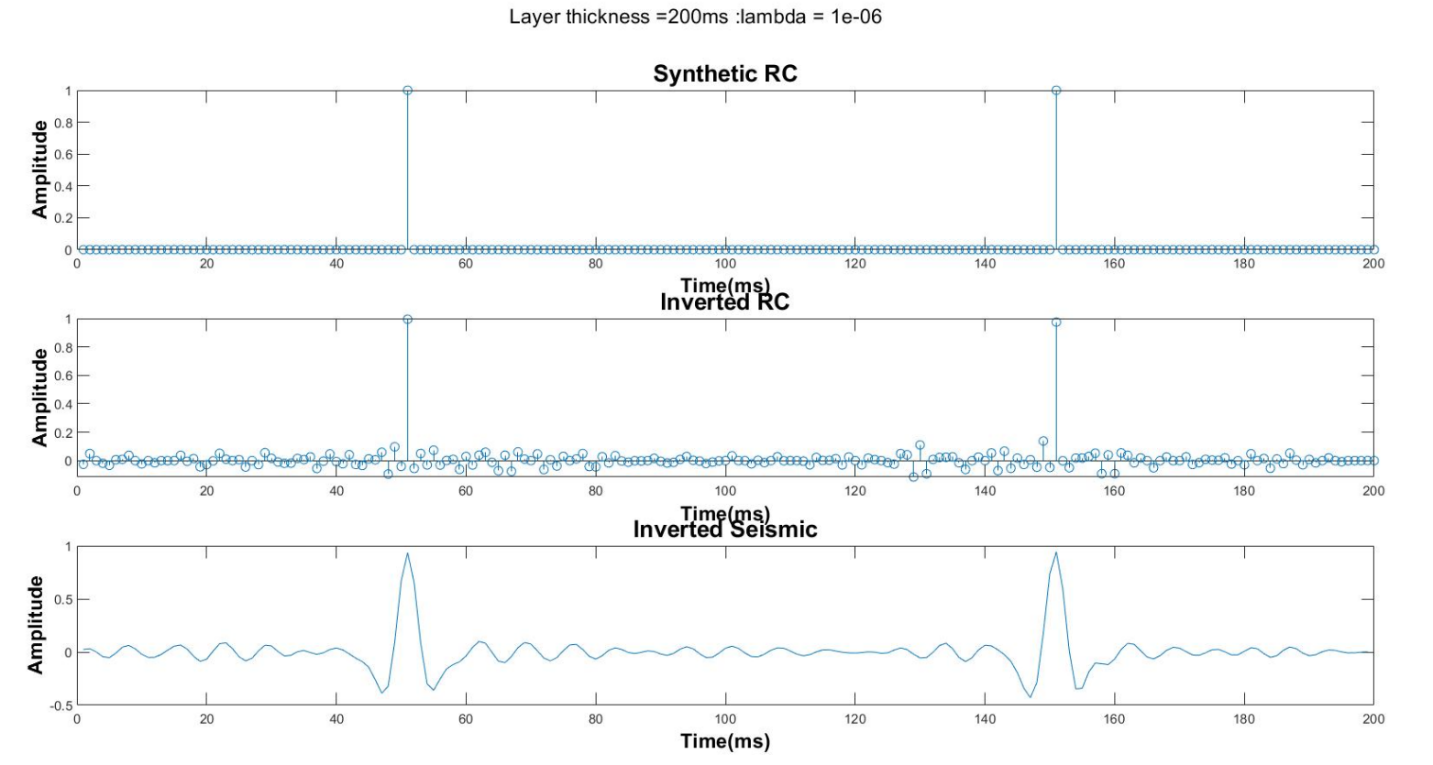
This is the error handling picture (wrong value for inline input is given) .A total 9 slides were prepared during my internship period. One of them is shown above (with error handling).

In addition to this there were various attempts to find a correlation between the inverted data and the well data with several lambda values. It was seen that this lambda values determine the weight between the two norms (L1 and L2 norms) or can be stated as the lambda values create a tradeoff between sparsity and accuracy. If the lambda values are decreased less than a certain threshold then it is seen that noise components become cruder, so a certain lambda is chosen for each inversion of the data. The effect of the above described notion is as follows:-



Lambda value = 10-4

 Lambda value = 10-5



Lambda value = 10-6

Conclusion

A new software that will be in favor of determining data on a new resolution scale is being developed. The resolution of the inverted data as compared to previous data will have a better resolution. The main motive is to use different wedge dictionaries to filter out the major layer responses and reconstruct the Seismic data using only thin layer responses only.

Generally, the thin layer responses are dominated by a thick layer. So in the process of data acquisition, such responses are hidden. So this software will help to construct the Seismic signal by elimination of the thick layer. This a new approach to the thin layer detection.

***My learnings***

Being a part of the above project was like discovering new areas. Indeed new areas of MATLAB was studied and work on GUI which was almost entirely a new area. Several ideas on Digital Signal Processing, sparse approximation problems, several approximation algorithms, Geophysics were studied.

References

1.) Introduction to seismic Inversion methods by Brian H Russell, Course Notes Series, No. – 2,

S.N Domenico, Series Editor.

2.) Seismic sparse-layer reflectivity inversion using basis pursuit decomposition, by Rui Zhang

and John Castagna, Volume 76, issue 6, SEG Library.

3.) Mathworks (for several topics on MATLAB and GUI).

4.) Overcomplete Dictionaries for Sparse Representation of Signals, Reasearch Thesis by

Michal Aharon.

5.) SIAM REVIEW, Vol-43, No. 1, pp. 129–159, 2001 Society for Industrial and Applied

Mathematics, “Atomic Decomposition by Basis Pursuit” by Scott Shaobing Chen, David L.

Donoho, Michael A. Saunders.