The Green's Function for the Diffusion Coefficient

by

Clarion Hadleigh Hess

Submitted to the Department of Earth, Atmospheric, and Planetary
Sciences
in partial fulfillment of the requirements for the degree of

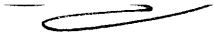Bachelor of Science in Earth, Atmospheric, and Planetary Sciences

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2012

Author ... **Signature redacted** ....
Department of Earth, Atmospheric, and Planetary Sciences
May 18, 2012

Certified by.. **Signature redacted** ....
Alison Malcolm
Assistant Professor, Department of Earth, Atmospheric, and
Planetary Sciences
Thesis Supervisor

Signature redacted

Accepted by .· ....
Professor Sam Bowring
Senior Thesis Coordinator, Department of Earth, Atmospheric, and
Planetary Sciences
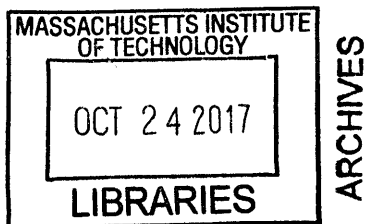
The Green's Function for the Diffusion Coefficient

by

Clarion Hadleigh Hess

Submitted to the Department of Earth, Atmospheric, and Planetary Sciences
on May 18, 2012, in partial fulfillment of the
requirements for the degree of
Bachelor of Science in Earth, Atmospheric, and Planetary Sciences

Abstract

The scattering diffusion coefficient between two points can theoretically be extracted from a random distribution of sources. An improved ability to measure the diffusion coefficient of the Earth's crust would simplify the process of characterizing the fracture network for applications in geothermal energy. This has the potential to make geothermal wells more economical to make, more efficient to operate, and longer lived. Previous work has shown the diffusion coefficient can be extracted from synthetic datasets in both one dimension and three dimensions using seismic interferometry. This paper attempts to recover the diffusion coefficient for a realistic source distribution taken from a microseismic dataset from a geothermal field in Indonesia. This dataset did not have an ideal distribution of sources, so the estimated diffusion coefficient did not match the expected value. A better estimate of the expected diffusion coefficient and an improved dataset with sources more evenly distributed in all directions around the receivers would likely give a better result.

Thesis Supervisor: Alison Malcolm
Title: Assistant Professor, Department of Earth, Atmospheric, and Planetary Sciences

Acknowledgments

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1   Seismic Interferometry

Seismic interferometry is the process of analyzing the interference patterns of seismic waves in order to estimate the properties of the interior of the Earth.[7, p. 1] Generally this is done for acoustic waves by using cross correlation, but it can also be done with deconvolution.[2] Seismic interferometry has evolved since 1968 when Jon Claerbout first presented his noniterative approach to analyzing a one dimensional, acoustically layered medium from seismogram data.[1] This process was successfully extended into three dimensions in 1999 by Rickett and Claerbout who used data from the Sun to reconstruct the Green's function between two points from the cross correlation of the noise signals at those two points.[5] In 2000, this theory was also shown to be able to investigate the interior of the Earth by Schuster and Rickett.[7, p. 16]

The Green's function is a fundamental characteristic of the medium between two points. The Green's function between any points A and B gives the response at point A to an impulse at point B.[3] This response is also the response at point B to an impulse at point A.[8, p. 309] This effectively places a virtual source at the location of a receiver.[12] This is a potentially important feature that could be applied to make seismic data acquisition easier as a receiver is much smaller and has less of an impact on the environment than a seismic source like dynamite.

The Green's function is dependent on D is the diffusion coefficient in units of

kilometers squared per second, t is the time vector in seconds, and r is the distance in kilometers between the two receivers. The one dimensional Green's function,

$$GF_{1D}(r,t) = \frac{1}{(4\pi Dt)^{1/2}} exp(-\frac{r^2}{4Dt}), \tag{1.1}$$

is used for one dimensional cases.[2] The three dimensional equation for the Green's function is

$$GF_{3D}(r,t) = \frac{1}{(4\pi Dt)^{d/2}} exp(-\frac{r^2}{4Dt}), \tag{1.2}$$

where d is the number of dimensions of the system.[4] The Paasschens paper included an additional term in the exponent of $\frac{ct}{l_a}$, where l_a is the absorption length.[4] This term was ignored because it is a function of dispersion rather than diffusion and this paper is focusing on diffusion.

Seismic interferometry has strict limitations on the source distribution. Equipartition of sources in which having sources located in every direction is equally likely is required in order to fully retrieve the Green's function.[12] This type of dataset could be constructed by either having an impulsive source that is recorded for a long enough period of time that the scatters enough that the scattered waves approach equally from all directions or having a randomly distributed sources surrounding the receivers. This even distribution is not generally the case in realistic data as receivers can rarely be surrounded by sources on all sides. This would depend on if the dataset is a microseismic dataset where the sources are underground or a surface generated source that would for example use dynamite or a vibroseis truck. In 2009, Fan and Snieder published a paper in which they examined the required source distributions for interferometry for diffusion. They determined that the one dimensional and three dimensional curves could be reconstructed by using a regular array of sources surrounding the pair of receivers.[2] Their paper was used as the basis for the one dimensional and three dimensional cases presented in this paper.

1.2 Diffusion Coefficient

Diffusion is the scattering of acoustic waves as it passes through a medium. The diffusion coefficient, D, with units of length squared over time, is given by

$$D = \frac{cl}{d},$$ (1.3)

where c is the velocity of the signal through the rock in kilometers per second, l is the mean free path for elastic, isotropic scattering in units of kilometers, and d is the unitless dimension of the system.[4] As a result, the diffusion coefficient is directly related to the mean free path of the material that the waves traveled through. The mean free path is the average distance that the signal can travel before it interacts with the medium and scatters.

These scattering events occur at a change of physical properties of the rock. This can include such things as a fracture in the rock that interrupts the transmission of the signal or a change of density or composition of the material. As a result, the diffusion coefficient helps determine where the Green's function peaks and the speed at which the Green's function subsides. The more that a signal scatters within a medium, the longer it will take a signal to travel away from any particular receiver. A large diffusion coefficient and the associated large mean free path will cause the function to fall off quickly and a small diffusion coefficient and mean free path will cause the function to fall off more slowly. This can be seen in Figure 1-1.

Assuming that scattering, as opposed to absorption, is the dominant source of energy loss in the medium, the diffusion coefficient can be correlated with the amount of fractures that are in the rock. As a result, the relative frequency of fractures in the medium could be predicted from an accurately estimated diffusion coefficient.

1.3 Applications to Geothermal Energy

Geothermal energy comes from thermal energy that is in the Earth's crust. This heat can be extracted by circulating water underground. This heated water can be

used directly as a heat source or indirectly to make electricity. Most of the United States would be able to generate electricity from geothermal energy if they were hydrofracked.[11] In the near future, this process could be implemented on a national scale to produce electricity at a price competitive with fossil fuels.[11] This paper discusses a technique that could make geothermal fields more productive by producing cheaper, more accurate map of the fracture network in the ground.

Geothermal energy comes from residual heat from the accumulation of the Earth and heat from radioactive decay. In the Earth's crust, the average thermal gradient ranges from of 15°C/km to 50°C/km.[11, p. 2-8] At 3.5 km, a depth commonly reachable by drilling, the rock in the continental United State ranges between 50°C and 150°C.[11, p. 2-15] Typically for large scale applications, this high temperature of the rock is recovered by circulating a medium such as water through the hot rock and back to the surface. The resulting hot water can be used to generate electricity.

There are other ways to take advantage of geothermal energy that work on a smaller scale. Some of these methods store heat underground via single wells. This type of energy may be appropriate for use in the seasonal cooling and heating of large commercial buildings. However, this direct use of the hot water is not suited for use on the national and global scale. I am focusing this paper on large scale applications, so I will not be referring to these cases.

Geothermal energy is most often found at depths ranging from 200 to 2000 meters. In order to extract the thermal energy from that depth at a commercial scale, multiple wells must be drilled. At least one of these wells will then be used to pump relatively cool water into the ground and at least one of these wells will be used to pump hot water out of the ground. This cycle extracts heat from the Earth's crust. The hot water that is extracted is then used either directly for heat or indirectly to generate electricity.[10]

Geothermal energy is a sustainable energy source, but it has drawbacks. In order to access the heat in the crust at the depths required, wells must be drilled. This requires land to temporarily hold a drilling rig, roads for large equipment and trucks, and an allowance for the noise and light pollution created from a drilling rig that

would be operated around the clock. If the hot rock is not permeable enough at the depth required, the rock will also need to be hydrofracked in order to allow water to circulate into and out of the ground on a timescale that is acceptable to the people running the geothermal operation.

This process of stimulation is known as hydrofracking or fracking (also fraking, fracing). The full effects of fracking are currently being debated. However, a properly sealed well does not allow either fracking fluid or any other substance to return up the well uncontrolled. This seal prevents any of these substances from coming into contact with the water table. Fracking also requires a large amount of water and chemicals. These chemicals could potentially spill into the environment at the surface. In drier areas, the sheer volume of water required to frack a well can be difficult to obtain, especially since most of the water is left in the well after fracking.

Once the well is actually constructed, geothermal energy has very little environmental impact. Depending on how electricity is extracted from the hot water that leaves the ground, the use of geothermal energy could have low to no carbon emissions. A unit of geothermal energy also ends up costing about the same amount as a unit of energy from fossil fuels. More research will need to be conducted prior to implementation on a national scale, but geothermal energy, as opposed to wind or solar, is competitively priced with respect to energy derived from fossil fuels.[11]

Geothermal energy is not widely used today partly because of the limited number of ideal areas to locate geothermal fields, especially on a large scale. An ideal area for geothermal energy needs to have hot rock, water, and a high porosity. Most of the existing geothermal plants are located in this type of situation. Although there are a relatively small number of areas with all three of these components, there are a large number of areas that have both hot rock and water. In these cases, porosity can be increased by hydrofracking to create and enhanced geothermal system (EGS). This type of system requires more investment and there is more uncertainty involved.

Geothermal energy is already commercially viable in Australia and Europe where the government financially supports geothermal energy.[11] However, more research is necessary in order to make enhanced geothermal systems competitive on a national

scale. More research is required to develop more cost effective drilling techniques, more reliable fracking technology, more accurate fracture characterization and more efficient heat conversion systems.[11]

Geothermal energy is worth developing as geothermal energy could be as cheap as energy from fossil fuels and it will not become scarce anytime in the near future. There are three orders of magnitude more geothermal energy available in EGS than the United States can use and the price of producing geothermal energy is competitive with the price of producing energy from fossil fuels.[11]

A better knowledge of the fracture characterization of the hot rock will significantly increase the efficiency of the geothermal plant. With a better knowledge of the physical location of the fractures, new wells can be accurately placed for the highest porosity. The high porosity will allow a larger volume of water to circulate through the system, increasing the total rate of heat extraction. Maximizing the number of fractures between wells will also maximize the amount of surface area that the water can contact the rock, potentially increasing the life of the well and increasing the total amount of heat that can be extracted. These factors increase the efficiency of the system, which makes the geothermal plant more economical and more competitive with other energy sources like fossil fuels. With geothermal energy as cheap to produce as energy derived from fossil fuels, geothermal energy could begin to take over the commercial sector. This would reduce pollution and the world's dependency on oil.
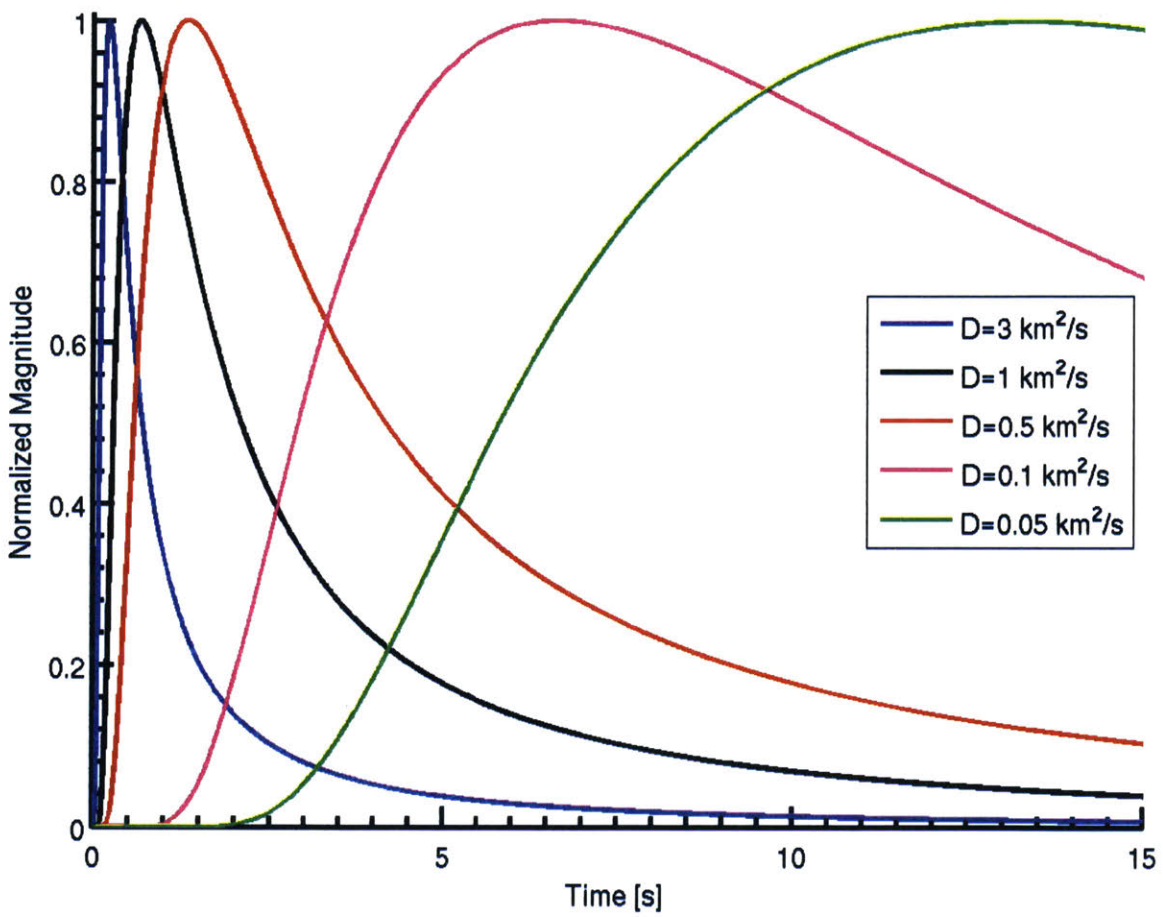
Figure 1-1: This figure shows the three dimensional Green's function for several different values of the diffusion coefficient. The larger the value of the diffusion coefficient, the earlier in time that the function peaks and the faster the curve falls off.

Chapter 2

Methods

2.1 Mathematical Calculation

For this paper, three different models were constructed based on the three different distributions of sources. The three models are the one dimensional idealized model, the three dimensional idealized model, and the three dimensional realistic model. The one dimensional idealized model contains sources distributed symmetrically and regularly around the receivers in a single line as shown in Figure 2-1. The three dimensional idealized model contains sources distributed symmetrically and regularly in a three dimensional grid around the receivers as shown in Figure 2-2. The three dimensional realistic model takes the locations of both its receivers and its sources from microseismic data from a geothermal field in Indonesia. This distribution can be seen in Figure 2-3.

These three models are processed in the same general manner. The Green's function (GF) between the two receivers is calculated directly using the one dimensional Green's function, Equation 1.1, or the three dimensional Green's function, Equation 1.2, as required by the model. This GF serves as the control to compare the result of the empirical calculation of the Green's function (EGF). The EGF between the two receivers is calculated using seismic interferometry. This EGF is then fit with a Green's function called the fitted Green's function (FGF) for the diffusion coefficient using nonlinear regression. The estimated diffusion coefficient of the FGF is then

compared to the the diffusion coefficient of the GF in order to determine the accuracy that the EGF approximates the GF. This result determines the quality of the source distribution.

For all of the following models unless otherwise specified time step of 0.005 seconds and the maximum time was 400 seconds. This time step of 0.005 seconds was chosen because 0.005 seconds is the time step of the waveforms from the microseismic dataset. The maximum time of 400 seconds was chosen because of the edge effects associated with the tail of the curve. This will be further examined in Section 4.1. The curves will only be shown for a time range of 0 to 15 seconds because that is the approximate range of waveforms for this dataset.

The values of the initial diffusion coefficient for the one dimensional case and the three dimensional case used in this paper were approximated from the physical properties of the region using Equation 1.3. The mean free path, l, is the inverse of the scattering coefficient, g_o, which has units of inverse length.[7] For a volcano, $g_o = 1\text{km}^{-1}$[6, p. 7], which approximates the composition of the rock in the geothermal field of the realistic data. Substituting a velocity of 3 km/s and a number of dimensions, d, of 1 or 3 depending on the number of dimensions of the model into Equation 1.3, the value for the diffusion coefficients are 3 km^2/s for one dimensional models and 1 km^2/s for three dimensional models.

The GF is calculated directly by using the diffusion coefficient, D, a time vector, t, and the distance between the two receivers, r. The exact equation for the GF that is used varies depending on the number of the dimensions of the model. One dimensional functions use Equation 1.1, and three dimensional functions use Equation 1.2. The absolute value of all functions is taken and then they are normalized by their maximum amplitude to a value of 1.

The EGF is calculated indirectly using cross correlation. The Green's functions from the first receiver to all of the sources are cross correlated with the Green's functions from the second receiver to all of the sources. In general, the cross correlation

of two functions $f(t)$ and $g(t)$ is given by

$$f(t) \star g(t) = f(-t) * g(t), \tag{2.1}$$

where $*$ is the convolution.[7, p. 6] The convolution is defined as

$$f(t) * g(t) = \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau, \tag{2.2}$$

where τ is a variable.[8, p. 229] These vectors were summed. The absolute value of the resulting EGF was normalized and fit with a FGF using the diffusion coefficient, D.

The EGF should approximate the GF. The accuracy of the EGF is determined by the nearness of the estimated diffusion coefficient to the diffusion coefficient of the GF. The amount that the estimated diffusion coefficient varies will determine how accurately the diffusion coefficient can be recovered by this method for this distribution of sources.

2.2   Idealized Models

The idealized models are used to determine the requirements of a regular source distribution to derive an accurate fit for the EGF. Multiple cases of each model were examined with different densities and ranges of sources.

2.2.1   1D Model

The physical setup of the one dimensional idealized model can be seen in Figure 2-1. For all one dimensional models, the first receiver is at -1 and the second receiver is at $+1$ with the sources arranged evenly about them. The parameters that control the distribution of the sources are the total range of the sources in kilometers, L, and the linear density of the sources in inverse cubed kilometers, ρ. Figure 2-1 shows Case 3 where L is 38 km and ρ is 1 km^{-1}. The parameters for the three cases presented in

23

1D Model	l [km]	ρ [km^{-3}]	D_o [km^2/s]
Case 1	14	1	3
Case 2	38	0.5	3
Case 3	38	1	3

Table 2.1: The parameters used to determine the source distributions for each idealized 1D case.

3D Model	l [km]	ρ [km^{-3}]	D_o [km^2/s]
Case 1	3	1	1
Case 2	12	0.5	1
Case 3	12	1	1

Table 2.2: The parameters used to determine the source distributions for each idealized 3D case.

this section can be seen in Table 2.1.

For each one dimensional distribution, the exact Green's function between the two receivers is calculated directly using Equation 1.1. For this model, the diffusion coefficient is 3 km^2/s from Equation 1.3. The resulting GF is used as the control for the EGF.

The EGF between the two receivers is calculated using seismic interferometry as described in Section 1.1. Then the EGF is fitted with a Green's function to produce the FGF. The estimated diffusion coefficient for FGF of this curve is compared to the initial diffusion coefficient of 3 km^2/s.

2.2.2 3D Model

The idealized model was expanded into three dimensions as shown in Figure 2-2. The receivers are located at [-1,0,0] and [1,0,0] and the sources are symmetrically arranged in a square grid centered around the two sources. The source distribution is determined by the length in kilometers of a side of the grid, L, and the volume density of sources in inverse cubed kilometers, ρ. In Figure 2-2, the parameters used are $L = 12$ km and $\rho = 0.5$ km^{-3}. The parameters for the three cases presented in this section can be seen in Table 2.2
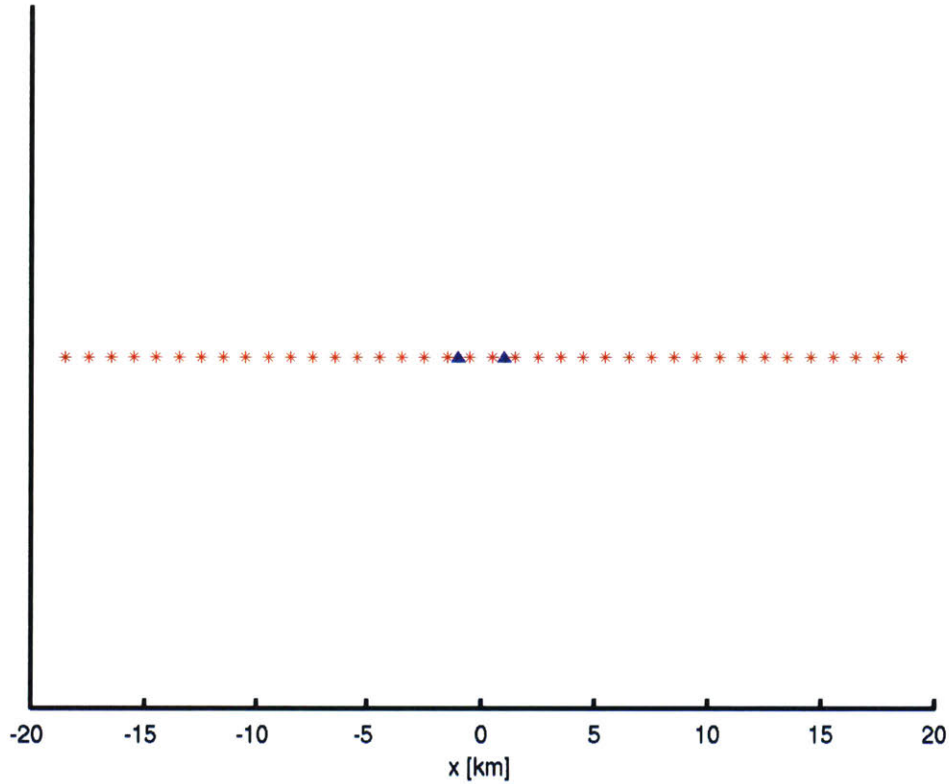
Figure 2-1: This figure shows the sources in red and receivers in blue for Case 2 of the idealized one dimensional model. The sources in this figure have a range of $L = 38$ km and a linear density of $\rho = 1$ km^{-1}.

2.3   Realistic Model

The locations of the sources and receivers for the realistic model were selected from a microseismic dataset. This dataset had multiple receivers operating at any point in time. The volume density of sources, ρ, and the maximum range of the sources, L, are important to be able to find the diffusion coefficient.[2] As a result, the pair of receivers that were chosen for this study had the largest number of receivers and the largest range. The locations of all of the events that occurred during the period of time that these two receivers were active were compiled together. This realistic distribution of sources that is used for this paper is shown in Figure 2-3.

With these locations, the model is processed as described in Section 1.1 using a
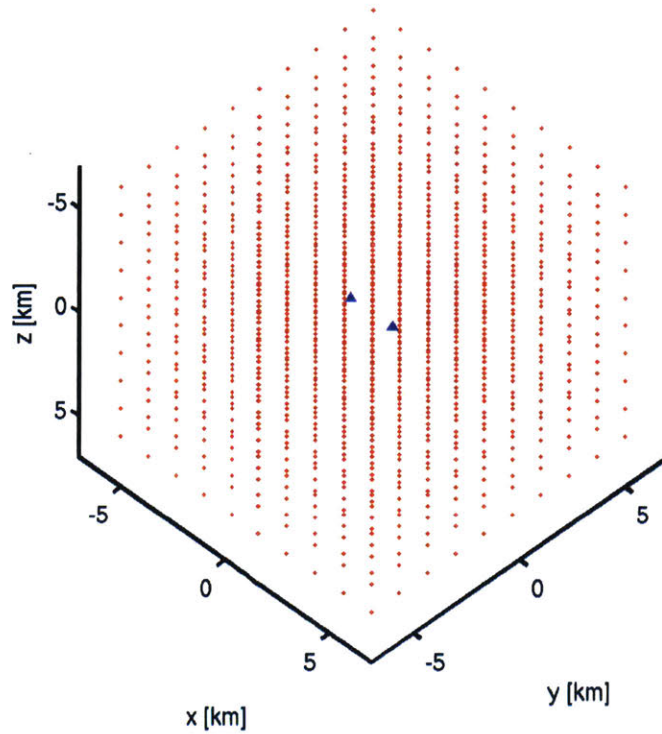
Figure 2-2: This figure shows the sources in red and receivers in blue for Case 2 of the idealized 3D model. The sources in this figure have a range of $L = 12$ km and a density of $\rho = 0.5$ km^{-3}. The point of view of this figure is 45° from each major axis.

time step of 0.005 seconds, a maximum time of 400 seconds and an diffusion coefficient of 1 km^2/s. The GF is calculated directly using Equation 1.2. The EGF is calculated using seismic interferometry. The FGF is fit to the EGF and the estimated diffusion coefficient from the FGF is compared to the diffusion coefficient.

Figure 2-3: The distribution of the sources and receivers for the realistic model.

Chapter 3

Results

3.1 Idealized Models

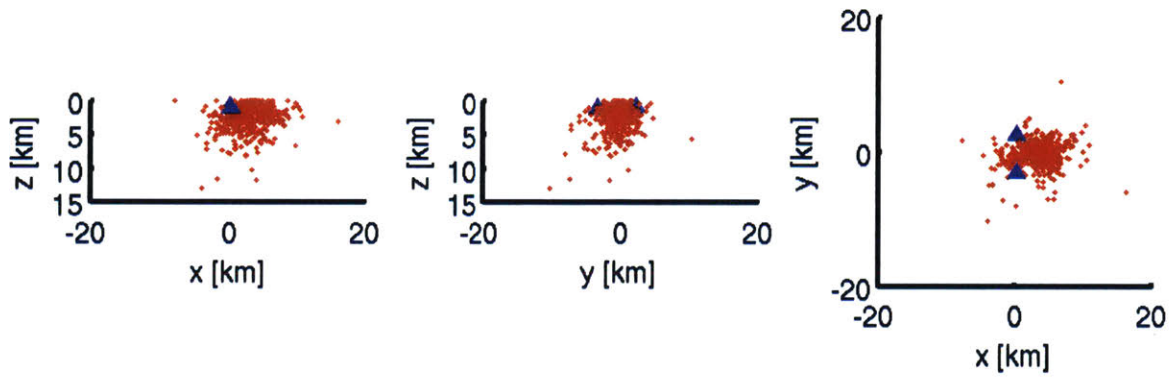### 3.1.1 1D Model

The one dimensional model was completed for a series of different values of the range of sources, L, and the linear density of sources, ρ, in order to determine the range and density required to accurately reproduce the Green's function. The results for the three one dimensional cases in this section are summarized in Table 3.1.

Case 1 is shown in Figure 3-1. This case has a relatively small range and high density of sources. The EGF closely matches the part of the GF curve to the left side of the maximum. The part of the EGF curve to the right side of the maximum of the GF curve does not match as the EGF falls off more quickly. The FGF is very inaccurate. The diffusion coefficient of the fit for this case is 14.45 km^2/s, a 381.7% error. This is a large total error even though the initial part of the EGF and GF

1D Model	L [km]	ρ [km^{-1}]	$D_\circ$ [km^2/s]	D_{est} [km^2/s]	Percent Error
Case 1	14	1	3	14.45	381.7%
Case 2	38	0.5	3	3.41	13.7%
Case 3	38	1	3	3.09	3.0%

Table 3.1: The parameters and error of the estimated diffusion coefficient for each idealized 1D case using a time step of 0.005 seconds and a maximum time of 400 seconds.
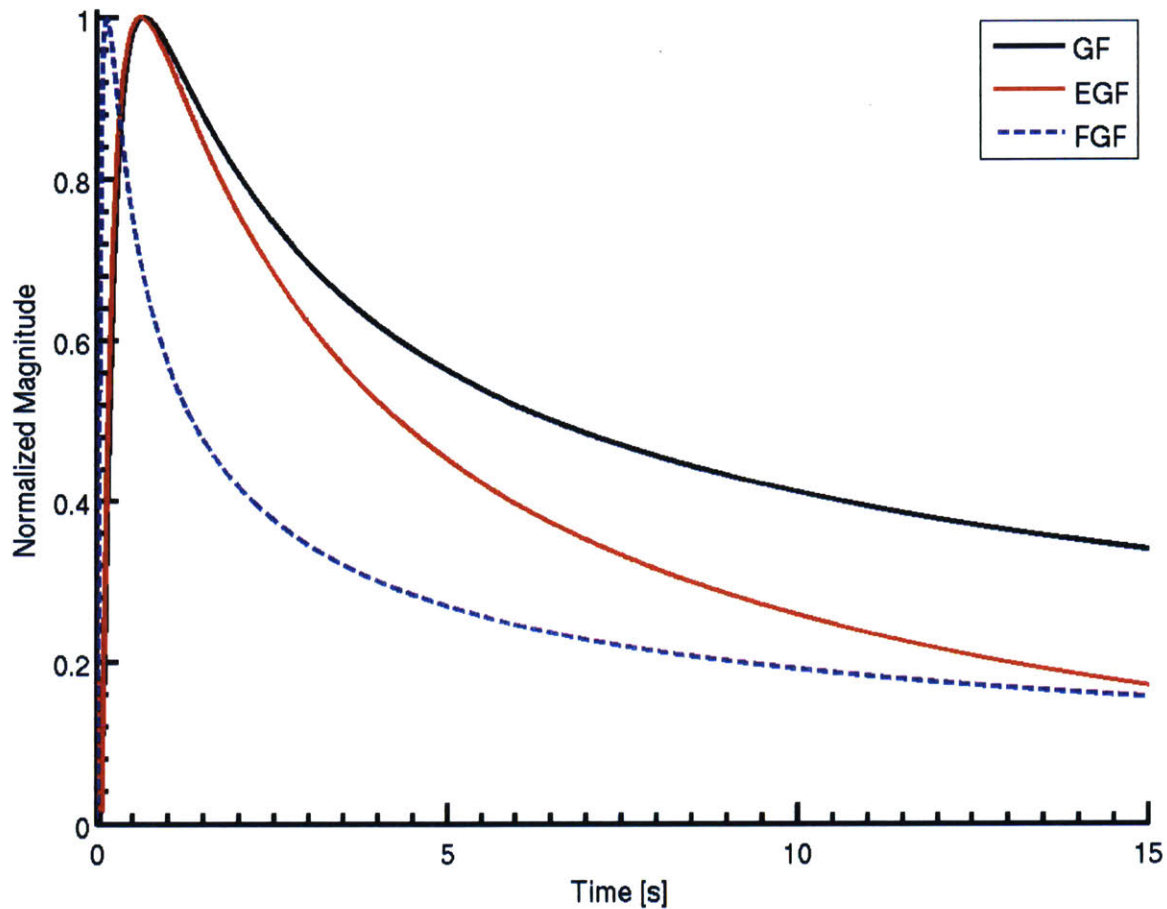
match.

Figure 3-1: This is Case 1 for the idealized 1D model with $L = 14$ km and $\rho = 1$ km^{-1}. The estimated diffusion coefficient is 14.45 km^2/s, a 381.7% error.

Case 2 can be seen in Figure 3-2. This case has 19 sources, five more than the Case 1, though the sources are spread out at half of the density of Case 1. The resulting set of curves is much more accurate than the first case. The EGF matches the GF on the right side of the peak fairly closely, especially at larger times. However, the left side of the peak of the EGF does not match the GF as well as it did in Case 1. The peak of the EGF also occurs earlier than that of the GF. The FGF matches the GF fairly closely. The estimated diffusion coefficient is 3.41 km^2/s. This is only 13.7% different from the diffusion coefficient of 3 km^2/s.

30

Figure 3-2: This shows the GF, EGF, and FGF of Case 2 for the idealized 1D model with $L = 38$ km and $\rho = 0.5$ km^{-1}. The FGF has an estimated diffusion coefficient of 3.41 km^2/s, an error of 13.7%.

Case 3 had 38 sources and can be seen in Figure 3-3. This case has both a high density of sources and a large range. In Case 3, the EGF most closely matches the GF. As a result, the FGF also closely matches the GF. The peak of the EGF still peaks slightly earlier than the GF and the FGF. The estimated diffusion coefficient is 3.09 km^2/s. This is only 3% larger than the diffusion coefficient.
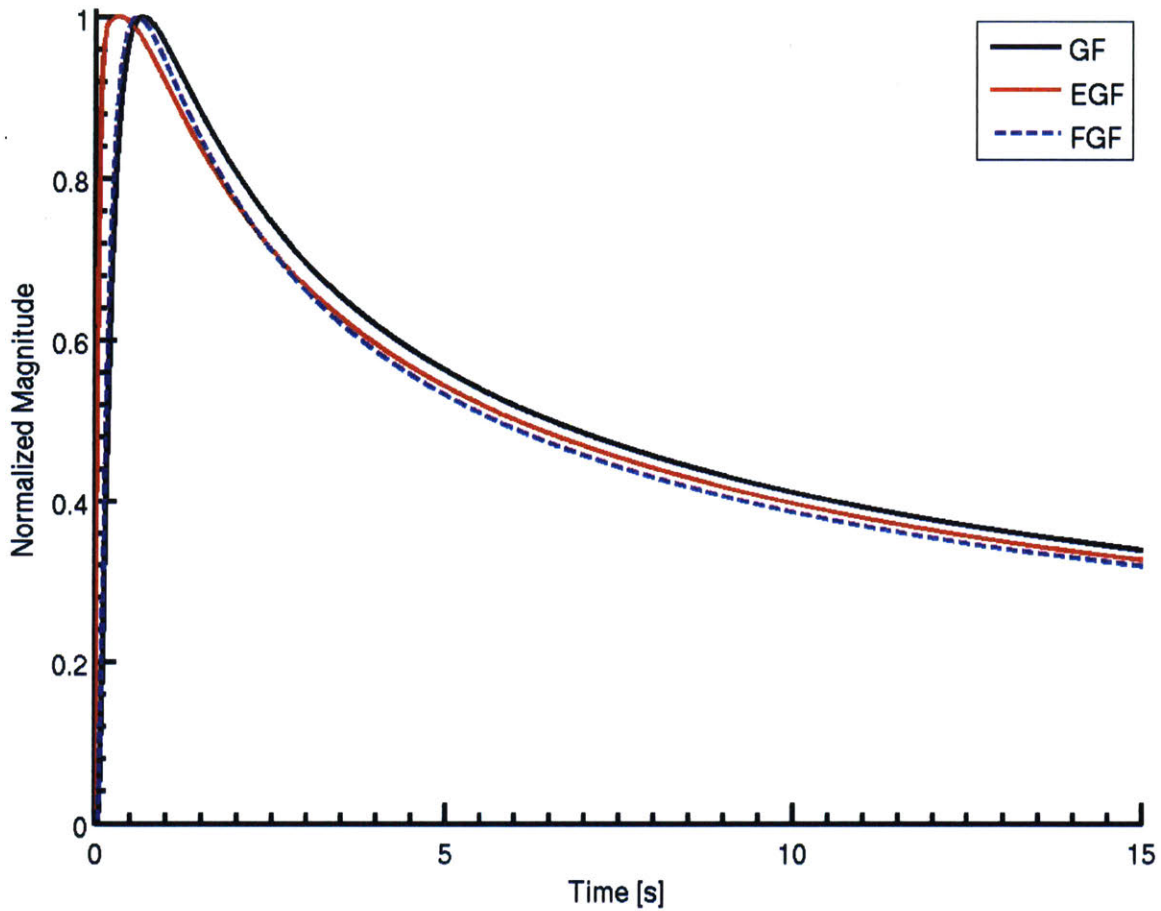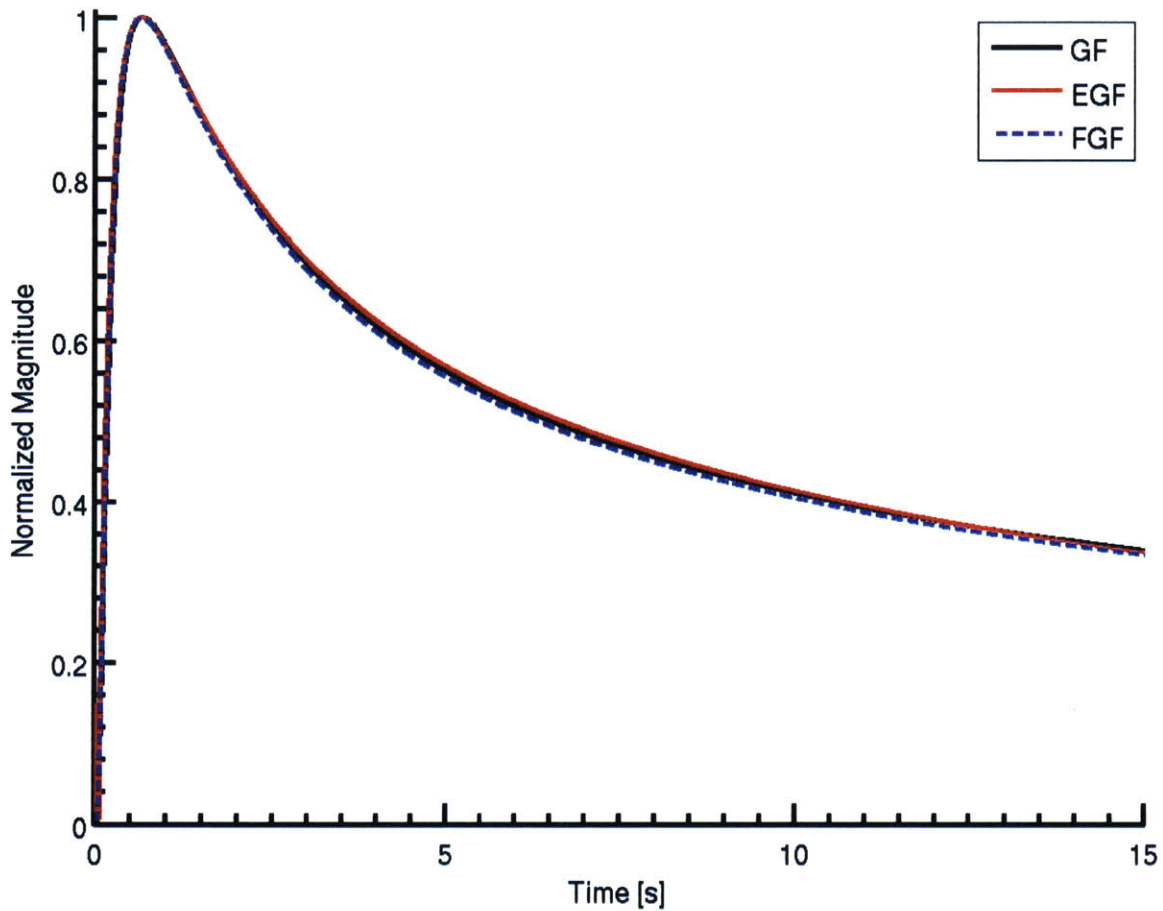
Figure 3-3: This shows the GF, EGF, and FGF for the Case 3 for the idealized 1D model with $L = 38$ km and $\rho = 1$ km^{-1}. The FGF has an estimated diffusion coefficient of 3.09 km^2/s, and error of 3.0%.

3.1.2   3D Model

The three dimensional idealized model was repeated for various values of L and ρ. The results from three examples included in this section are shown in Table 3.2. In a similar manner to the one dimensional cases, Case 1 and Case 3 cases have the same density and range over a much different volume. These two cases were compared with the same GF that was calculated using a diffusion coefficient of 1 km^2/s, the distance between the two receivers, and the time vector.

Case 1 can be seen in Figure 3-4. The sources for this case have a relatively small

3D Model	L [km]	ρ [km^{-3}]	$D_\circ$ [km^2/s]	D_{est} [km^2/s]	Percent Error
Case 1	3	1	3	1.36	36.34%
Case 2	12	0.5	3	0.83	16.6%
Case 3	12	1	3	1.01	1.16%

Table 3.2: The parameters and error of the estimated diffusion coefficient for each idealized 3D case using a time step of 0.005 seconds and a maximum time of 400 seconds.

range. The FGF generally matches the EGF except for a small difference at times larger than 4 seconds. These two curves do not match the GF. They are fairly close on the left side of the peak. The GF falls faster than the EGF on the right side. The peak of the EGF and FGF also lead the GF slightly. The estimated diffusion coefficient is 1.36 km^2/s. This has an error from the initial diffusion coefficient of 36.34%.

Case 2 is shown in Figure 3-5. The source distribution for this case has a large extent and a smaller density. The FGF matches the EGF very closely. The EGF more closely resembles the GF for this case. The three curves almost match on the left side of the peak, but there is some separation in the curves from the peak through the right side of the graph. The peak of the GF leads the peak of the EGF and FGF and the GF falls off more steeply to the right of the peak. There is still a significant difference between the FGF and the GF. The estimated diffusion coefficient of the is 0.83 km^2/s, an error of 16.6%.

Case 3 is shown in Figure 3-6. Case 3 has a source distribution that is both dense and has a large range. The GF, EGF, and FGF all overlap for the range of this graph. The estimated diffusion coefficient is 1.01 km^2/s, only 1.01% different from the original value of 1 km^2/s.

3.2 Realistic Model

The realistic distribution of sources was used for this section. The GF, EGF, and FGF for this model can be seen in Figure 3-7. The EGF is not very similar to the GF. The FGF also does not closely match the EGF, even though they meet at the
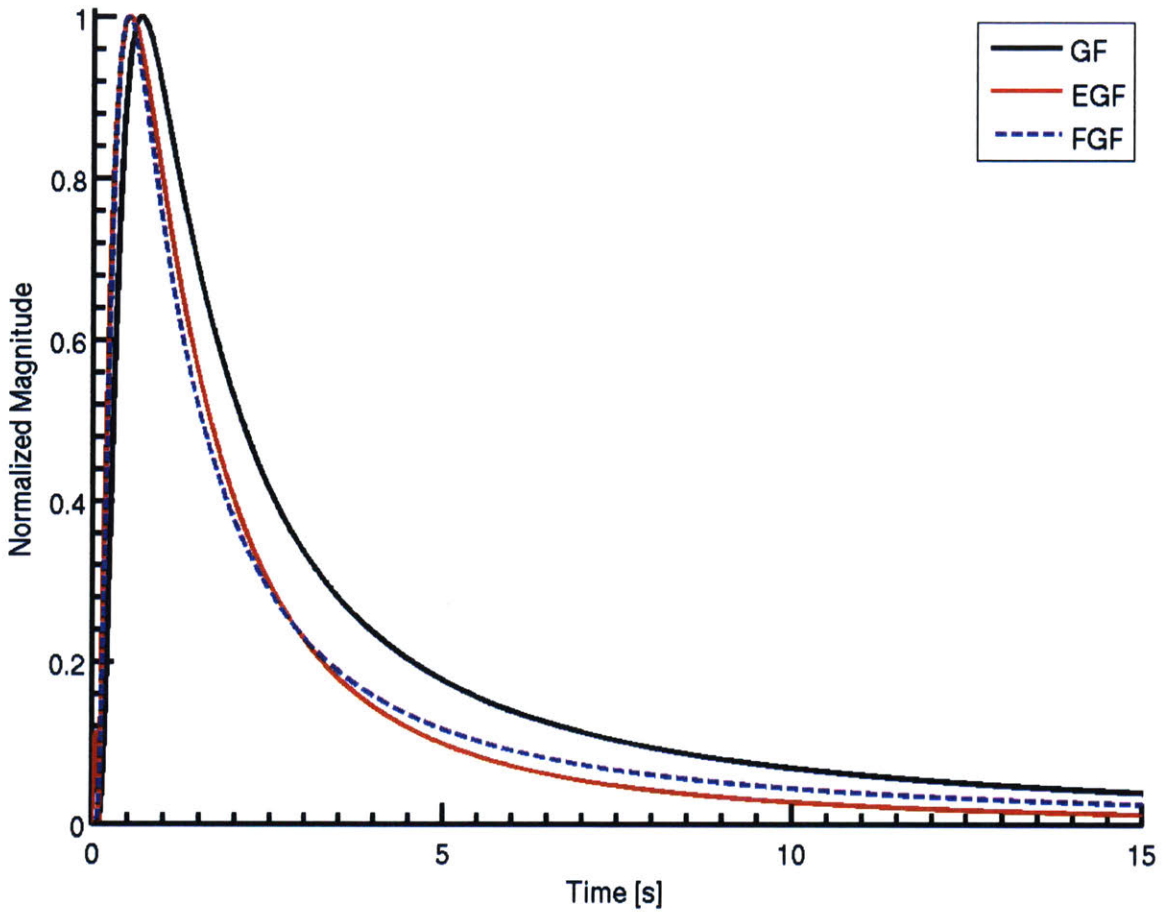
33

Figure 3-4: This shows the GF, EGF, and FGF of Case 1 for the idealized 3D model with $L = 3$ km and $\rho = 1$ km^{-3}. The estimated diffusion coefficient of the FGF is 1.36 km^2/s, an error of 36.34%.

peak and the edges of the graph. The width of the EGF is wider than the GF. The estimated diffusion coefficient is 2.49 km^2/s. This is an error of 149%.
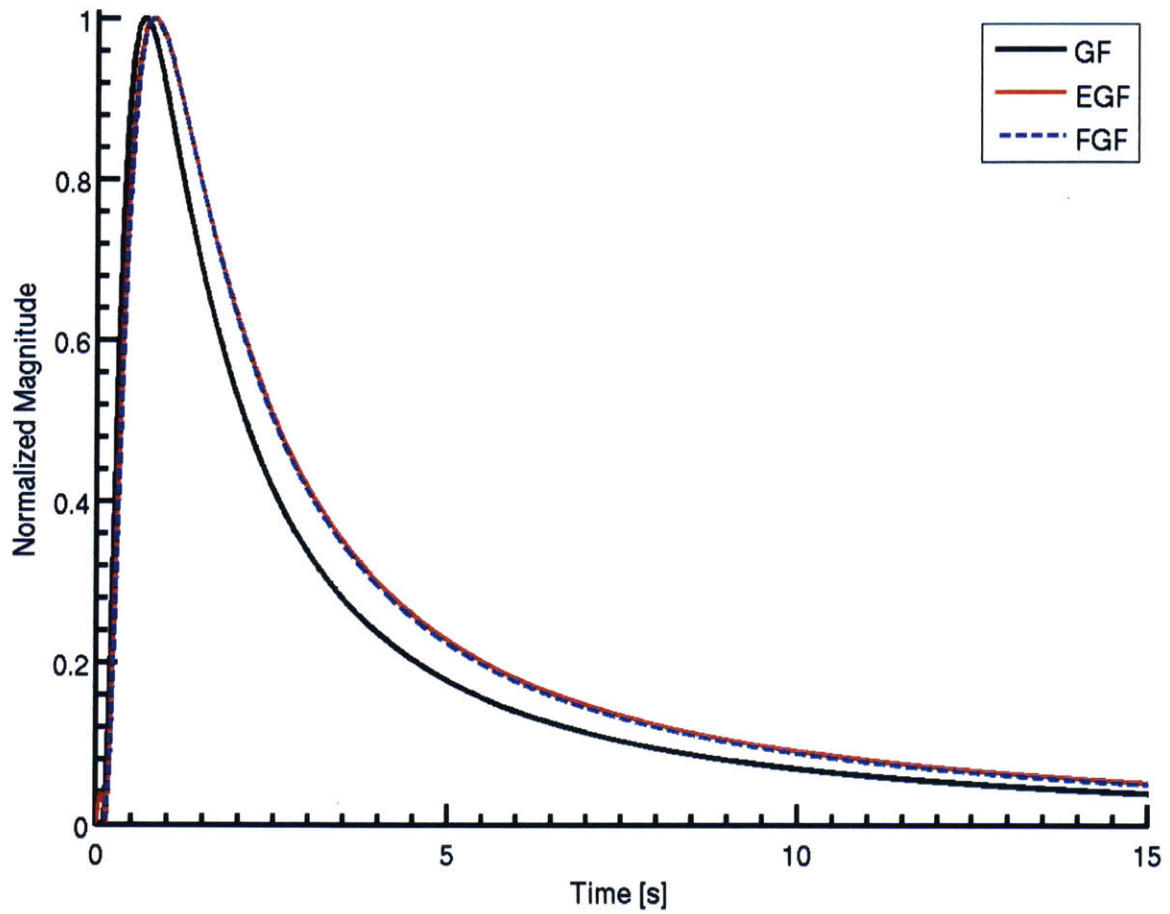
Figure 3-5: This shows the GF, EGF, and FGF of Case 2 for the idealized 3D model with $L = 12$ km and $\rho = 0.5$ km^{-3}. The FGF has an estimated diffusion coefficient of 0.834 km^2/s, an error of 16.6%.
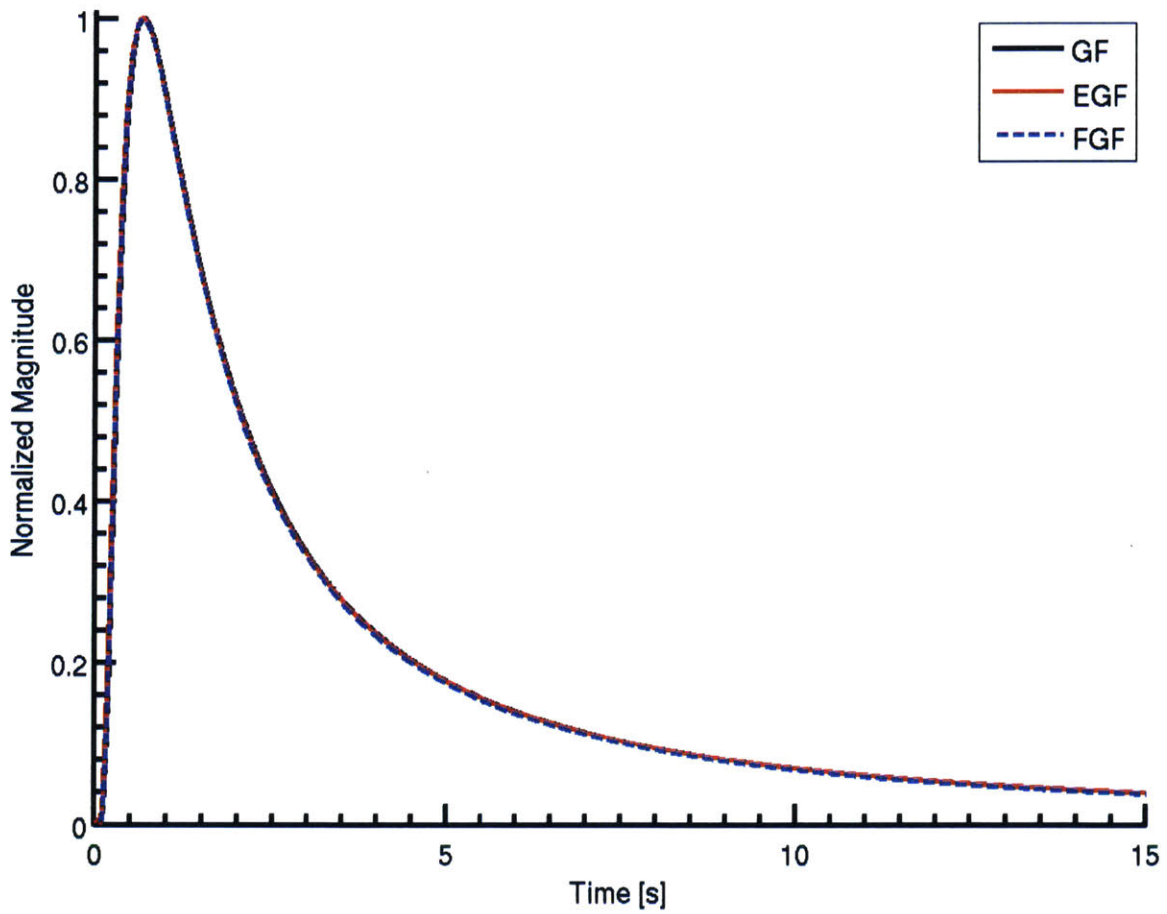
Figure 3-6: This shows the GF, EGF, and FGF of Case 3 for the idealized 3D model with $L = 12$ km and $\rho = 1$ km^{-3}. The FGF has an estimated diffusion coefficient of 1.01 km^2/s, an error of 1.16%.
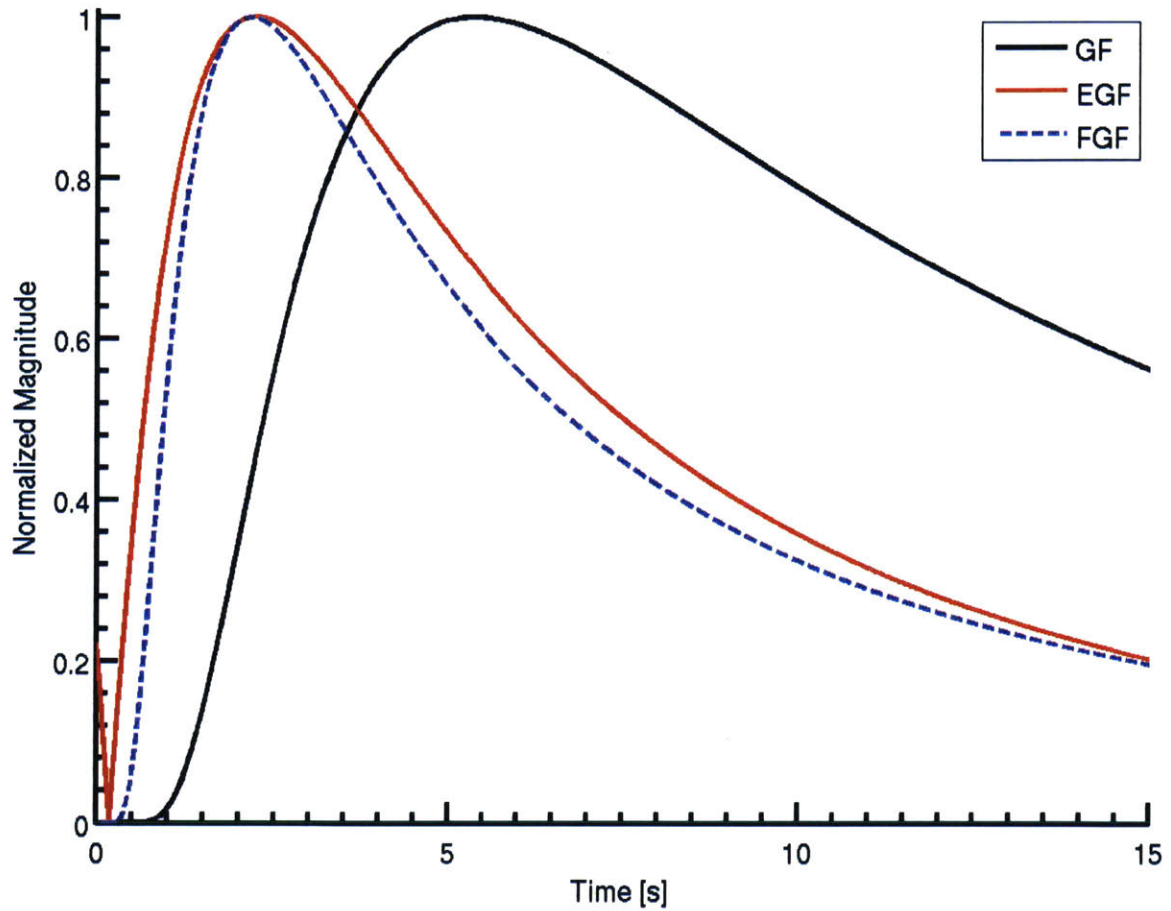
Figure 3-7: These are the GF, EGF, and FGF curves for the realistic 3D model. The estimation of the diffusion coefficient is 2.49 km^2/s, a 149% error.

Chapter 4

Discussion

4.1 Time Dependence and Artifact

A fixed time vector length was used for this paper. The maximum time affects how the EGF is ultimately shaped. The EGF curve in Figure 4-1 and Figure 4-2 that used a maximum time of 15 seconds is compared against the original EGF curve with a maximum time of 400 seconds. At a smaller maximum time value, the behavior of the tail of the EGF curve with a maximum time of 15 seconds decreases and finally drops quickly to zero near the maximum time due to the edge effects of cross correlation. The 400 second curve also drops off, but it drops off around 400 seconds. The height of the curve at 400 seconds is much smaller than it is at 15 seconds, so it has a much smaller effect on the graph. Additionally, the 15 second curve begins to fall below the 400 second curve around the 8 second mark. The 400 second graph does not fall like this over the range graphed. It is a subtle difference, but this large time value was used for this paper in order to prevent this processing artifact from affecting the outcome of the study. The full range of the 400 second curve can be seen in Figure 4-3 and Figure 4-4.

The time step that is used in this paper was held constant for the paper at 0.005 seconds. Figure 4-5 shows Case 3 of the 3D idealized model in which the EGF and the GF appear to overlap. The third curve is the EGF with a time step of 0.1 seconds. The smaller time step causes the EGF to become less accurate, especially near the
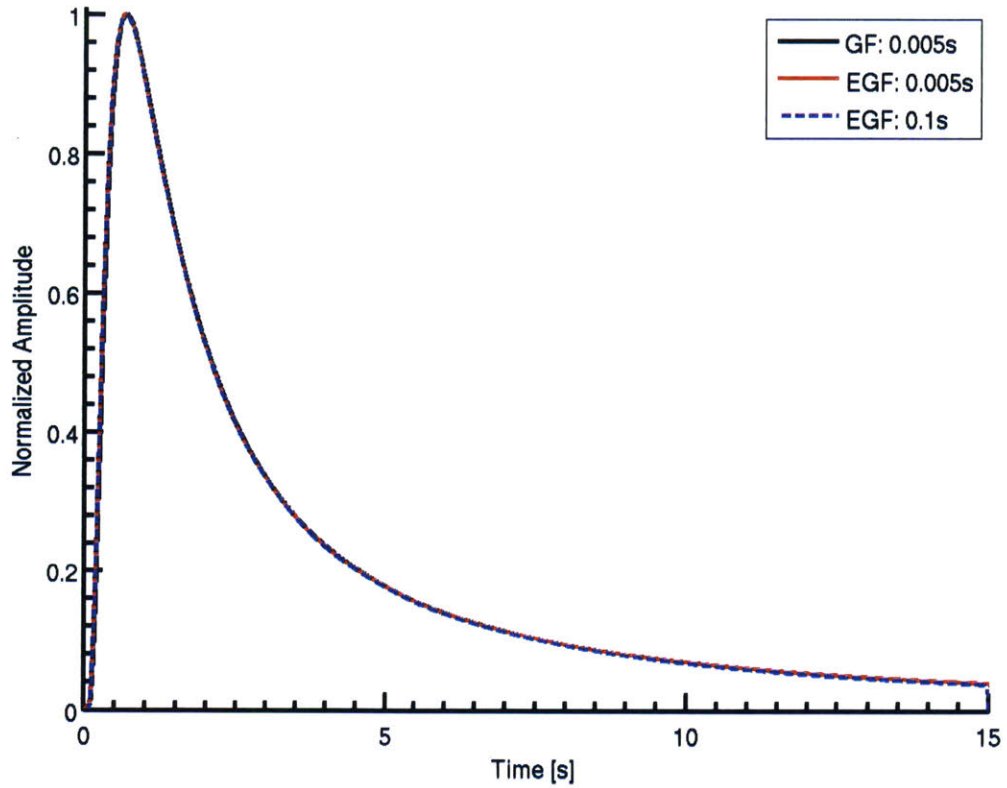
Figure 4-1: Two EGF curves for the 3D Case 3 that demonstrate the effects of using a shorter maximum time. EGF_{long} has a maximum time of 400 seconds and EGF_{short} has a maximum time of 15 seconds.

origin.

Several of the graphs that have been included in this paper have an obvious anomaly near the origin. This anomaly, shown in Figure 4-6, is an artifact of processing. It is likely caused by either processing by using circular cross correlation or not setting the negative Green's function to 0 for negative times. It may also be an artifact of the time window starting at zero. This artifact appears in the EGF for the all of the one dimensional and three dimensional models. It is potentially affecting the location of the peak of the EGF, which could in turn affect the error of the FGF and the estimated diffusion coefficient.
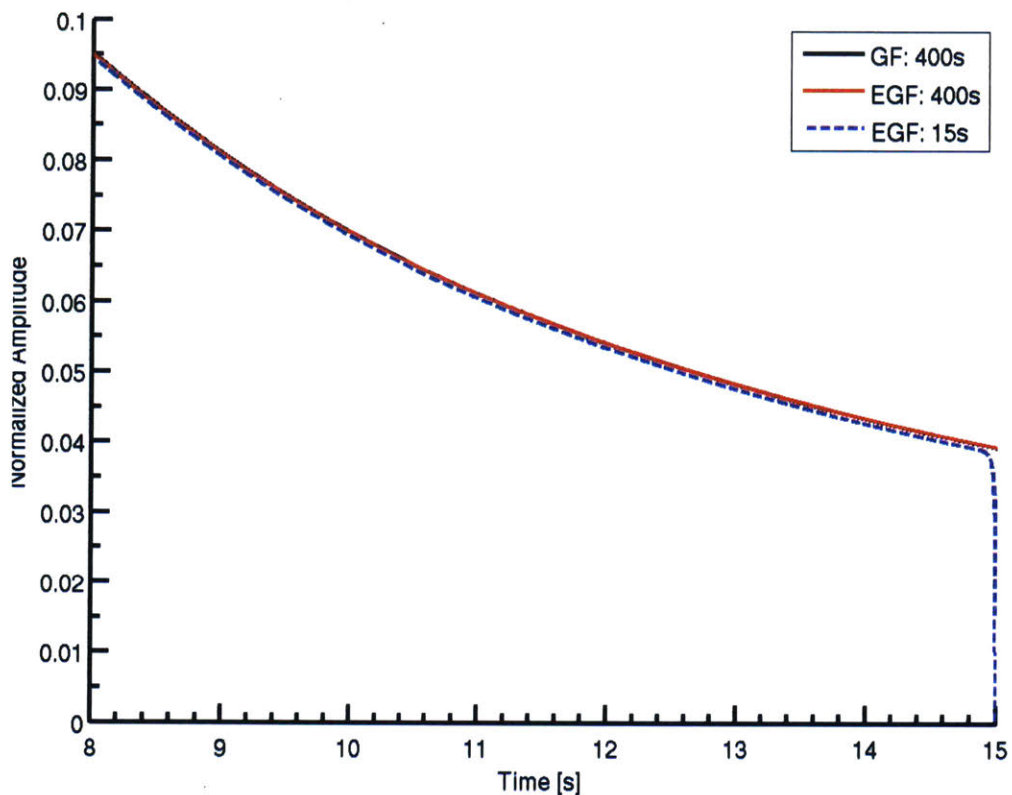
Figure 4-2: Two EGF curves for the 3D Case 3 that demonstrate the effects of using a shorter maximum time. EGF_{long} has a maximum time of 400 seconds and EGF_{short} has a maximum time of 15 seconds. This figure shows the curve from 10 seconds to 15 seconds.

4.2 Idealized Models

4.2.1 1D Model

Three cases for the idealized one dimensional model demonstrated in Section 3.1.1 demonstrate the one dimensional dependence of the estimated diffusion coefficient on the source distribution. Case 1 has a small range and a high density of sources. Case 2 has a large range and a low density of sources. Case 3 has a large range and a large density of sources. These different source distributions strongly affect the accuracy of the EGF and FGF. The summarized one dimensional results can be seen in Table 4.1
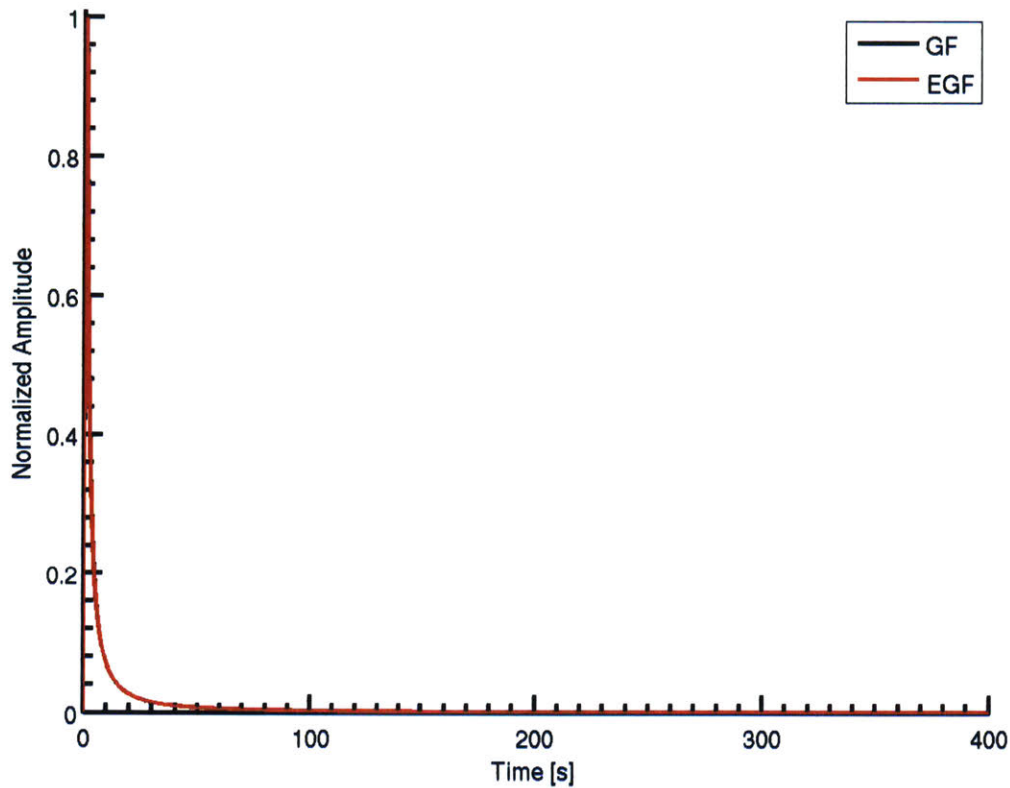
41

Figure 4-3: These are the full calculated curves from 0 to 400 seconds for the GF and the EGF between the two receivers for Case 3 of the idealized 3D model.

In Case 1, as seen in Figure 3-1, the error of 381.7% is very high. This is the error for both the estimation of the diffusion coefficient and the estimation of the mean path length since the two variable are linearly related. The EGF matches the left side of the GF peak closely, but the EGF falls off more quickly on the right side. The result is a large difference in the value of the diffusion coefficient between the GF and the FGF. The estimated diffusion coefficient of 14.45 km^2/s is much greater than that of the initial diffusion coefficient of 3 km^2/s. Using Equation 1.3 and assuming a velocity of 3 km/s, this means that instead of a mean free path of 1 km, the mean free path would be 4.82 km.

The GF, EGF, and FGF curves in Case 2 in Figure 3-2 much more closely resemble one another. In this case, the tails are very similar, but the EGF does not fit the GF very well. The estimated diffusion coefficient for this case is 3.41 km^2/s as compared
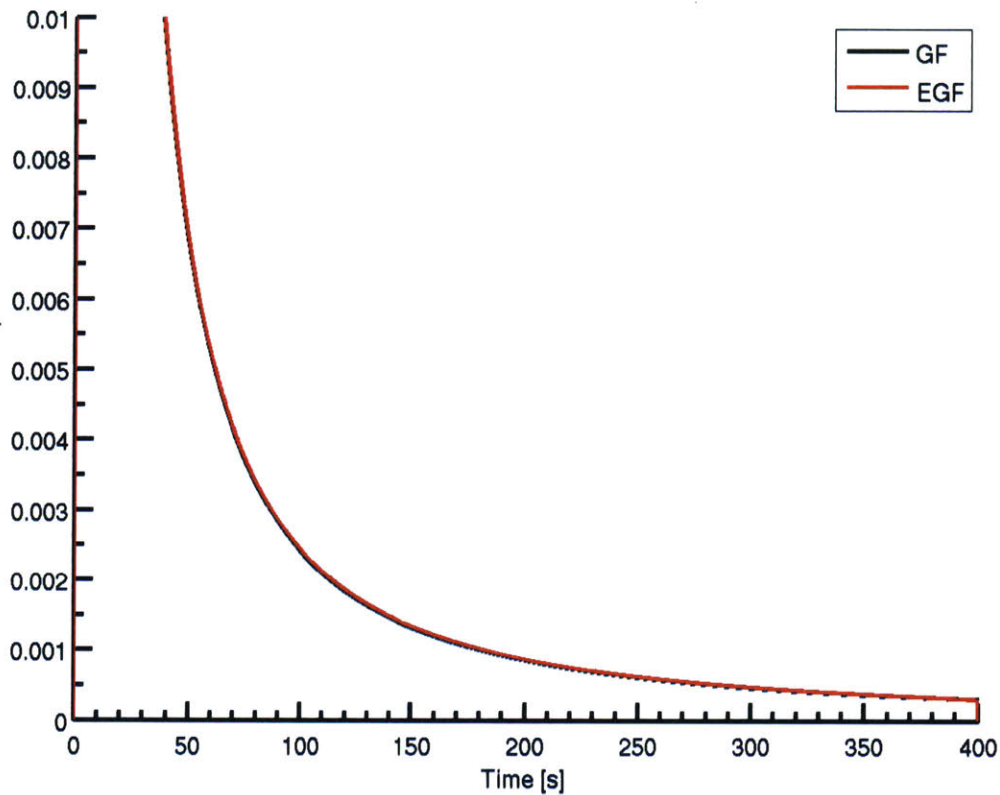
Figure 4-4: These are the full calculated curves from 0 to 400 seconds for y values from 0 to 0.1 of the GF and the EGF between the two receivers for Case 3 of the idealized 3D model.

with the diffusion coefficient for the GF of 3 km^2/s. This is a 13.7% error and this FGF would result in a mean free path of 1.13 km.

Case 3 in Figure 3-3 has both a large range of sources and a high density of sources. The EGF and FGF almost completely agree with the GF. Consequently, the estimated diffusion coefficient of 3.09 km^2/s is only 3% larger than the diffusion coefficient of 3 km^2/s. The resulting mean free path of the FGF would be 1.03 km. In this case, the diffusion coefficient could be recovered from one dimensional sources evenly spaced and distributed around the two receivers if the sources are spaced closely together and for a large range.

Based on the results from these three cases, the source distribution determines the accuracy of the fit to the GF. A high density of sources relatively near to the receivers

43
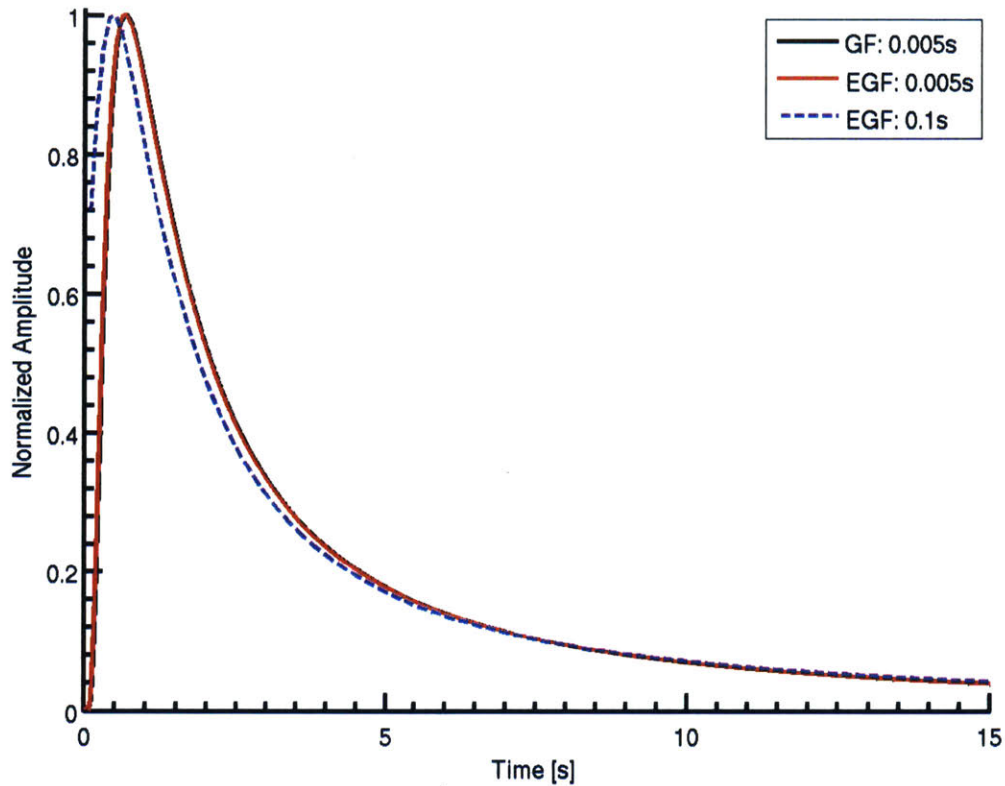
Figure 4-5: Two EGF curves for the 3D Case 3 with different values for time steps that demonstrates the effects of using a larger time step. EGF_{small} has a time step of 0.005 seconds like the curves in the rest of this paper. EGF_{large} has a time step of 0.1 seconds.

as found in Case 1 and Case 3 is necessary to accurately fit the left side of the curve. A large extent of sources as found in Case 2 and Case 3 is necessary in order to match the shape of the tails of the EGF. This agrees with the one dimensional results by Fan and Snieder.[2] The most accurate fits will have both a high density of sources near the receiver and a large extent of sources. The diffusion coefficients that are recovered in this model from such accurately fitted curves are also very accurate.

4.2.2   3D Model

Three cases for the idealized three dimensional model in Section 3.1.2 demonstrate the three dimensional dependence of the estimated diffusion coefficient on the source

44
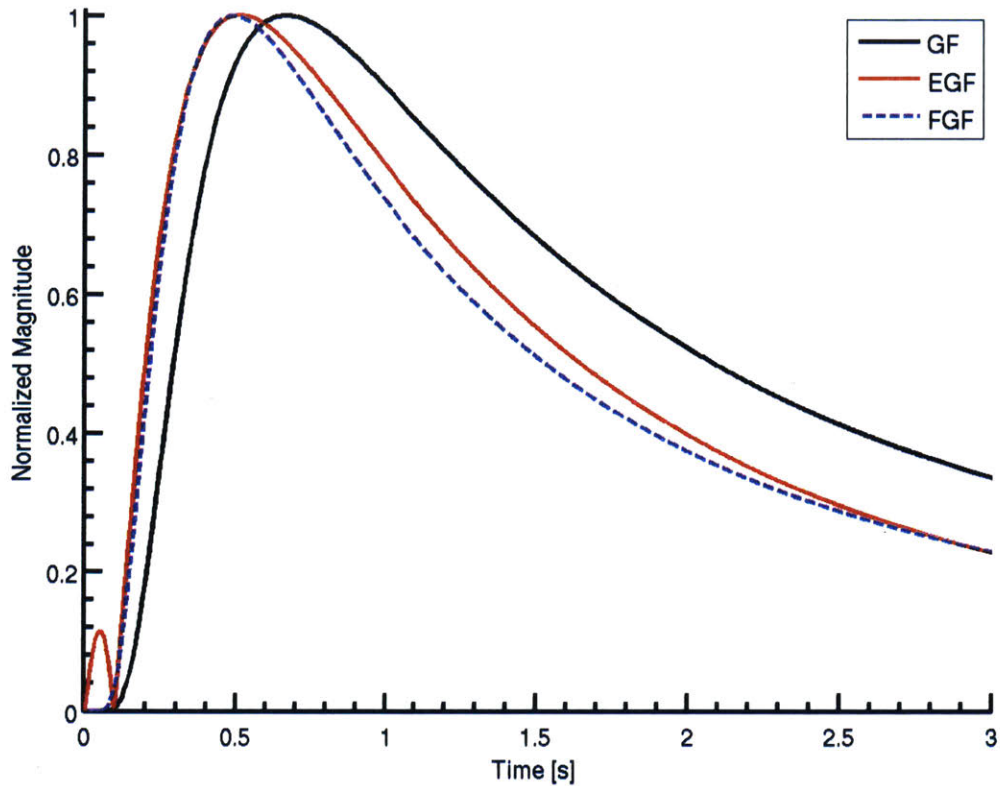
Figure 4-6: This is the first three seconds of Case 3 for the idealized three dimensional model for 0 to 3 seconds. The feature from 0 to 0.2 seconds is an artifact of the processing as discussed in Section 4.1.

distribution. The first case has a small range and a high density of sources. The second case has a large range and a low density of sources. The third case has a large range and a high density of sources. These different source distributions strongly affect the accuracy of the EGF and FGF. The summarized three dimensional results can be seen in Table 4.2

Case 1 had a high density of sources over a small extent. The curves for this example are given in Figure 3-4. This distribution gave a much more reasonable response than the similar distribution than the 1D Case 1 as this case only had an error of 36.4%.

Case 2, shown in Figure 3-5, demonstrates the effects of a distribution with a large range and a small density. The FGF matches the EGF, but since the EGF does not

1D Model	L [km]	ρ [km^{-1}]	D_o [km^2/s]	D_{est} [km^2/s]	l_o [km]	l_{est} [km]	Percent Error
Case 1	14	1	3	14.45	1	4.82	381.7%
Case 2	38	0.5	3	3.41	1	1.14	13.7%
Case 3	38	1	3	3.09	1	1.03	3.0%

Table 4.1: The parameters and error of the estimated diffusion coefficient for each idealized 1D case using a time step of 0.005 seconds and a maximum time of 400 seconds.

3D Model	L [km]	ρ [km^{-3}]	D_o [km^2/s]	D_{est} [km^2/s]	l_o [km]	l_{est} [km]	Percent Error
Case 1	3	1	1	1.36	1	1.36	36.34%
Case 2	12	0.5	1	0.83	1	0.83	16.60%
Case 3	12	1	1	1.01	1	1.01	1.16%

Table 4.2: The parameters and error of the estimated diffusion coefficient for each idealized 3D case using a time step of 0.005 seconds and a maximum time of 400 seconds.

match the GF well, there is error in the fit. The error for this case is 16.6%. This is a relatively small error in general and it gives a reasonable approximation of the initial diffusion coefficient. However, given the choice of data, the third case would be a much better option.

Case 3 had a very effective distribution of sources that resulted in a very good approximation of the estimated diffusion coefficient, as seen in Figure 3-6. For this distribution, the estimated diffusion coefficient between the two receivers could be extracted using seismic interferometry. Ideally, this is the type of distribution that a real dataset would have in order to accurately extract the diffusion coefficient. This third case was also run with the receivers at the top of the three dimensional grid of sources instead of in the center of the grid. These two cases did not differ significantly from each other, however this is likely due to the fact that the plane containing the two receivers still had many sources in it.

These results demonstrate that, like the one dimensional model, the three dimensional model needs both a high density and a large extent in order to accurately approximate the EGF.

Realistic Model	Time Min [s]	Time Max [s]	$D_\circ$ [km^2/s]	D_{est} [km^2/s]	$l_\circ$ [km]	l_{est} [km]	Percent Error
FGF	0	400	1	2.49	1	2.49	149%
FGF$_{left}$	0.2	1.9	1	3.35	1	3.35	235%
FGF$_{right}$	2.5	15	1	2.28	1	2.28	128%

Table 4.3: The parameters and error of the estimated diffusion coefficient for the realistic case using a time step of 0.005 and a maximum time of 400 seconds.

4.3 Realistic Model

The EGF in Figure 3-7 is not an accurate model of the GF. The FGF matches the EGF at both its peak and edges of this graph. Fitting the left and right sides of the peak of the EGF instead of the entire curve do not give a better value for the diffusion coefficient. The fit of the left side of the peak from 0.2 to 1.9 seconds is shown in Figure 4-7 and has an estimated diffusion coefficient of 3.35 km^2/s and an error of 235%. The fit for the right side in Figure 4-8 for 2.5 to 15 seconds is shown in Figure 4-8 and has an estimated diffusion coefficient of 2.28 km^2/s and an error of 128%. The estimated diffusion coefficient for the full FGF curve was 2.49 km^2/s, so fitting the individual sides does not allow an estimation of the diffusion coefficient to be significantly closer to 1 km^2/s, the value of the diffusion coefficient for the GF. These results are summarized in Table 4.3.

The potential to accurately fit a Green's function to part of the EGF could help extract a more accurate estimated diffusion coefficient from a less than ideal dataset in terms of either its density or its extent. The EGF from a dataset that had a dense number of events relatively close to the receivers is reasonably well fit by the diffusion model for the left side of the peak. The EGF from a dataset that had less densely spaced sources that extended over a greater distance appears to give an accurate estimate of the diffusion coefficient for the right side of the peak of the curve. Targeting the time window to be fit in this manner depending on the dataset could lead to more accurate estimations of the diffusion coefficient in cases where the type of data limits the ability of seismic interferometry to extract the estimated diffusion coefficient from the data.
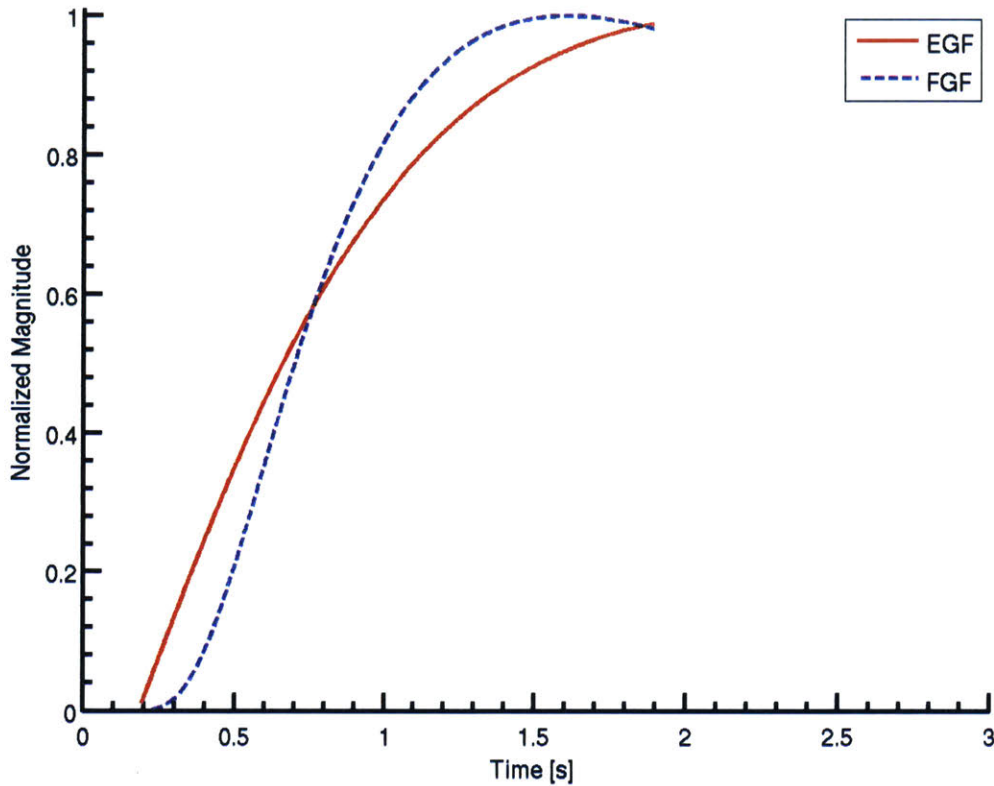
Figure 4-7: The FGF for the left side of the EGF for realistic data. The estimated diffusion coefficient of the FGF is 3.35 km^2/s, an error of 235%.

The difference between the diffusion coefficients for the GF and the FGF is very large at 149%. Using Equation 1.3 and holding the velocity and the number of dimensions constant, the relative mean free path for each value can be calculated. A diffusion coefficient of 2.49 km^2/s is equal to a mean free path of 2.49 km. A diffusion coefficient of 1 km^2/s is equal to a mean free path of 1 km. the mean free path is the average distance between the points at which the signal is scattered. Assuming that these scattering event represent fractures, the GF is predicting that the fractures are 1 km apart and the approximation is predicting that the fractures are 2.49 km apart, an error of 149%.

The range of measured values for g_o ranges from 1^{-6} km^{-1} in the lower mantel to 10^{-3} in the lithosphere to 1 km^{-1} in the case of a volcano.[6, p. 7] The lithosphere
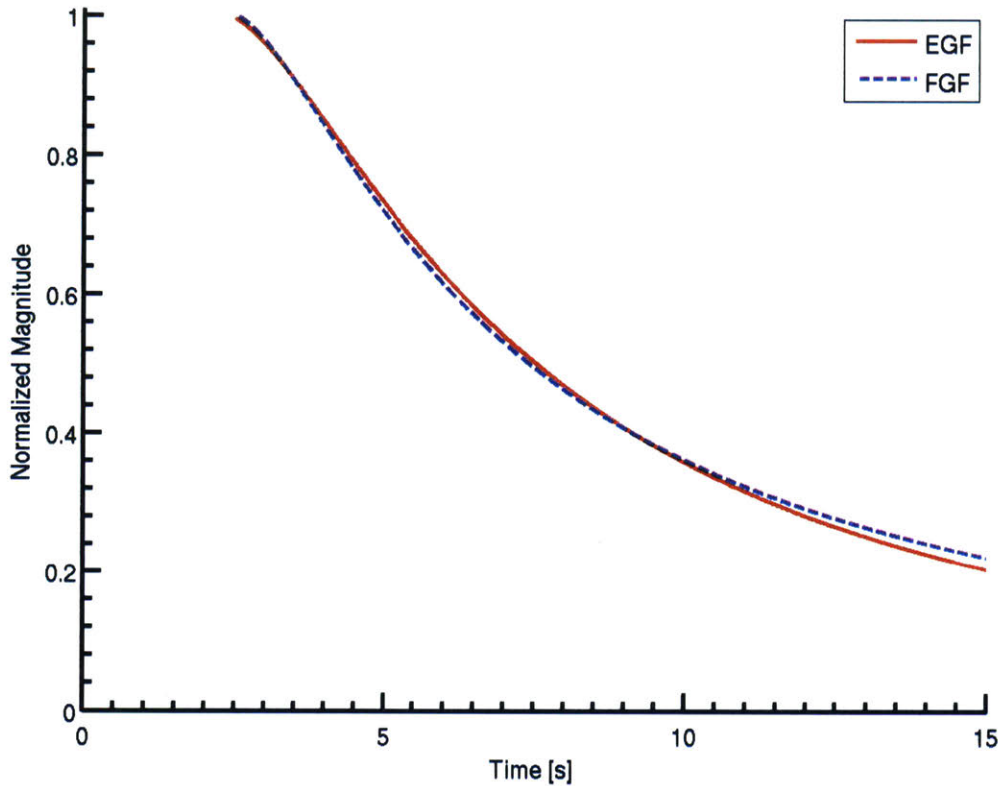
Figure 4-8: The FGF of the right side of the EGF for realistic data. The estimated diffusion coefficient of the FGF is 2.28 km^2/s, an error of 128%.

extends as much as 100 km into the Earth's crust and the mantle is even deeper.[6, p. 1] This is well below the depth of the events from the microseismic dataset in the geothermal field used for this study where the deepest event was at 15 kilometers with most events much shallower. As a result of this depth and relative scale of scattering, these measurements are generally taken at lower frequencies than the frequencies that are important in the geothermal field. The lower frequencies have a longer wavelength which is suitable for resolving images on a larger scale. Higher frequencies have shorter wavelengths and are suitable for resolving images on a finer scale. This higher frequency measurement will detect more scattering events. As a result, the mean free path would be much smaller than it currently is. This would make the initial diffusion coefficient of the GF even smaller and move the GF even

farther away from the FGF and the EGF.

The EGF and the GF with the diffusion coefficient that was examined in Section 3.2 are very different from one another. The approximate peak of the EGF leads the peak of the GF by about three seconds. The GF or the EGF or both curves could be responsible for the error. From the results of the idealized models, the approximation should fit the value of the Green's function between the two receivers, given a source distribution that is both dense and broad enough.

The GF could have the wrong diffusion coefficient. The diffusion coefficient was calculated from Equation 1.3 in Section 2.3 from the values of $c = 3$ km/s, $l = 1$ km, and $d = 3$. If the true value of the velocity of the mean free path is larger, the diffusion coefficient would also be larger. This would bring the exact curve in Figure 3-7 closer to the approximate fit that was calculated from the actual data. The velocity of the acoustic waves underground could be as high as 4-6 km/s rather than the 3 km/s that was used for this paper. If the velocity is actually 6 km/s, then the initial diffusion coefficient would be 2 km^2/s and the estimated diffusion coefficient would only have an error of 49%.

A potentially more accurate approximation of the diffusion coefficient could be obtained by using the multiple lapse-time window analysis (MLTWA). This method uses the amplitude of the early part of the S-wave seismogram that is dominated by the direct S-wave to determine the attenuation of the medium and the amplitude of the coda of the S-wave to approximate the total scattering coefficient.[6] This method is based on using information from real data to find the diffusion coefficient rather than the roughly estimated value that was used for this paper, so the multi-window method would likely give a more accurate result for the diffusion coefficient. The more accurate estimation of the diffusion coefficient would likely move the GF closer to the EGF.

The diffusion coefficient for the GF could also be calculated by using the cross correlation method to find the diffusion coefficient between a source and a receiver. However, this would likely be inaccurate for the same reason that obtaining the fitted diffusion coefficient did not work well.

This method of approximating the diffusion coefficient from microseismic events has potential for application to the real world in the future after further research. The model as it currently stands can give a very limited amount of information from the approximated diffusion coefficient. The diffusion coefficient is not well estimated with the dataset used for this paper, although better datasets may yield better results. The strict source distribution requirements necessary to be able to derive the diffusion coefficient from the distribution of events could make it difficult for this model to be applied in the same manner as this paper. Further study could result in modifications to this model that would make it able to handle more sparse datasets that more accurately reflect a realistic dataset.

4.4   Future Work

The realistic model in this paper could be improved with better information. More sources could be included in the model. The dataset that the source locations were taken from included additional data that was not included in this study because information on the which receivers were active during that time period were not readily available. If this information could be found and included in the model, the increased number and potentially wider distribution of the sources could make the EGF more accurate.

Alternatively, the source distribution could be too biased. Equipartition of energy in the source distribution is important for extracting an accurate diffusion coefficient in models.[12] Given the high density of sources near the two receivers and especially concentrated below the receivers, it might prove worthwhile to either select a certain number of sources per area or weight the contributions of sources from different areas. This could better approximate a random distribution of sources with an equipartition of energy for the receivers.

A different dataset that has a source distribution that is more similar to the source distribution of the third case for the three dimensional model could have more accurate results. The first order difference between the sources that were used for

51

this paper and the sources for the third case of the three dimensional model was the location of the receiver with respect to the sources. The three dimensional model had receivers distributed equally both above and below the two receivers. The receivers for the realistic model were centered much closer to the top of the source distribution. This could account for at least some of the inaccuracy of the EGF.

The relationships between the source distribution and the approximation of the Green's function was only examined qualitatively in this paper. Further work could be focused on creating a quantitative description of the relationships between the source distribution and their affect on the EGF. This would not be a straightforward problem as there are likely several more variables that need to be taken into consideration including angular location of the sources with respect to the receivers, the distribution of sources between the two receivers, and the density and extent of the sources that were examined in this paper. Such quantitative work could be used to determine exactly what this model is calculating. The EGF could be calculated for different initial diffusion coefficients in order to determine if there is a constant factor between the initial diffusion coefficient and the estimated diffusion coefficient or if this method always overestimates the estimated diffusion coefficient or if there is any relationship at all. Determining the source of the artifacts and negating them could also make this method significantly more reliable. This information could be used to make the estimated diffusion coefficient approximate the initial diffusion coefficient more accurately. This knowledge would help the geothermal industry be able to have a better understanding of their fields, which would help improve production and decrease costs. This would help make geothermal energy competitive with energy derived from fossil fuels both on cost and volume.

Bibliography

[1] Jon F. Claerbout. Synthesis of a Layered Medium from its Acoustic Transmission Response. *Geophysics*, 33(2):264–269, 1968.

[2] Yuanzhong Fan and Roel Snieder. Required Source Distribution for Interferometry of Waves and Diffusive Fields. *Geophysical Journal International*, 179(2):1232–1244, 2009.

[3] Pierre Gouedard, Philippe Roux, Michel Campillo, and Arie Verdel. Convergence of the two-point correlation function toward the Green's function in the context of a seismic-prospecting data set. *Geophysics*, 73:V47–V53, December 2008.

[4] J. C. J. Paasschens. Solution of the time-dependent Boltzmann equation. *Phys. Rev. E*, 56:1135–1141, Jul 1997.

[5] James Rickett and Jon Claerbout. Acoustic Daylight Imaging via Spectral Factorization; Helioseismology and Reservoir Monitoring. *The Leading Edge*, 18(8):957–960, 1999.

[6] Haruo Sato, Michael Fehler, and Takuto Maeda. *Seismic Wave Propagation and Scattering in the Heterogeneous Earth*. Springer, second edition, 2012.

[7] Gerard Thomas Schuster. *Seismic interferometry*. Cambridge University Press, Cambridge, 2009.

[8] Roel Snieder. *A Guided Tour of Mathematical Methods for the Physical Sciences*. Cambridge University Press, Cambridge, UK ; New York, second edition, 2004.

[9] Roel Snieder. Retrieving the green's function of the diffusion equation from the response to a random forcing. *Phys. Rev. E*, 74:046620, Oct 2006.

[10] Jefferson W. Tester. *Sustainable Energy: Choosing Among Options*. MIT Press, Cambridge, Mass., 2005.

[11] Jefferson W. Tester. The Future of Geothermal Energy: Impact of Enhanced Geothermal Systems [EGS] on the United States in 21st Century, 2007.

[12] K. van Wijk. On Estimating the Impulse Response Between Receivers in a Controlled Ultrasonic Experiment. *Geophysics*, 71(4):SI79–SI84, 2006.