UNIVERSITY OF MISKOLC

Faculty of Earth Science and Engineering

Department of Geophysics

INVERSION-BASED METHOD DEVELOPMENT TO IMPROVE THE INTERPRETATION OF POTENTIAL FIELD GEOPHYSICAL DATA

PhD THESIS

By

MAHMOUD IBRAHIM MAHMOUD ABDELAZIZ

Scientific Supervisors:

Prof. Dr. Mihály Dobróka

Dr. Endre Turai

MIKOVINY SÁMUEL DOCTORAL SCHOOL OF EARTH SCIENCES

Head of the Doctoral School: Prof. Dr. Péter Szűcs

Miskolc, 2022

HUNGARY

ADVISOR'S FOREWORD

for the (PhD) thesis

"INVERSION-BASED METHOD DEVELOPMENT TO IMPROVE THE INTERPRETATION OF POTENTIAL FIELD GEOPHYSICAL DATA"

by

Mahmoud Ibrahim Mahmoud Abdelaziz

The topic of the Candidate's thesis - inversion based geophysical data processing - is in focus of international research. The new method developments introduced by the Candidate in the thesis belong to the range of modern data processing tools of applied geophysics.

In accordance with the main objectives of the thesis, the Candidate investigated the inversion-based Fourier transformation method on synthetic and in-situ measured datasets in three distinct categories:

- reducing the outlier sensitivity by applying the inversion-based Fourier transformation to the synthetic 1D wavelet and 2D magnetic and gravity datasets,
- analyzing the inversion approach in processing the non-regularly sampled complete datasets in 1D and 2D,
- analyzing the incomplete sampling problem at various degrees of missing data.

The Candidate completed his PhD doctoral research with a high degree of independence and precision and reached scientifically valuable new results. His continuous efforts towards scientific research, his creativity, and the results presented in this thesis prove the scientific knowledge and the suitability of the Candidate for independent research. In our opinion, the Candidate's results are worth to be published in ranked international journals of applied geophysics. We certify that this dissertation contains only valid data and the presented results are representing the Candidate's own work. In our opinion, it is fully adequate in scope and quality required by the Mikoviny Sámuel Doctoral School of Earth Sciences. Based on the above, we support and recommend the public defence of the thesis and the award of the PhD title.

22/09/2022, Miskolc

Dr. Endre Turai

University private professor

Dr. Mihály Dobróka

professor emeritus

LIST OF CONTENTS

Chapter 1
INTRODUCTION

It is generally recognized that geophysical measurements are frequently hampered by a variety of geological, geophysical, and environmental conditions, all of which cause the data to be contaminated with varying degrees of noise. Also, some of these noises are artificially created during the field data acquisition, such as those associated with instrumental component industrial noise, while others are naturally controlled due to diurnal variation effects on the potential fields. These noises strongly affect the quality of the geophysical outputs and have a significant impact on how the measured geophysical data is processed and interpreted. Therefore, the raw or field observations must be effectively processed before they can be used to interpret the subsurface variations. Data processing is a critical stage between the geophysical data gathering in the field and analyzing it in order to solve the noise or error prevalence problem accompanied by the field data measurements and attenuate the negative consequences of a noisy signal. Several sources and ways that deal with geophysical data processing have been thoroughly studied in the literature (Claerbout, 1985; Reynolds, 1997; Kearey et al, 2002; Milsom, 2003; Bevan, 2010; Hinze et al, 2013). Even though noise levels are expected to decrease further in the future as survey design and data processing methods improve, contemporary geophysical measurements frequently contain high enough amounts of noise to impede qualitative and quantitative interpretations. This motivates us to look into noise-reduction strategies and see how effective they are at removing unwanted noise from collected geophysical datasets.

In geophysical exploration, the data is sometimes incomplete or sporadically sampled during data collecting due to physical and budgetary constraints. Obstacles, no-permit zones, and hardware issues with the equipment employed are all factors to consider. It raises a more complicated or an ill-posed problem of the missing data either the measuring points are taken over uniform (regular) or random (non-regular) grids. Missing datasets means that there are large substantial omissions in relevant data in the areas under investigation and hence, aliasing might occur, resulting in low resolution. As a result, we are directing our emphasis to the recovery of the missing data problem. A more tremendous trend in recent years has been the adoption of various processing strategies for handling incomplete sampling designs (Sacchi et al, 1998; Trad, 2009; Ma, 2013; Chai et al, 2020).

In accordance with the above-mentioned concerns, we introduce the Fourier transform as an effective signal processing tool that may be designed to boost both 1D wavelets and 2D potential magnetic and gravity datasets. Among numerous geophysical data processing methods, the Fourier transform plays a vital role in signal processing to enhance the quality of measured datasets that contain simple or complex noises during the field survey. We shall encounter the Fourier transform as a filtering technique for improving the performance of the 1D and 2D geophysical data outputs. The discrete Fourier transformation is used to convert discrete noiseless or noisy time-domain datasets into a frequency-dependent representation. The discrete Fourier transform has a wide range of applications including polynomial multiplication estimate, acquiring a large number of objects via radar echoes, correlation, and spectral analysis, as well as determining the signal's frequency spectrum, which is our work's primary geophysical purpose. Moreover, the Fourier Transform enables us to evaluate waveforms uniquely and powerfully, such as sound waves, electromagnetic fields, elevation from hill to location, and so on.

In fact, the traditional Fourier transform (DFT-FT) based on simple Least square method can give optimal solution in the case of Gaussian noisy data, while it is not effective enough when dealing with more complicated noisy data containing outliers (e.g. Cauchy noise-contaminated data). In addition, the traditional DFT approach is not applicable and inefficient causing several shortcomings when analyzing incomplete or missing datasets. These are the main reasons that prompt us to use the deviation vector of a weighted norm to be minimized through the inversion method of iteratively reweighted least-squares Fourier transformation (2D IRLS-FT) (Dobróka et al. 2016). In that regard, the introduced inversion algorithm proceeds to datasets sampled in both 1D and 2D. Accordingly, we can summarize the objectives of this thesis as follows: (1) reducing the outlier sensitivity using the IRLS inversion-based FT compared to the traditional discrete Fourier transform (DFT), (2) applying the inversion algorithm for handling the non-regular sampling problem, and (3) processing of the incomplete or missing datasets sampled both equidistantly and non-equidistantly utilizing our newly developed inversion approach. For noise reduction, Dobróka et al (2015) handled a one-dimensional (1D) inversion-based Fourier transformation (S-IRLS-FT) method which was developed and generalized to two-dimensional (2D) inversion, providing accurate and efficient results in the field of the reduction to the pole of the magnetic datasets (Dobróka et al 2017). The inversion method's efficacy, precision, and noise reduction capacity were all proved by the results. In the framework of this inversion-based Fourier transformation method, the continuous

Fourier spectrum is discretized using the series expansion method to solve our over-determined inverse problem in the form of the expansion coefficients. The series expansion-based inversion methods were effectively employed for the interpretation of borehole geophysical data (Szabó, 2011 and 2015) as well as processing of induced polarization (IP) datasets (Turai and Dobróka, 2011). Moreover, where integration is necessary, the elements of the Jacobian matrix are slowly estimated and hence, require more time to calculate. This problem can be solved by using the Hermite functions as basis functions making use of the favor that they are eigenfunctions of the Fourier transformation which allow quick and accurate determination of the elements of the Jacobian matrix. To make the inversion algorithm more robust and resistant, the Cauchy weights (Amundsen, 1991) are often integrated into an Iteratively Reweighted Least Square (IRLS) algorithm (Scales et al, 1988), but in this case, the computation of the scale parameters is problematic because they should be known in advance. As a result, the most frequent value (MFV) method (Steiner, 1988 and 1997) is used to handle this problem by iteratively calculating the Cauchy Steiner weights through an internal iteration loop instead of Cauchy weights in a manner that minimizes data loss. As proved by Dobróka et al (1991), combining the algorithm of Iteratively Reweighted Least Square (IRLS) with Steiner weights given by the most frequent value (MFV) method is a particularly beneficial operation. Inserting the weights into the IRLS inversion process leads each data point's error margin to contribute to the solution resulting in a more accurate, stable, and robust inversion outcome as well as successful and effective outliers reduction. The applications of the most frequent value (MFV) have been introduced by many authors demonstrating numerous benefits using least-squares or other traditional statistical procedures (Ferenczy et al. 1990; Steiner and Hajagos 1994; Szucs and Civan 1996; Szucs et al, 2006; Szegedi and Dobróka, 2014; Zhang, 2017).

In this study, MATLAB-based codes for the proposed methods of conventional discrete Fourier transform (2D DFT) and inversion-based Fourier transformation (2D IRLS-FT) are implemented on synthetic 1D wavelets, synthetic 2D magnetic and gravity datasets, and real gravity field data measurements. First, the outlier sensitivity rejection for 1D data sampled equidistantly is assessed using both the DFT and IRLS-FT methods. Moreover, solving the non-regular sampling problem by our inversion method is evaluated on 1D complete datasets sampled randomly as well as random walk measurement positioning. In addition, a newly developed inversion-based Fourier transformation algorithm is performed on 1D incomplete or missing datasets sampled both regularly and non-regularly, and it is effectively extended to the incomplete block sampling designs. On the other hand, for noise rejection analyzing purposes,

equidistantly sampled 2D noise-free and noisy synthetic magnetic datasets are subjected to the above-mentioned noise filtering-based 2D Fourier transformation techniques in the framework of reduction to the pole. The inversion method is also tested on the non-regular complete magnetic datasets sampled with the same two configurations as 1D data. We then evaluate the incomplete sampling problem of the 2D noise-free and noisy magnetic data taken over both regular and random-walk sampling grids. Furthermore, both the traditional DFT and IRLS inversion-based FT methods are tested and evaluated on 2D noisy and noise-free synthetic gravitational datasets generated using a model of a right rectangular prism. The model is initially applied to a regular grid before being randomized to produce irregularly sampled gravity intervals. Similar to the magnetic methodology, the DFT method is only applied to gravitational datasets equidistantly sampled, as it is ineffective for irregular intervals, and the results are then compared to those found by the inversion algorithm (2D IRLS-FT) applied to datasets sampled both equidistantly and non-equidistantly. New geophysical applicability in the field of low-pass filtering of such 2D gravitational datasets is successfully introduced in the present investigation. Besides, the 2D incomplete noise-free and noisy gravity data are then comprehensively subjected to the newly proposed inversion-based Fourier transformation algorithms on an equidistantly and non-equidistantly sampled basis. Finally, a field example is introduced as a case study in the western central part of Sinai Peninsula, Egypt to effectively evaluate the noise sensitivity reduction of the low-pass filtering-based Fourier transformation on real equidistantly and non-equidistantly sampled gravitational field datasets. Such real field gravity datasets measured over both equidistant and non-equidistant grids are also involved to assess the incomplete or missing sampling problem using our newly developed inversion-based low-pass filtering. The results of 1D and 2D synthetic datasets, as well as the gravity field measurements, indicated the newly developed inversion method's efficiency, resistance, and noise rejection capacity in both the space and frequency domains rolling out a new and cost-effective means of coping with incomplete sample issues that may arise during the collection of a wide range of field data.

Chapter 2
THE GEOPHYSICAL INVERSION METHODS

Geophysics' purpose is to use observations of physical phenomena to determine the characteristics of the interior of the earth from the surface and/or boreholes. Thus, the manner how determining the interested targets of the subsurface structure and rock properties from the surface geophysical measurements is a great challenge for all geophysicists. Deriving the earth's inner model from the surface data collected during field acquisition through some quantitative tools applications is the aim of the inverse problem. Before doing it the selection of a suitable geophysical model, and the specification of its parameters is necessary in formulating a forward problem. It's relationship can be used to calculate theoretical data for the assumed earth model. *Figure 1* depicts two charts highlighting the distinctions between forward and inverse geophysical problems.

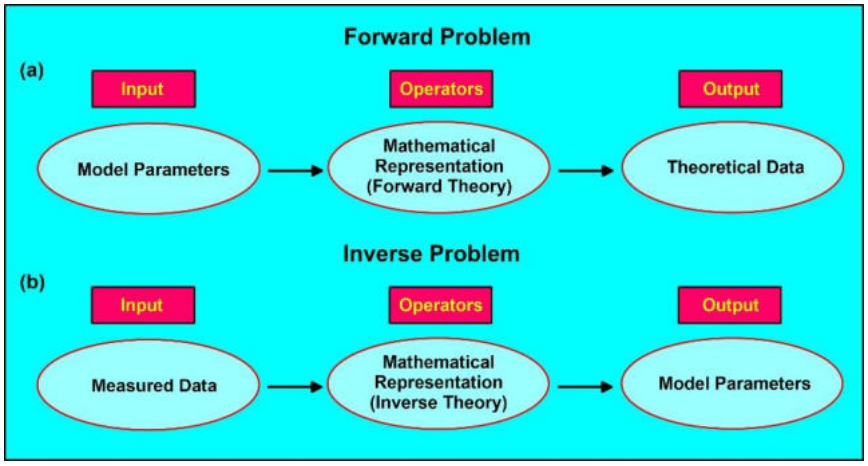

Figure 1. *a) Forward problem: synthetic data computation using known input model parameters, b) Inverse problem: model parameters estimation from the observational field datasets.*

The geophysical inversion methods have been thoroughly explained in many kinds of literature (Menke, 1984; Tarantola 1987; Meju, 1994; Oldenburg and Yaoguo Li, 2005; Sen and Stoffa, 2013; Menke, 2018). The geophysical inversion methods can be categorized into two main items according to (Sen and Stoffa, 2013). The first refers to the direct inversion methods in which the model parameters can be retrieved directly from the data observed. On the other hand, model-based inversion approaches are of particular importance since they compare theoretical data for a supposed model with observed ones. In our study, the following section introduces and briefly discusses some local inversion methods.

2.1. THE STRAIGHTFORWARD INTERPRETATION METHODS

It is important to note that the terminology ‘straightforward’ or ‘direct’ is commonly dedicated to some inversion approaches because of the direct derivation of the earth's model parameters from the data. In these methods, the values of the model parameters should be initially set such as velocity, density, etc, however, the model's variability in one or two-dimensional domains must be taken into our consideration. Furthermore, the physics of the forward problem is used to create direct inversion approaches (Sen and Stoffa, 2013). Because of their speed and ease of implementation, direct inversion methods play an important role in a variety of geophysical applications. On the other hand, these methods are disadvantageous and not the optimal solution in some cases when compared to other inversion methods due to their inaccuracy and limited reconstruction (Kazufumi and Bangti, 2014). Several geophysical applications of these inversion approaches, particularly in the context of seismology, have been explained through much literature (Fuchs and Muller 1971; Yagle and Levy, 1985; Singh et al. 1989). It is well known that the direct- or operator-based inversion methods are highly sensitive and affected by the data that is incomplete and mostly contaminated with noise. The main reason since it's not always evident how to condition the operator to minimize the negative impacts of applying these operators to field data. This problem can be effectively solved by employing the data uncertainties in the inversion procedure as proven by the model-based inversion methods (Sen and Stoffa, 2013).

2.2. THE MODEL-BASED INVERSION METHODS

In the model-based inversion approach, the model is regarded as the solution to the inverse problem if the match between observed and synthetic data created for an assumed model is satisfactory. On the other hand, when the misfit between the measured and calculated data is high, the actual values of the model parameters are modified in such a way to agree with the real field datasets. In that case, the synthetic data are recalculated and evaluated to the observations once more. This procedure is progressively continued and repeated over several iteration steps until an acceptable fit is achieved between the predictions and observed data. The matched criteria in the last iteration step refer to the most promising model that simulates the real subsurface geological structure and hence it can be accepted as the inverse problem solution. Therefore, the inversion can be thought of as an optimization process aimed at finding the optimum model that effectively describes the observations. The algorithm of the inversion process can be mathematically started by considering a model of the earth's interior whose parameters are set as a column vector as follows:

$$\vec{m} = \{m_1, m_2, m_3, \ldots\ldots, m_M\}^T, \quad (1)$$

where M is the total number of model parameters and T is a matrix transpose. Forward computation with the above-mentioned column vector of model parameters m is used to produce synthetic data $\vec{d}_{syn}$ through the so-called response function as:

$$\vec{d}_{syn} = \vec{g}(\vec{m}), \quad (2)$$

where $\vec{g}$ is the forward modeling operator which is a non-linear operator in most geophysical problems. In that case, the synthetic or theoretical data can be estimated in a discrete form of N dimensional column vector as follows:

$$\vec{d}_{syn} = \left\{d_1^{(syn)}, d_2^{(syn)}, d_3^{(syn)}, \ldots\ldots, d_N^{(syn)}\right\}^T. \quad (3)$$

Such synthetic datasets are then compared to those measured during the process of field geophysical survey $\vec{d}_{meas}$ at N number of ground stations:

$$\vec{d}_{meas} = \left\{d_1^{(meas)}, d_2^{(meas)}, d_3^{(meas)} \ldots\ldots, d_N^{(meas)}\right\}^T. \quad (4)$$

The measure of acceptability of an earth model illustrated in Eq. 2 can be evaluated by estimating the misfit between the measured and synthetic datasets. This can be performed with the help of objective function or energy function as follows:

$$\vec{e} = \vec{d}_{meas} - \vec{d}_{syn} = \vec{d}_{meas} - \vec{g}(\vec{m}). \quad (5)$$

Regarding L_p norm (Menke 1984), the error function can be introduced as:

$$\|e\|_p = [\sum_{i=1}^{s} |e_i|^p]^{1/P}, \quad (6)$$

while in the common case of L_2 norm, it can be defined as:

$$\|e\|_2 = [\sum_{i=1}^{s} |e_i|^2]^{1/2}, \quad (7)$$

or in vectorial form as:

$$\|e\|_2 = \left[\left(\vec{d}_{meas} - \vec{g}(\vec{m})\right)^T \left(\vec{d}_{meas} - \vec{g}(\vec{m})\right)\right]^{1/2}. \quad (8)$$

L_2 norm is widely used in most geophysical applications while in some cases L_1 norm can be applied as introduced by (Claerbout and Muir 1973). The most commonly model-based inversion methods applied for geophysical problems are the linear or linearized inversion optimization approaches. Therefore, this study will briefly shed light on some of the linear inversion approaches that are also essential to reaching our goal.

2.2.1. THE LINEAR INVERSION OPTIMIZATION APPROACHES

Among all geophysical approaches, linearized inversion methods are the most well-known in the last decades which assign the solution to a local optimum of the objective function. The major assumption of these linear inversion methods is based on the fact that the model parameters have a linear impact on the data. It's worth noticing that by inserting a matrix $\underline{\underline{G}}$ instead of the forward modeling vector-vector function $\vec{g}$ stated above in Eq. 2, the linearization processes significantly simplify the response function in the following manner:

$$\vec{d} = \underline{\underline{G}}\, \vec{m}. \tag{9}$$

The data-model relationship can be successfully linearized if specific conditions are met, which vary from one geophysical approach to another. The strategy of the linearization process can be extremely understood by considering small disturbances $\delta\vec{m}$ (model correction vector) added to the reference initial model $\vec{m}_0$ to compute the observed datasets $\vec{d}_{obs}$ using the forward modeling operator $\vec{g}$ as follows:

$$\vec{d}_{obs} = \vec{g}\,(\vec{m}_0 + \delta\vec{m}), \tag{10}$$

while the synthetic data can be defined as:

$$\vec{d}_{syn} = \vec{g}\,(\vec{m}_0). \tag{11}$$

In the first-order term, the observed data obtained in Eq. 10 can be distributed by the Taylor series expansion method concerning the start model $\vec{m}_0$ as:

$$d_k\,(\vec{m}_0 + \delta\vec{m}) = g_k\,(\vec{m}_0) + \sum_{j=1}^{M}\left(\frac{\partial_{gk}}{\partial_{mj}}\right)_{\vec{m}_0} \delta m_j, \tag{12}$$

where $k = 1, 2, \ldots, N$ and $\left(\frac{\partial_{gk}}{\partial_{mj}}\right)_{\vec{m}_0}$ is the jacobian matrix that constitutes the partial derivatives of synthetic data concerning model parameters. In that case, the observed data can be given as:

$$\vec{d}_k = \vec{d}_k^{(0)} + \sum_{j=1}^{M}\left(\frac{\partial_{gk}}{\partial_{mj}}\right)_{\vec{m}_0} \delta m_j. \tag{13}$$

By simplifying the jacobian matrix as $\underline{\underline{G}}$ and $\delta\vec{d}$ as the difference between the measured and synthetic data

$$\delta\vec{d} = \underline{\underline{G}}\, \delta\vec{m}, \tag{14}$$

we can now formulate a linear expression for the data and model parameters disturbances. Referring to Eq. 5, the deviation vector can be defined as:

$$\vec{e} = \vec{d}_{obs} - \vec{d}_{syn} - \underline{\underline{G}}\,\delta\vec{m}, \tag{15}$$

and it can be simply written with the linearized inverse problem as:

$$\vec{e} = \vec{d}_{meas} - \underline{\underline{G}}\,\delta\vec{m}, \tag{16}$$

where $\delta \mathrm{d}_{\mathrm{meas}} = \vec{d}_{meas} - \vec{d}_{syn}$. Over several iterations, the response function is linearized subsequently where the initial model $\vec{m}_0$ is replaced by a new model m in each iteration step until a stop criterion is met. Finally, the inverse problem can be solved by inverting $\underline{\underline{G}}$ and hence, we can estimate an update to the model vector as follows:

$$\vec{m} = \underline{\underline{G}}^{-1}\,\vec{d}, \tag{17}$$

where $\underline{\underline{G}}^{-1}$ refers to the inverse matrix. One of the most important aspects when dealing with the geophysical inversion methods is to accurately determine the type of inverse problem that has to be solved. Depending on the relationship between the number of data and model parameters, the inverse problem can be classified into three main categories. When the number of field data points exceeds that of the model parameters, the inverse problem is said to be over-determined. On the other hand, the inverse problem can be realized as underdetermined if the model has more parameters than the number of observed data. Combing the characteristics of both inversions leads to the mixed-determined inverse problem. Several inversion methods have been introduced by many authors to solve the above-mentioned inverse problems such as (Backus and Gilbert, 1967; Backus,1970; Parker, 1977; Jackson, 1979).

2.2.1.1. Solution to the Inverse Problem with an Over-determination

As described above, the inverse problem can be recognized with an over-determination aspect when we have enough data (*N*) collected during the field geophysical survey compared to the number of unknown model parameters (i.e. $N > M$). In such a situation, all model parameters are represented in the data and hence the observed data points should be located on a straight line to be effectively fitted by the equations in a linear system. In that sense, the methods of least squares are the optimum procedures for obtaining the most suitable fit to the data with the smallest non-zero error value (Sen and Stoffa, 2013).

2.2.1.1.1. The Gaussian Least Squares (LSQ)

The Gaussian Least Squares can be considered as one of the most significant linear inversion methods based on minimizing the L_2 norm of the deviation vector. Gauss (1809) was the first to introduce and formulate the mathematical expressions of this inversion method for solving the over-determined inverse problem. The objective function E that has to be minimized for estimating the misfit between the measured and theoretical data can be given as:

$$E = \sum_{k=1}^{N} e_k^2 = \sum_{k=1}^{N} (d_k - \sum_{i=1}^{M} G_{ki}\ m_i)\left(d_k - \sum_{j=1}^{M} G_{kj}\ m_j\right), \tag{18}$$

which can be then differentiated in terms of the model parameters with derivatives equal to zero to get it's minimal $\frac{\partial E}{\partial m_l} = 0$ ($l=1,...,M$). The differentiation process is applied separately to all terms of Eq. 18 resulting in an expression as follows:

$$\sum_{i=1}^{M} m_i\ \sum_{k=1}^{N} G_{ki}\ G_{kl} = \sum_{k=1}^{N} d_k G_{kl}, \tag{19}$$

which can be written in a vectorial form as:

$$\underline{\underline{G}}^T \underline{\underline{G}}\ \vec{m} = \underline{\underline{G}}^T\ \vec{d}. \tag{20}$$

By multiplying both sides by $\left(\underline{\underline{G}}^T \underline{\underline{G}}\right)^{-1}$ and introducing generalized inverse matrix $\underline{\underline{G}}^{-g} = \left(\underline{\underline{G}}^T \underline{\underline{G}}\right)^{-1} \underline{\underline{G}}^T$, the least-squares estimate of the model can be defined as:

$$\vec{m} = \underline{\underline{G}}^{-g}\ \vec{d}. \tag{21}$$

Moreover, the matrix $\left(\underline{\underline{G}}^T \underline{\underline{G}}\right)^{-1} \underline{\underline{G}}^T$ is used to estimate the model parameters through inverting the linear equation system. It is important to note that the elements of a matrix $\underline{\underline{G}}$ are consisting of $(N \times M)$, while those for the symmetric square matrix $\underline{\underline{G}}^T \underline{\underline{G}}$ are $(M \times M)$. Depending on the matrix $\left(\underline{\underline{G}}^T \underline{\underline{G}}\right)^{-1}$, the solution of the least-squares can be successfully found taking into our consideration that the model parameters should be represented by several data vector information.

2.2.1.1.2. The Weighted Gaussian Least Squares (WLSQ)

It is commonly known that field survey data is contaminated with a lot of noises or inaccuracies, resulting in data uncertainty. These uncertainties of the measured data can be handled by inserting a weighting matrix $\underline{\underline{W}}$ into the inversion algorithm. The data's weights are provided in its major diagonal with quantity commensurate to the degree of uncertainty.

Because of the weighting procedure, noisy data with large errors are incorporated into the inversion with less weight, which is extremely advantageous, especially when dealing with outliers. In that regard, the misfit between the measured and calculated data can be minimized in a weighted norm of the deviation vector as:

$$E = \sum_{k=1}^{N}(d_k - \sum_{i=1}^{M} G_{ki}\ m_i) \sum_{s=1}^{N} W_{ks} \left(d_s - \sum_{j=1}^{M} G_{sj}\ m_j\right). \quad (22)$$

In a similar manner to the Gaussian Least Squares method, the derivatives concerning the model parameters using the property of $\frac{\partial E}{\partial m_l} = 0$ can be estimated as:

$$\sum_{j=1}^{M} m_j \sum_{k=1}^{N} G_{kl} \sum_{s=1}^{N} W_{sk}\ G_{sl} = \sum_{k=1}^{N} G_{kl} \sum_{s=1}^{N} W_{ks}\ d_s, \quad (23)$$

and in a vectorial form as:

$$\underline{\underline{G}}^T\ \underline{\underline{W}}\ \underline{\underline{G}}\ \vec{m} = \underline{\underline{G}}^T\ \underline{\underline{W}}\ \vec{d}. \quad (24)$$

In that case, the weighted least-squares approach is solved using the generalized inverse matrix $\underline{\underline{G}}^{-g} = \left(\underline{\underline{G}}^T\ \underline{\underline{W}}\ \underline{\underline{G}}\right)^{-1}\ \underline{\underline{G}}^T\ \underline{\underline{W}}$ to give the model vector as follows:

$$\vec{m} = \underline{\underline{G}}^{-g}\ \vec{d}. \quad (25)$$

2.2.1.1.3. The Iteratively Reweighted Least Squares (IRLS)

The Iteratively Reweighted Least Squares (IRLS) can be considered as one of the most powerful optimization inversion methods based on solving the L_p approximation problem. In this method, the weighting matrix is recalculated iteratively over several steps. More importantly, by increasing the iterative procedures, the deviation between the measured and predicted data becomes lower and hence, the noisy datasets have less contribution in the linear inverse problem solution. The IRLS approach is effectively implemented to minimize the objective function of L_p norm as follows:

$$E_p = \sum_{k=1}^{N} |e_k|^p = \sum_{k=1}^{N} \left|d_k - \sum_{i=1}^{M} G_{ki}\ m_i\right|^p. \quad (26)$$

Using $\frac{\partial \mathrm{E}}{\partial \mathrm{m_l}} = 0$, the above-mentioned formula is differentiated to the model parameters to get the derivatives as:

$$\sum_{\mathrm{j=1}}^{\mathrm{M}} \mathrm{m_j} \sum_{\mathrm{k=1}}^{\mathrm{N}} \mathrm{G_{ki}}\ \mathrm{W_{kk}} \mathrm{G_{kl}} = \sum_{\mathrm{k=1}}^{\mathrm{N}} \mathrm{G_{kl}}\ \mathrm{W_{kk}} \mathrm{d_k}, \quad (27)$$

with the same vectorial form stated above in Eq. 24 as:

$$\underline{\underline{G}}^T\ \underline{\underline{W}}\ \underline{\underline{G}}\ \vec{m} = \underline{\underline{G}}^T\ \underline{\underline{W}}\ \vec{d}. \quad (28)$$

When p equals to 2 in the first step of the IRLS method, the weights (W_{kk}) are assigned to an identity matrix ($\underline{I}$) resulting in the Gaussian least-squares model parameters vector as follows:

$$\vec{m}^{(1)} = \left(\underline{G}^T \underline{G}\right)^{-1} \underline{G}^T \vec{d}, \tag{29}$$

which is then used to estimate the deviation vector $e_k^{(1)} = d_k - \sum_{i=1}^{M} G_{ki}\ m_i^{(1)}$ and a weighting matrix $W_{kk}^{(1)} = |e_k|^{p-2}$. In the second step, the model vector can be calculated as:

$$\vec{m}^{(2)} = \left(\underline{G}^T \underline{W}^{(1)} \underline{G}\right)^{-1} \underline{G}^T \underline{W}^{(1)} \vec{d}. \tag{30}$$

This process is continued over several q-th iteration steps according to the following equation:

$$\underline{G}^T \underline{W}^{(q-1)} \underline{G}\, \vec{m}^{(q)} = \underline{G}^T \underline{W}^{(q-1)} \vec{d}. \tag{31}$$

Furthermore, the method of Least Absolute Deviations (LAD) which depends mainly on the minimization of L_1 norm ($p = 1$) numerically proceeds in most of the geophysical application problems (Scales et al. 1988). Although its wide extent of applicability, the Least Absolute Deviations method can practically give more accurate results only when the data is contaminated by a lesser number of major errors.

2.2.1.2. Solution to the Inverse Problem with an Under-determination

Sometimes the field data measured in the field are not enough to give a clear comprehensive picture of the earth's subsurface geology. This situation causes an increase in the number of model parameters relative to the observed datasets (i.e. $N < M$) making the inverse problem to be under-determined (Aki and Richards, 1980; Menke, 1984; Tarantola, 1987). In this case, there are an infinite number of solutions to the inverse problem that satisfy the experimental data especially when the observations are contaminated with errors (Meju, 1994). Thus, several attempts are exerted to find a unique or most feasible solution among all equivalent ones, and therefore to overcome the ill-posed problem of underdetermination. One of these ways is to reduce the number of model parameters by discretizing the continuously earth's model into multiple separate ones, to achieve the desired goal of overdetermination. Moreover, this can be solved by adding some additional constraints or prior information to the problem that is introduced independently of the data d such as the previously studied geological, geophysical, and/or borehole data (Jackson, 1979). Mathematically, a solution to the under-determined inverse problem can be simply introduced by minimizing the objective function characterizing

the Euclidean norm of the model vector ($L = \|\vec{m}\|$) together with the Lagrange multipliers $\vec{\lambda}$ as follows:

$$E = L + \sum_{k=1}^{N} \lambda_k \, e_k = \sum_{j=1}^{M} m_j^2 + \sum_{k=1}^{N} \lambda_k \, (d_k - \sum_{i=1}^{M} G_{ki} \, m_i), \tag{32}$$

where λ_k is the k-th multipliers. By applying $\frac{\partial E}{\partial m_l} = 0$, the model parameter estimates can be defined as:

$$m_l = \frac{1}{2} \sum_{k=1}^{N} \lambda_k \, G_{kl}, \tag{33}$$

which is simplified in a vectorial form as:

$$\vec{m} = \frac{1}{2} \underline{\underline{G}}^T \, \vec{\lambda}, \tag{34}$$

this formula in conjunction with the data vector $\vec{d} = \frac{1}{2} \, \underline{\underline{G}} \, \underline{\underline{G}}^T \, \vec{\lambda}$ can help drive the multipliers vector as follows:

$$\vec{\lambda} = 2 \left(\underline{\underline{G}} \, \underline{\underline{G}}^T \right)^{-1} \vec{d}, \tag{35}$$

resulting in the following model vector as a solution to the under-determined inverse problem

$$\vec{m} = \underline{\underline{G}}^T \left(\underline{\underline{G}} \, \underline{\underline{G}}^T \right)^{-1} \vec{d}, \tag{36}$$

which can be simply written using the generalized inverse matrix $\underline{\underline{G}}^{-g} = \left(\underline{\underline{G}}^T \underline{\underline{G}} \right)^{-1} \underline{\underline{G}}^T$ as:

$$\vec{m} = \underline{\underline{G}}^{-g} \, \vec{d}. \tag{37}$$

Additionally, one of the most important aspects when dealing with the under-determined inverse problem is to improve the relationship between the model and data errors. This can be effectively accomplished by providing extra constraints to the model space, such as the weighting matrix, like the over-determined inverse problem stated above. This is since we occasionally obtain valuable knowledge about some model parameters which should remain constant or do not change much during the inversion procedure, as opposed to other mysterious model parameters that must be varied or modified from the initial ones (Sen and Stoffa, 2013). Therefore, the weighted norm of the deviation vector is minimized, and a weighting matrix $\underline{\underline{W}}$ is inserted to give a model vector $\vec{m} = \frac{1}{2} \underline{\underline{G}}^T \underline{\underline{W}}^{-1} \vec{\lambda}$ and multipliers vector $\vec{\lambda} = 2 \left(\underline{\underline{G}} \, \underline{\underline{W}}^{-1} \underline{\underline{G}}^T \right)^{-1} \vec{d}$ as driven before in Equations 34 and 35 respectively. This leads to estimating a simple solution to the under-determined inverse problem as follows:

$$\vec{m} = \underline{\underline{W}}^{-1}\,\underline{\underline{G}}^{T}\left(\underline{\underline{G}}\,\underline{\underline{W}}^{-1}\,\underline{\underline{G}}^{T}\right)^{-1}\,\vec{d}, \tag{38}$$

which simplified using the generalized inverse matrix $\underline{\underline{G}}^{-g} = \underline{\underline{W}}^{-1}\,\underline{\underline{G}}^{T}\left(\underline{\underline{G}}\,\underline{\underline{W}}^{-1}\,\underline{\underline{G}}^{T}\right)^{-1}$ to:

$$\vec{m} = \underline{\underline{G}}^{-g}\,\vec{d}. \tag{39}$$

2.2.1.3. Solution to the Inverse Problem with a Mixed-determination

In fact, geophysicists are interested in measuring sufficient datasets during the field acquisition to fully understand the subsurface and achieve their objectives. This vast amount of data makes the inverse problem to be either over-determined or mixed-determined (Menke, 1989). The mixed-determined inverse problem is based on the fact that there is no such inverse problem as being entirely over-determined or completely under-determined in the geophysical practice, but a combination of both components is found. Due to the partial underdetermination, an infinite number of solutions to the inverse problem can be obtained. On the other hand, some errors are yielded between the measured and calculated data due to the partial overdetermination and hence the deviation vector never equals zero. This situation can be effectively improved by maintaining the over-determined unknowns separate from the under-determined unknowns and a new linearly combined set of the model parameters is constructed. In that case, the over-determined model parameters can be estimated with the help of the least-squares strategy while identifying those with the minimum L_2 solution length for under-determination (Menke, 2018). In our investigation, a linearized method known as Damped Least Squares (DLSQ) or Levenberg-Marquardt is introduced as an efficient mixed-determined inversion procedure that relies primarily on L_2 norm minimization. Using the minimized objective functions of the linear over-determined and under-determined inverse problems in Equations 18 and 32 respectively, the combined objective function to be minimized in the mixed-determined case can be given as follows:

$$\emptyset = E + \varepsilon^2\,L = \sum_{k=1}^{N}\left(d_k - \sum_{i=1}^{M} G_{ki}\;m_i\right)\left(d_k - \sum_{j=1}^{M} G_{kj}\;m_j\right) + \varepsilon^2\;\sum_{q=1}^{M} m_q^2, \tag{40}$$

where ε^2 is the damping or weighting factor. This factor is vitally pertinent in error detection as well as solution length thus, its value should be accurately selected while inserted in the inversion procedure. For instance, when a large value of ε is chosen, not only the under-determined component is minimized but also a minimization of the over-determined components can be accompanied. This leads to solutions with unminimized error E and hence, inaccurate estimations of the true model parameters are produced. On the other hand, if ε has a value of zero, the error can be minimized but unfortunately, the under-determined unknowns

can not be solved due to the lack of prior information. Therefore, a proper selection of ε can be accomplished by a further trial which grantee an approximate minimization of error E and the solution's under-determined portion (Menke, 2018). By differentiating Eq. 40 in terms of the model parameters $\frac{\partial E}{\partial m_l} = 0$, the model parameter estimates can be defind as:

$$\sum_{i=1}^{M} m_i \ (\sum_{k=1}^{N} G_{ki} \ G_{kl} + \ \varepsilon^2 I) = \sum_{k=1}^{N} \ d_k \ G_{kl}, \tag{41}$$

or in a vectorial form as follows:

$$\left(\underline{\underline{G}}^T \underline{\underline{G}} + \ \varepsilon^2 \ \underline{\underline{I}}\right) \vec{m} = \ \underline{\underline{G}}^T \ \vec{d}, \tag{42}$$

where $\underline{\underline{I}}$ is the identity matrix. In that case, the generalized inverse matrix is introduced as $\underline{\underline{G}}^{-g} = \left(\underline{\underline{G}}^T \underline{\underline{G}} + \ \varepsilon^2 \ \underline{\underline{I}}\right)^{-1} \ \underline{\underline{G}}^T$ to give a simple solution as:

$$\vec{m} = \ \underline{\underline{G}}^{-g} \ \vec{d}. \tag{43}$$

Similar to that applied in both over-determined and under-determined inverse problems, the weighted norm solution of the mixed-determined inverse problem by the least squares in data space gives the following model vector:

$$\vec{m} = \ \left(\underline{\underline{G}}^T \ \underline{\underline{W}} \ \underline{\underline{G}} + \ \varepsilon^2 \ \underline{\underline{I}}\right)^{-1} \ \underline{\underline{G}}^T \ \underline{\underline{W}} \ \vec{d}, \tag{44}$$

which can be simplified using the generalized inverse matrix $\underline{\underline{G}}^{-g} = \left(\underline{\underline{G}}^T \ \underline{\underline{W}} \ \underline{\underline{G}} + \ \varepsilon^2 \ \underline{\underline{I}}\right)^{-1} \ \underline{\underline{G}}^T \ \underline{\underline{W}}$ as:

$$\vec{m} = \ \underline{\underline{G}}^{-g} \ \vec{d}. \tag{45}$$

2.2.1.4. Assessing the Inversion Procedure's Quality

Testing and evaluating the proposed inversion method on synthetic datasets before dealing with real field data measurements is an essential procedure of geophysical inversion practice. Therefore, the synthetic datasets generated for solving the geophysical problems are always contaminated with various quantities of noise to imitate field data. As a result, the influence of rising associated errors on model parameter estimation may be seen clearly. Thus, the calculation of the random noise and quantification of these estimation errors are of great importance in the practice of geophysical inversion. In that regard, several model acceptance criteria are commonly analyzed for assessing the reliability and goodness of the estimated inversion results. Using the L_2 norm in data space, the misfit between the observed $d_k^{(observed)}$

and calculated $d_k^{(calculated)}$ datasets can be quantified by the normalized data distance as follows:

$$d = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(\frac{d_k^{(observed)} - d_k^{(calculated)}}{d_k^{(calculated)}}\right)^2} * 100\ (\%), \tag{46}$$

where N is the total number of inverted data. On the other hand, the relative model distance is established to numerically investigate the model parameters in synthetic inversions concerning i-th exactly known parameters $m_i^{(exact)}$ and computed ones $m_i^{(estimated)}$ as follows:

$$D = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\frac{m_i^{(estimated)} - m_i^{(exact)}}{m_i^{(exact)}}\right)^2} * 100\ (\%), \tag{47}$$

where M is the total number of model parameters. Furthermore, a mathematical expression of a covariance matrix (Menke, 1984) to evaluate the quality of model parameter estimates can be introduced as:

$$COV = \sigma_d^2 \left(\underline{\underline{G}}^T \underline{\underline{G}}\right)^{-1}, \tag{48}$$

where σ_d^2 is the variance in the data space. The variance of i-th model parameter is given by the i-th main diagonal elements of the model covariance matrix whilst others distributed outside the diagonal indicate the model parameters correlation strength which can be formulated by the following correlation matrix as:

$$corr_{ij}^{(m)} = \frac{cov_{ij}^{(m)}}{\sqrt{cov_{ii}^{(m)}\, cov_{jj}^{(m)}}}, \tag{49}$$

with this model correlation matrix, the elements of its main diagonal are always equal to 1, whereas those on the outside diagonal range from -1 to +1. The best correlation coefficients can be estimated when their values range from 0 to ± 0.5 and hence, a more stable and accurate solution to the inverse problem is presented. When these values are close to -1 or +1, the solution is said to be highly correlated. To overcome a large number of model correlation matrix elements, it can be enhanced as a single scalar known as the mean spread value (Menke, 1984; Salat et al, 1982) as follows:

$$\mathrm{S} = \sqrt{\frac{1}{\mathrm{M}\,(\mathrm{M}-1)}\sum_{\mathrm{j}=1}^{\mathrm{M}}\sum_{\mathrm{i}=1}^{\mathrm{M}}\left(\mathrm{CORR}_{\mathrm{ij}} - \delta_{\mathrm{ij}}\right)^2}, \tag{50}$$

where δ_{ij} is the Kronecker delta symbol. This scalar, which has a range of [0;+1], can also be used as a reasonable indicator of the inversion accuracy.

2.3. THE ALGORITHM OF THE SERIES EXPANSION INVERSION PROCEDURE

The series expansion-based inversion method is considered to be one of the best achievements among several inversion approaches that developed in the last few decades by the Department of geophysics at the University of Miskolc. In the framework of the series expansion procedure, the expansion coefficients are established as the unknown model parameters. This inversion method is highly effective when the underlying geological structures to be investigated and interpreted are characterized by inhomogeneity, as defined by complicated lateral and/or vertical variations (Fancsik et al, 2021). To deal with the complexity of the earth's interior, it is preferable to discretize the subsurface into a large number of cells that are spread both laterally and vertically. Each cell attains a physical parameter giving a huge number of unknowns and therefore, the inverse problem to be solved becomes under-determined. As previously stated, this situation necessitates providing extra constraints or prior information to the problem separate from the observational data. However, it is not always the optimal solution, especially if we deal with a pure under-determination inverse problem. Therefore, converting the inverse problem from under-determination to over-determination is of great importance in most geophysical applications in general and potential field theories in particular. In this regard, the series expansion is extremely advantageous as one of the most useful discretization processes since it enables us to reduce the unknowns of series expansion coefficients compared to those found at each data point. This means that the inverse problem turns to an easily solvable overdetermined one and hence, it is possible to get extremely high resolution and accuracy where the results are based mainly on the measured datasets. The series expansion-based discretization inversion method has been applied to solve a variety of geophysical problems such as gravitational (Dobroka and Volgyesi, 2008, 2010), magnetic data processing (Dobroka et al. 2015, 2017), magnetotelluric investigation (Dobroka et al. 2013), DC geoelectric (Gyulai et al, 2010, 2017), induced polarization (Turai et al. 2010; Turai and Dobroka, 2011), and borehole geophysics (Dobroka and Szabo 2010; Dobroka et al. 2016).

The algorithm of the series expansion-based discretization of the model parameters was handled by the Department of Geophysics at the University of Miskolc as follows:

$$p(x,y,z) = \sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sum_{k=1}^{N_z} B_l\, \Psi_i(x)\, \Psi_j(y)\Psi_k(z), \tag{51}$$

where $\Psi_1 \ldots\ldots \Psi_N$ are the chosen basis functions, N_x, N_y, N_z are the series expansion limits, and B_l are the series expansion coefficients $(l = i + (j-1) * N_x + (k-1) * N_x * N_y)$. The stability of the inversion procedure is substantially influenced by the basis functions used. In that regard, great efforts have been exerted to enhance several types of basic functions such as power functions and Legendre polynomials which must be selected carefully to solve the prospective problem. Moreover, the series expansion coefficients B_l of numbers $M = N_x\, N_y\, N_z$ are the unknown model parameters to be determined as a solution to the overdetermined inverse problem. During the inversion process, these expansion coefficients are calculated iteratively, and new ones $\vec{B}^{new} = \vec{B}^{old} + \delta\vec{B}$ are generated at each step to obtain new theoretical data $p^{new}(x, y, z)$. This method allows us to estimate successfully the model unknown values in the half-space introduced by the finite difference (FDM) algorithm without incurring any concessions in the precision of the forward problem.

The ambiguity or instability of the over-determined series expansion-based inversion can be extremely avoided when a priori information about the area under exploration is provided. The most important factor is the subsurface geological structure such as layering which aids in reducing the number of inverse problem unknowns. So, we introduce some of the homogeneous and inhomogeneous structural models that have a great influence on series expansion-based inversion. In the first case when a q-th layer interface involved in a 3D homogeneous model geometry is represented, the series expansion discretization can be defined as:

$$z = f_q(x, y) = \sum_{i=1}^{N_x^{(q)}} \sum_{j=1}^{N_y^{(q)}} C_l^{(q)} \Psi_i(x)\, \Psi_j(y), \tag{52}$$

the number of unknown expansion coefficients $C_l^{(q)}$ is given as $N_x^{(q)} N_y^{(q)}$ which can be written in p layered model case assuming one physical parameter per layer with the total number of model parameters:

$$M = \sum_{q=1}^{p} N_x^{(q)}\, N_y^{(q)} + p + 1. \tag{53}$$

On the other hand, when we have a vertically inhomogeneous layered model, the physical parameters of the q-th are defined as follows:

$$p_q(z) = \sum_{i=1}^{N_q^{(p)}} D_l^{(q)}\, \Psi_i(z), \tag{54}$$

with the number of unknowns provided as:

$$M = \sum_{q=1}^{p} \left(N_x^{(q)}\, N_y^{(q)} + N_q^{(p)} \right) + 1, \tag{55}$$

where $l = L_q + i$. Hence, a homogeneous half-space is found with L_q representing an initial index in the q-th layer. Furthermore, the discretized physical parameters through series expansion are introduced in the context of a laterally inhomogeneous model in the q-th layer as:

$$p_q(x,y) = \sum_{i=1}^{N_{p,x}^{(q)}} \sum_{j=1}^{N_{p,y}^{(q)}} D_l^{(q)} \Psi_i(x)\, \Psi_j(y), \tag{56}$$

where $l = L_q + i + (j-1)\, N_x$. The number of unknown expansion coefficients in the p layered model case is given as:

$$M = \sum_{q=1}^{p} \left(N_x^{(q)}\, N_y^{(q)} + N_{p,x}^{(q)}\, N_{p,y}^{(q)} \right) + 1. \tag{57}$$

Finally, when we have an inhomogeneous layered model that varies both vertically and laterally, the physical parameters can be expressed using a series expansion in the following way:

$$p_q(x,y,z) = \sum_{i=1}^{N_{p,x}^{(q)}} \sum_{j=1}^{N_{p,y}^{(q)}} \sum_{k=1}^{N_{p,z}^{(q)}} B_l^{(q)}\, \Psi_i(x)\, \Psi_j(y) \Psi_k(z), \tag{58}$$

where $l = L_q + i + (j-1)\, N_{p,x}^{(q)} + (k-1)\, N_{p,x}^{(q)}\, N_{p,y}^{(q)}$. In that case, the number of unknowns for p layered model is expressed as:

$$M = \sum_{q=1}^{p} \left(N_x^{(q)}\, N_y^{(q)} + N_{p,x}^{(q)}\, N_{p,y}^{(q)}\, N_{p,y}^{(q)} \right) + 1. \tag{59}$$

In all of the above-mentioned circumstances, the number of data is higher than the number of unknowns ($N > M$), whether homogeneous or inhomogeneous structural models, implying that the inverse problem is usually over-determined. This prompts us to provide the proposed inversion algorithm of series expansion with a priori information bearing in mind the incredible performance of inversion in computing procedures, particularly when using finite difference (FDM) or finite element (FEM) methods.

2.4. THE INVERSION-BASED FOURIER TRANSFORMATION

As discussed above in chapter 1, we introduce the inversion-based Fourier transformation (IRLS-FT) algorithm as an effective approach for geophysical data processing. This procedure evaluates noise reduction capabilities and determines the most likely quality of observed data that has been contaminated by complicated noises such as outliers. In this inversion-based filtering, the series expansion-based Fourier transformation is considered as the weighted inverse problem using the weights calculated in the framework of Steiner's Most Frequent

Value (MFV) method (Steiner, 1988 and 1997). Moreover, the Hermite functions are employed as basis functions, together with the eigenfunctions' advantage characteristic that allows us to speed up the calculation of the Jacobian matrix elements. In the following section, 1D and 2D algorithms of the IRLS Fourier Transformation (Dobróka et al. 2015, 2017), as well as the IRLS inversion methodology are presented.

2.4.1. 1D ALGORITHM OF THE IRLS FOURIER TRANSFORMATION

In the framework of signal processing and interpretation enhancing methods, the Fourier transformation algorithm can be performed to drive the data in frequency from time-domain datasets, and therefore, one can get the most important features of interest. The Fourier transformation in a one-dimensional case can be expressed as follows:

$$U(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u(t)\, e^{-j\omega t}\, dt, \tag{60}$$

where $U(\omega)$ is the Fourier spectrum that converted from a phenomenon carried by some time-domain observations $u(t)$, ω and j are the angular frequency and imaginary unit respectively. On the other hand, a time-domain signal is established using an inverse Fourier transformation as:

$$u(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} U(\omega)\, e^{j\omega t}\, d\omega. \tag{61}$$

To facilitate the procedures of data processing and interpretation, we use the series expansion method to discretize the continuous Fourier frequency spectrum $U(\omega)$ as follows:

$$U(\omega) = \sum_{i=1}^{M} B_i\, \Psi_i(\omega), \tag{62}$$

where B_i represents the expansion coefficients that are calculated as the solution of the overdetermined inverse problem and $\Psi_i(\omega)$ are the chosen basis functions. Inserting Eq. 62 into Eq. 61, one can estimate the theoretical data as shown in the following formula:

$$u_k^{theor} = \sum_{i=1}^{M} B_i \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Psi_i(\omega)\, e^{j\omega t_k}\, d\omega, \tag{63}$$

where the following term

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \Psi_i(\omega)\, e^{j\omega t_k}\, d\omega,$$

is a component of the *NxM* Jacobian matrix $G_{k,i}$ (that consists of N elements of time-domain data and M elements of unknown expansion coefficients). It is highly significant to take in mind that the Jacobian matrix can be considered as the inverse Fourier transform of the basis function Ψ_i and hence, the theoretical data can be summarized as:

$$u_k^{theor} = \sum_{i=1}^{M} B_i \; G_{k,i}. \tag{64}$$

Because the frequency spectrum is given as an integral in the interval (-∞,∞), the basis functions should be given in the same domain, which can be achieved by using the Hermite functions and polynomials. The mathematical background of the Hermite polynomials was established by Laplace (1812) and further refined by Hermite (1864) to incorporate multidimensional polynomials. Since the frequency covers such a wide range in most of all geophysical applications, scaling has a vital role in modifying the Hermite functions executed. In that regard, the scaled Hermite functions of the inverse Fourier transformation can be defined as:

$$H_n(\omega,\alpha) = \frac{e^{\frac{-\alpha\omega^2}{2}} h_n(\omega,\alpha)}{\sqrt{\sqrt{\frac{\pi}{\alpha}}\, n!\, (2\alpha)^n}} \qquad \text{where } h_n(\omega,\alpha) = (-1)^n \, e^{\alpha\omega^2} \left(\frac{d}{d\omega}\right)^n e^{-\alpha\omega^2}, \tag{65}$$

where α is the scale factor. To find a speedy solution to the forward problem, Dobróka et al. (2017) employed the nonscaled Hermite function $H_n^{(0)}$ as eigenfunction of the inverse Fourier transformation as follows:

$$G_{kn} = \frac{1}{\sqrt[4]{\alpha}} (j)^n \, H_n^{(0)} \left(\frac{t}{\sqrt{\alpha}}\right). \tag{66}$$

It can be seen that the Jacobian matrix does not seem to involve integration, which makes the inversion procedures faster and less time-consuming. It is obvious that the easy computation of the Jacobian matrix elements enables us to produce theoretical data in a linear form as introduced in Eq. 63. In this case, the general element of the deviation vector can be expressed as:

$$e_s = u_s^{measured} - \sum_{i=1}^{M} B_i \; G_{k,i}. \tag{67}$$

By calculating the deviation vector, the inverse problem can be continued straightforwardly.

2.4.2. 2D ALGORITHM OF THE IRLS FOURIER TRANSFORMATION

The two-dimensional algorithm of the inversion-based Fourier transformation (2D IRLS-FT) can subsequently be driven in the same way as the previously stated one-dimensional inversion. The 2D Fourier transformation algorithm is used to convert the data from the space domain to the spatial frequency domain as follows:

$$U\left(\omega_x,\omega_y\right) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} u\,(x,y)\, e^{-j(\omega_x x+\omega_y y)}\, dx\, dy, \tag{68}$$

while the 2D space domain data can be acquired through its inverse as:

$$u\,(x,y) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} U\left(\omega_x,\omega_y\right) e^{j(\omega_x x+\omega_y y)}\, d\omega_x\, d\omega_y. \tag{69}$$

The angular spatial frequencies of the 2D spatial frequency spectrum are denoted by ωx and ωy, whereas x and y are the spatial coordinates. Furthermore, the continuous 2D Fourier transform $U\left(\omega_x,\omega_y\right)$ is discretized utilizing the series expansion method as follows:

$$U\left(\omega_x,\omega_y\right) = \sum_{n=1}^{N}\sum_{m=1}^{M} B_{n,m}\, \Psi_{n,m}\left(\omega_x,\omega_y\right). \tag{70}$$

The model parameters of the over-determined inverse problem are calculated as expansion coefficients $B_{n,m}$, whilst the appropriately chosen basis functions are $\Psi_{n,m}\left(\omega_x,\omega_y\right)$. In our linear inversion procedure, the Jacobian matrix needed for theoretical data calculations can be thought of as the inverse Fourier transform of the basis function as:

$$G_{k,l}^{n,m} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \Psi_{n,m}\left(\omega_x,\omega_y\right) e^{j(\omega_x x+\omega_y y)}\, d\omega_x\, d\omega_y, \tag{71}$$

where $k = 1,2,\dots,K$ and $l = 1,2,\dots,L$ refer to the x and y coordinates of the measurement points. At this stage, Hermite functions and polynomials are used to supply the basis functions with the same domain of the integral's interval (-∞,∞) of the frequency spectrum. Similar to that illustrated above in the 1D inversion algorithm case, the following two equations describe the two-dimensional scaled Hermite functions using scale factors α and β as follows:

$$H_n(\omega_x,\alpha) = \frac{e^{\frac{-\alpha\omega_x^2}{2}} h_n(\omega_x,\alpha)}{\sqrt{\sqrt{\frac{\pi}{\alpha}}\, n!\,(2\alpha)^n}} \qquad \text{where } h_n(\omega_x,\alpha) = (-1)^n\, e^{\alpha}\, \omega_x^2 \left(\frac{d}{d\omega_x}\right)^n e^{-\alpha\,\omega_x^2} \tag{72}$$

$$H_m(\omega_y,\beta) = \frac{e^{\frac{-\beta\omega_y^2}{2}} h_m(\omega_y,\beta)}{\sqrt{\sqrt{\frac{\pi}{\beta}}\, m!\,(2\beta)^m}} \qquad \text{where } h_m(\omega_y,\beta) = (-1)^m\, e^{\beta}\, \omega_y^2 \left(\frac{d}{d\omega_y}\right)^m e^{-\beta\,\omega_y^2}. \tag{73}$$

On the other hand, non-scaled Hermite functions $H_n^{(0)}$and $H_m^{(0)}$ are inserted as eigenfunctions of the inverse Fourier transformation by Dobróka et al (2017) to fasten the forward problem solution as follows:

$$G_{k,l}^{n,m} = \frac{(j)^{n+m}}{\sqrt[4]{\alpha\beta}}\, H_n^{(0)}\left(\frac{x_k}{\sqrt{\alpha}}\right) H_m^{(0)}\left(\frac{y_l}{\sqrt{\beta}}\right). \tag{74}$$

In this case, the theoretical data at a given measuring point (x_k, y_l) can be expressed as:

$$u\,(x_k,y_l)^{theor} = \sum_{n=1}^{N}\sum_{m=1}^{M} B_{n,m}\, G_{k,l}^{n,m}, \tag{75}$$

which can be simply reformulated using the notations $u\,(x_k\,, y_l) = u_s$, $B_{n,m} = B_i$ and $u_s = \sum_{i=1}^{I} B_i\, G_{s,i}$ as:

$$u_s^{theor} = \sum_{i=1}^{I} B_i\, G_{s,i}, \tag{76}$$

where $i = 1, \dots, I$ and $s = 1, \dots, S = KL$ and hence, the inverse problem can be straightforwardly solved by minimizing a proper norm of the deviation vector's general element as:

$$e_s = u_s^{measured} - \sum_{i=1}^{I} B_i\, G_{s,i}. \tag{77}$$

2.4.3. THE IRLS INVERSION METHODOLOGY

As a method of linearized geophysical inversion, the least-squares method can be used to solve sets of linear equations for quick procedures. Legendre (1805) was the first to publish a clear and concise exposition of the least-squares method. It is well known that the least-squares inversion method is more effective when dealing with data of Gaussian noise (regular noise). It can be used to evaluate the misfit between the measured and predicted data by minimizing the L_2 norm of the deviation vector as follows:

$$E_2 = \sum_{k=1}^{N} e_k^2, \tag{78}$$

giving the minimized normal equations as:

$$\underline{\underline{G}}^T\, \underline{\underline{G}}\, \vec{B} = \underline{\underline{G}}^T\, \vec{u}^{measured}. \tag{79}$$

Because most geophysical data measured during field surveys contain varying levels of irregular noise, such as outliers, the deviation vector of another norm called the weighted norm is minimized like the following:

$$E_w = \sum_{k=1}^{N} w_k\, e_k^2. \tag{80}$$

Employing Cauchy weights w_k improves the stability and robustness of the inversion procedures, and it can be expressed as:

$$w_k = \frac{\sigma^2}{\sigma^2 + e_k^2}, \tag{81}$$

where σ^2 is the scale parameter, the value of which must be known as an a priori provided quantity for the inversion process to continue. As a result, the most frequent value (MFV) approach of Steiner must be employed to solve this problem, with the Cauchy-Steiner weights being estimated (Steiner, 1997) as follows:

$$w_k = \frac{\varepsilon^2}{\varepsilon^2 + e_k^2}, \tag{82}$$

where ε^2 is the dihesion that can be computed iteratively through an internal loop as:

$$\varepsilon_{j+1}^2 = 3\,\frac{\sum_{k=1}^{N} \frac{e_k^2}{\left(\varepsilon_j^2 + e_k^2\right)^2}}{\sum_{s=1}^{N} \left(\frac{1}{\varepsilon_j^2 + e_s^2}\right)^2}. \tag{83}$$

It is obvious that at the starting step ($j = 0$), the Steiner scale factor $\varepsilon_0 \leq \frac{\sqrt{3}}{2}\,(e_{max} - e_{min})$ can be used to calculate the Steiner scale factor of the next step ε_{j+1}^2 and this process continues until the stop-criterion is met, which can be also achieved by using a fixed iteration number. At this stage, the inverse problem has become nonlinear because of using the Cauchy-Steiner weights, thus the iteratively reweighted least-squares (IRLS) method (Scales et al, 1988) must be used. In this inversion procedure, the predicted data at the 0th step can be calculated using the expansion coefficients $\vec{B}^{(0)}$ which derived from the non-weighted least-squares method $u_k^{(0)} = \sum_{i=1}^{M} B_i^{(0)}\, G_{ki}$. Then the deviation vector can be computed from $e_k^{(0)} = u_k^{measured} - u_k^{(0)}$ yielding the following weights:

$$w_k^{(0)} = \frac{\varepsilon^2}{\varepsilon^2 + \left(e_k^{(0)}\right)^2}. \tag{84}$$

Now the misfit between the measured and calculated data can be determined at the first iteration step as the following:

$$E_w^{(1)} = \sum_{k=1}^{N} w_k^{(0)}\, e_k^{(1)^2}, \tag{85}$$

which is minimized to give linear normalized equations as:

$$\underline{\underline{G}}^T\, \underline{\underline{W}}^{(0)}\, \underline{\underline{G}}\; \vec{B}^{(1)} = \underline{\underline{G}}^T\, \underline{\underline{W}}^{(0)}\, \vec{u}^{measured}, \tag{86}$$

where $\underline{\underline{W}}^{(0)}$ is the weighting matrix. It can be seen that the expansion coefficients at the 1st iteration step $\vec{B}^{(1)}$ calculated from the above equation of the linear weighted least-squares method are subsequently used to estimate new predicted data $u_k^{(1)} = \sum_{i=1}^{M} B_i^{(1)}\, G_{ki}$, and hence, the deviation vector is $e_k^{(1)} = u_k^{measured} - u_k^{(1)}$. Similarly, the weights $w_k^{(1)} = \frac{\varepsilon^2}{\varepsilon^2 + \left(e_k^{(1)}\right)^2}$ are calculated to compute the new misfit function $E_w^{(2)} = \sum_{k=1}^{N} w_k^{(1)}\, e_k^{(2)^2}$ resulting in $\vec{B}^{(2)}$ through

its minimization. This process is continued and repeated over the j-th iteration step as shown in the normal equation:

$$\underline{\underline{G}}^T \, \underline{\underline{W}}^{(\mathrm{j}-1)} \, \underline{\underline{G}} \;\; \vec{B}^{(\mathrm{j})} \;=\; \underline{\underline{G}}^T \, \underline{\underline{W}}^{(\mathrm{j}-1)} \, \vec{u}^{measured}, \tag{87}$$

until a stop criterion is met at the last iteration step and therefore, the series expansion coefficients are accepted as a solution of the inverse problem.

Chapter 3
NUMERICAL INVESTIGATION OF THE INVERSION-BASED 1D FOURIER TRANSFORMATION

The application of the Fourier transform is beneficial in most physical sciences, including engineering, physics, chemistry, and applied mathematics. Therefore, we introduce the inversion-based Fourier transformation in our study as a robust and reliable methodology for improving the processing and interpretation of geophysical datasets. Before moving on to two-dimensional datasets, this chapter discusses in brief several benefits of the inversion method applied to one-dimensional datasets such as noise reduction capabilities, non-regular sampling problems as well as a new promising direction for datasets sampled incompletely within the area under investigation.

3.1. TESTING THE NOISE REDUCTION CAPABILITIES

One of the most serious issues is geophysical datasets that have been corrupted with various amounts of noise during the field survey. This noise prevalence must be taken into account before dealing with the different techniques of data processing and interpretation. Thus, putting our inversion-based Fourier transformation approach to the test to see if we can improve the quality of geophysical signal outputs is a key topic of interest.

3.1.1. REDUCING OUTLIER SENSITIVITY IN IRLS-FT

In this section, we shed the light on the comparison between the traditional Fourier transformation DFT and the inversion-based Fourier transformation IRLS-FT for outlier sensitivity rejection when applied to 1D equidistantly sampled datasets. Furthermore, the algorithm of the one-dimensional inversion-based Fourier transformation (S-IRLS-FT) method was handled by the Department of Geophysics, University of Miskolc, which was followed by some PhD students, as seen in Daniel Oduro Boatey Nuamah's and Hajnalka Szegedi's PhD dissertations. In addition, this work has been already studied in many kinds of literature (Dobroka et al, 2012; Szegedi and Dobróka, 2014; Dobroka et al, 2015; Nuamah and Dobróka, 2019; Nuamah et al, 2021), therefore, in this thesis, we briefly present the signal processing applicability of the inversion-based Fourier transformation to 1D datasets. In that regard, a regular sampling interval of $\Delta t = 0.005$ sec was chosen to create a time-domain signal over 401 measuring points in a tested area covering a time range of [-1 to +1] as shown in *Figure 2*. This Morlet wavelet is generated using the following formula:

$$u(t) = k\, t^{\eta}\, e^{-\lambda t}\, sin(\mu t + \emptyset), \tag{88}$$

with a given time t and signal parameters input values $k = 738.91$, $\eta = 2$, $\lambda = 20$, $\mu = 40\pi$, and $\emptyset = \pi/4$.

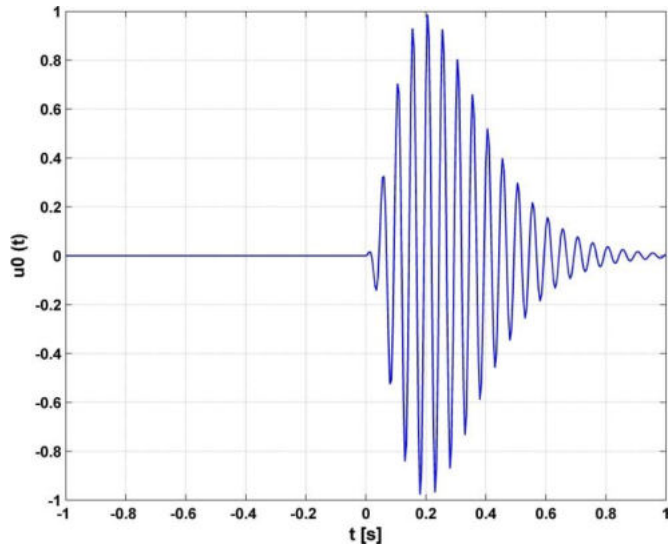

Figure 2. *Noise-free wavelet in the time domain.*

In this case of equidistantly sampled noise-free datasets, the time-domain signal is converted to a frequency domain to give the real and imaginary components of the Fourier frequency spectrum by the traditional (DFT) and inversion (IRLS) methods as clearly seen in *Figure 3a* and *Figure 3b* respectively. It is demonstrated that both converted spectra are shown to have identical shapes and amplitudes, proving that the two techniques are very applicable to noise-free datasets.

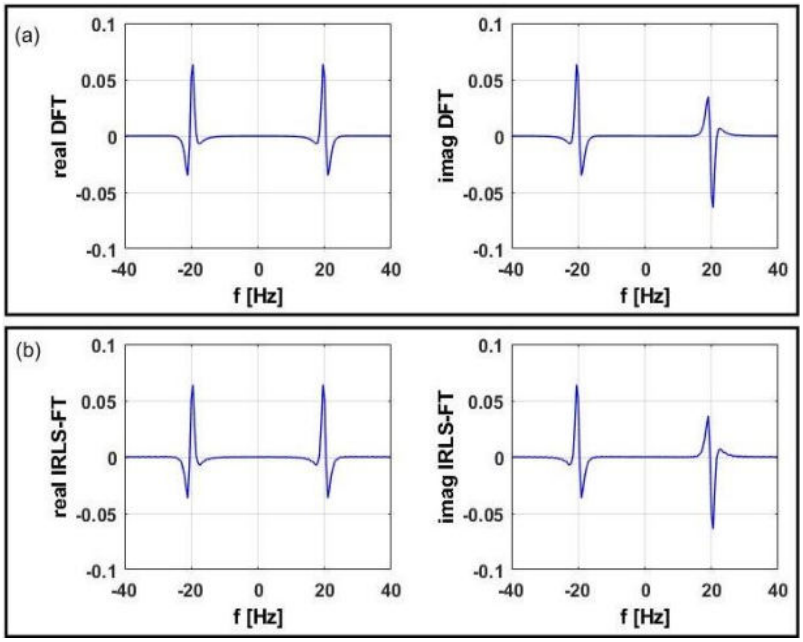

Figure 3. *a) DFT spectrum of the Noise-free signal in the frequency domain, b) IRLS-FT spectrum of the Noise-free signal in the frequency domain.*

To assess the noise reduction capacity of both DFT and S-IRLS-FT methods, this noise-free wavelet is contaminated with both Gaussian noise using zero mean and standard deviation $\sigma = 0.01$ *(Figure 4a)* and a random noise of Cauchy distribution to imitate real field data measurements *(Figure 4b)*.

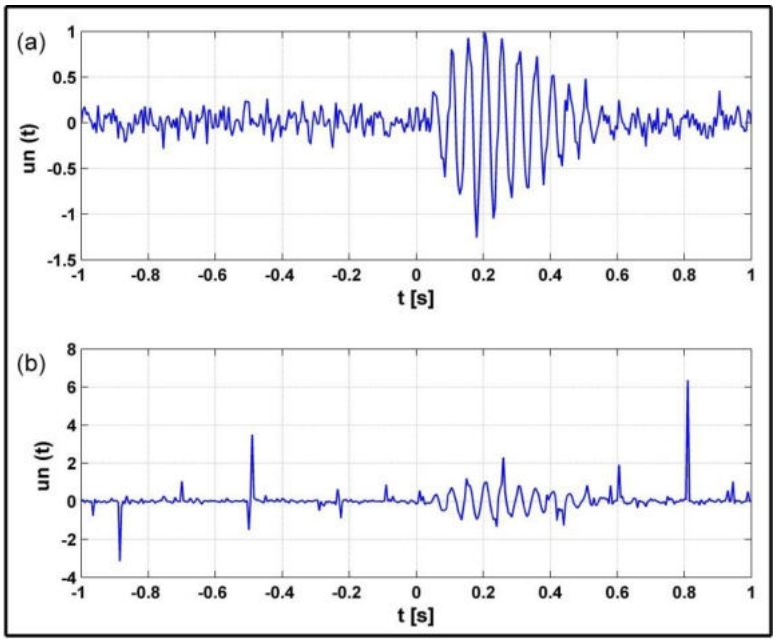

Figure 4. *a) Gaussian noise-contaminated wavelet in the time-domain, b) Cauchy noise-contaminated wavelet in the time-domain.*

On the other hand, to compare the inversion method's applicability to that of noise-free datasets, the noisy signal containing Gaussian noise *(Figure 4a)* is subjected to the Fourier transformation yielding in 1D noisy frequency spectra by DFT and inversion-based (IRLS) methods as introduced in *Figure 5a* and *Figure 5b* respectively. In addition, *Figure 6a* and *Figure 6b* represent the 1D noisy Fourier frequency spectra obtained by applying the DFT and inversion method to the Cauchy noise-contaminated data respectively. It is noticed that the one-dimensional noisy frequency spectra produced by the inversion method either with Gaussian or Cauchy noises are highly enhanced and improved compared to those found by the traditional DFT method to be nearly similar to that of the noise-free Fourier spectrum (*Figure 3a)* reflecting the professional noise rejection capability of the inversion used.

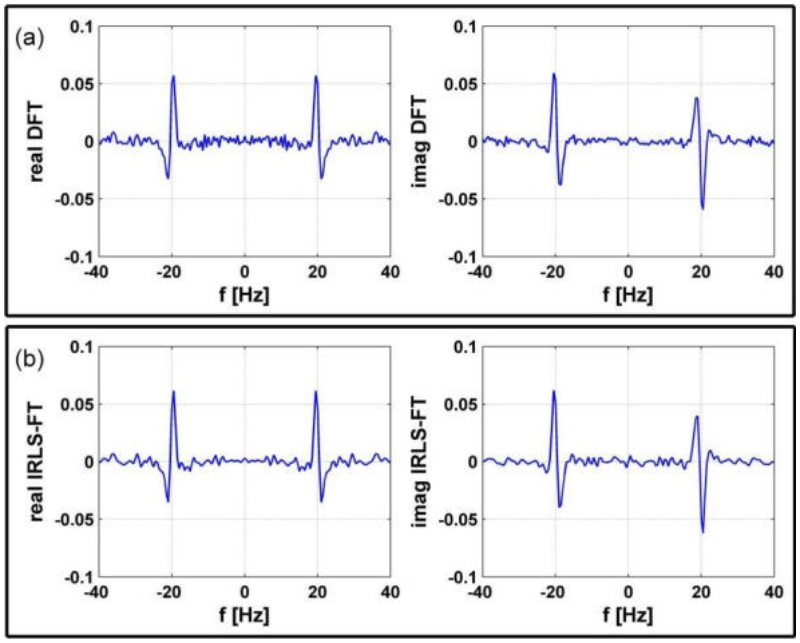

Figure 5. *a) DFT spectrum of the Gaussian noise-contaminated signal in the frequency domain, b) IRLS-FT spectrum of the Gaussian noise-contaminated signal in the frequency domain.*

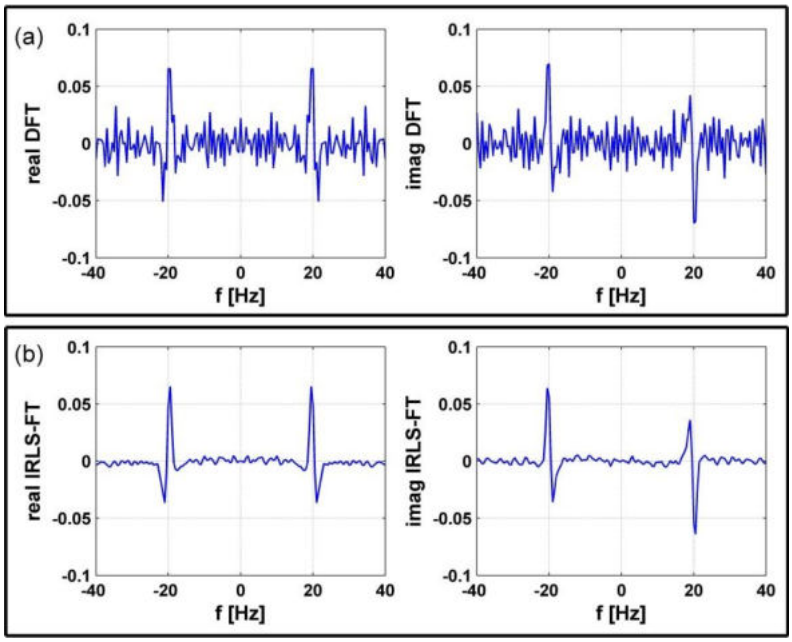

Figure 6. *a) DFT spectrum of the Cauchy noise-contaminated signal in the frequency domain, b) IRLS-FT spectrum of the Cauchy noise-contaminated signal in the frequency domain.*

For quantitative interpretation, the output signals can be evaluated mathematically by calculating the misfit between the noisy and noise-free datasets. In the context of the time-domain, the data distance assessed for the noiseless and noisy signals may therefore be as follows:

$$d = \sqrt{\frac{1}{N}\sum_{i=1}^{N}[u^{noiseless}(t_i) - u^{noisy}(t_i)]^2}, \quad (89)$$

while for the frequency domain accuracy, the model distance between the spectra can be expressed as:

$$D = \left[\frac{1}{M}\sum_{i=1}^{M}\left(Re\left[U^{noisy}(f_i)\right] - Re\left[U^{noiseless}(f_i)\right]\right)^2 + \frac{1}{M}\sum_{i=1}^{M}\left(Im\left[U^{noisy}(f_i)\right] - Im\left[U^{noiseless}(f_i)\right]\right)^2\right]^{\frac{1}{2}}, \quad (90)$$

where N and M are the associated numbers of the examination area's time and frequency points. In this case, the data distance calculated between the noise-free *(Figure 2)* and Gaussian noise-contaminated data *(Figure 4a)* is 0.1032, whilst for Cauchy noise *(Figure 4b)* is 0.4554. Moreover, in the case of Gaussian noise, the model distances between the noise-free *(Figure 3a)* and noisy spectra produced by the DFT *(Figure 5a)* and the inversion-based IRLS-FT *(Figure 5b)* are numerically computed as 0.0041 and 0.0027 respectively. However, the noiseless spectrum *(Figure 3a)* and the spectrum-based Cauchy noise obtained by the DFT method *(Figure 6a)* are separated by a model distance of D= 0.0182, whereas the estimated value for that produced by the inversion (IRLS-FT) method *(Figure 6b)* has a model distance of D= 0.0021. According to these RMS distance values as well as the resolution of signals in both time and frequency domains, one can prove that the inversion-based 1D Fourier transformation is extremely effective, efficient, and gives sufficient improvements in noise reduction compared to the conventional Fourier transformation method (DFT) especially when we deal with outliers of Cauchy noise-contaminated data.

3.2. TESTING THE NON-REGULAR SAMPLING PROBLEM

The problem of non-regular sampling may be accompanied during the procedures of the field data acquisition due to some restrictions encountered in the investigated area, such as the presence of unpaved roads or any natural or artificial conditions. This was a great motivation to implement the proposed inversion-based Fourier transformation (IRLS-FT) method on data sampled non-regularly as previously discussed by Nuamah and Dobróka (2019). In this section, we present an application of the inversion method to two different strategies of one-dimensional non-regularly sampled datasets; the first is when complete data are randomly selected while the other is concerned with the random walk measurement positioning.

3.2.1. COMPLETE DATASET WITH RANDOM MEASUREMENT LOCATIONS

The above-mentioned 1D equidistantly dataset, sampled over 401 measuring points in an interval of [-1 to +1], can then be used to create the random measurement placement scenario. In that regard, a random selection of the observing points is constructed by means of randomizing their interval length. In this investigation, two times the half sampling distance around the exact or non-shifted position is presented. *Figure 7* depicts a plot of randomly sampled measuring points on a non-regular grid.

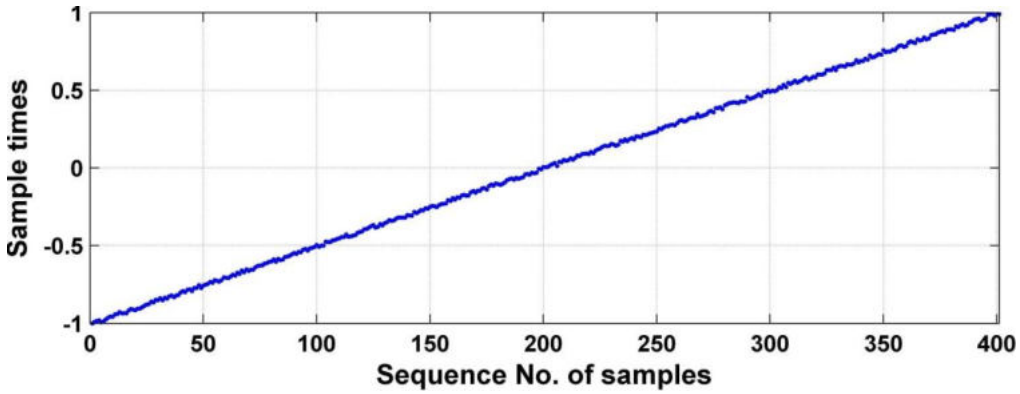

Figure 7. *Plot of randomly sampled measuring points.*

It can be seen that the sampling points deviate from the straight line due to the randomizing procedure and the resultant Morlet wavelet in the time domain can be displayed in *Figure 8*.

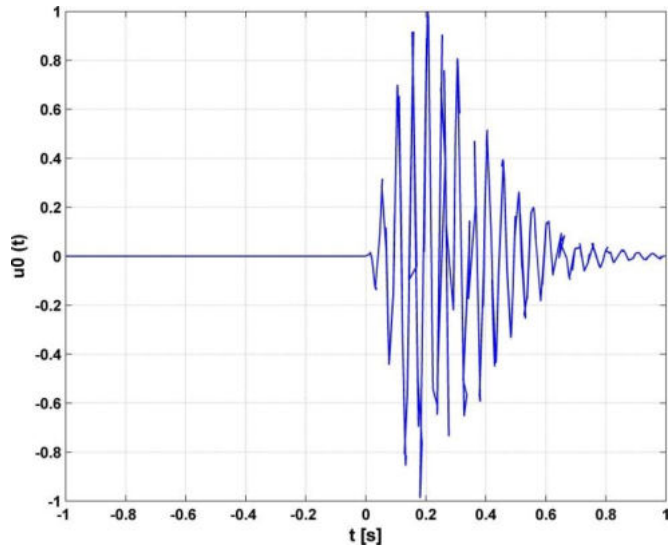

Figure 8. *Non-regular sampling wavelet in the time domain.*

The inversion algorithm is applied to the above-mentioned time-domain signal to verify the effectiveness and accuracy of the inversion-based 1D Fourier transformation (IRLS-FT)

method. This results in the real and imaginary parts of the Fourier frequency spectrum for 1D datasets sampled non-equidistantly as demonstrated in *Figure 9*.

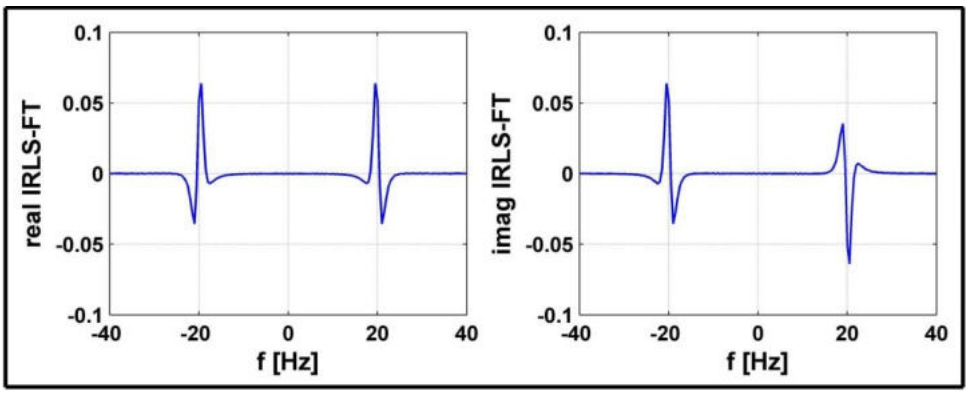

Figure 9. *IRLS-FT spectrum of the non-regular signal in the frequency domain.*

It is observed that the Fourier frequency spectrum based on non-regular sampling datasets has similar shapes and amplitudes when compared to those obtained in the regular sampling case using either the conventional DFT *(Figure 3a)* or inversion-based (IRLS) *(Figure 3b)* methods. The inversion method's efficiency is further emphasized by the results of the 1D inverse Fourier transformation, in which the above frequency-domain spectrum is transformed into a calculated wavelet in the time domain *(Figure 10)*.

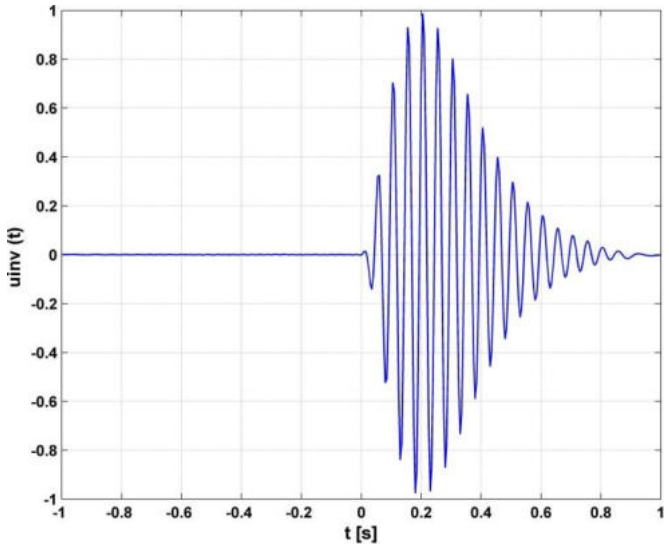

Figure 10. *The calculated wavelet by IRLS-FT in the time domain.*

It is seen that the calculated time-domain signal is nearly identical to that generated for the 1D equidistant case *(Figure 2)*, proving great enhancements by the inversion method. Numerically, the calculated data distance between the wavelet-based inversion *(Figure 10)* and the non-regular time-domain signal *(Figure 8)* demonstrates a very low value of 0.0033. These results reveal that the inversion-based Fourier transformation is highly effective in both time

and frequency domains when dealing with random measurement positioning of complete datasets.

3.2.2. COMPLETE DATASET WITH RANDOM-WALK MEASUREMENT POSITIONS

Random walk measurements that are carried out over non-equidistant grids are frequently proceeded in most geophysical field data acquisition due to a variety of natural or man-made impediments, such as caves, road accessibility, buildings, and so on. Furthermore, the great developments in the geophysical survey's equipment over the last few decades have encouraged many geophysicists to measure datasets at random locations in the field and therefore, help them save time and effort during field data gathering. As a result, our objective is to show how effective can the inversion-based Fourier transformation (IRLS-FT) deal with the random walk measurements in one-dimension (Nuamah and Dobróka, 2019). On behave of this random walk sampling, the above-mentioned Morlet wavelet-based 1D complete equidistant datasets that are initially created, across 401 measuring points in an interval of [-1 to +1], are further randomized to produce non-equidistant intervals. *Figure 11* shows the grid of non-equidistantly sampled measuring stations in one dimension.

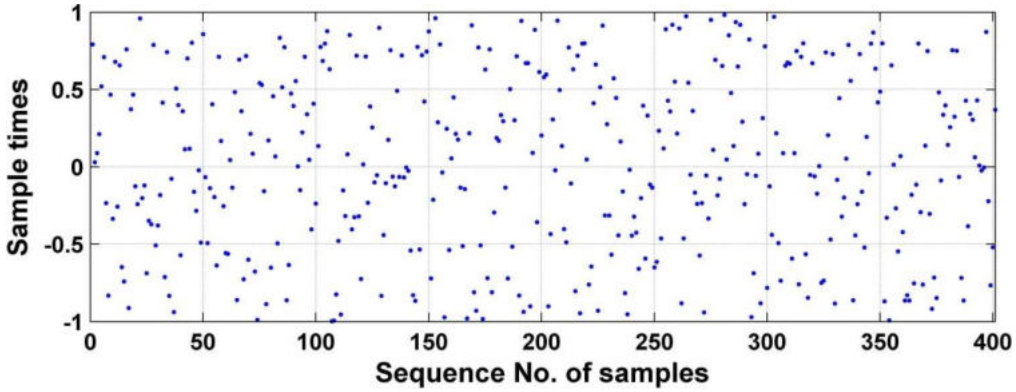

Figure 11. *The randomized measuring points that required for non-equidistant wavelet generation.*

The randomly selected sampling points are then subjected to the algorithm of the inversion-based 1D Fourier transformation to give a one-dimensional Fourier frequency spectrum as clearly shown in *Figure 12*.

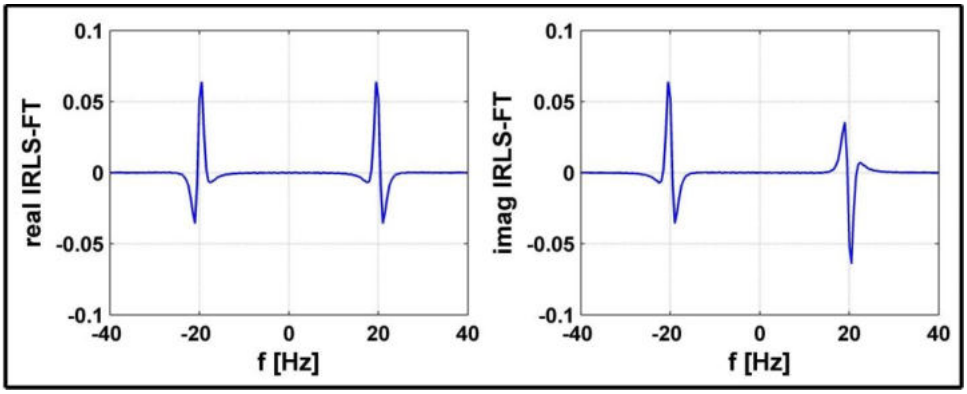

Figure 12. *IRLS-FT spectrum of the randomly sampled signal in the frequency domain.*

One can notice that the 1D frequency spectrum of the non-equidistantly sampled signal by the inversion method demonstrates sufficient improvements. The shapes and amplitudes of both real and imaginary parts of the IRLS spectrum are very similar to that obtained by the traditional Fourier transformation approach (DFT-FT) in an equidistantly sampling case *(Figure 3a)*. Moreover, a wavelet in the time-domain is converted by means of the 1D inverse Fourier transformation to confirm the inversion's efficiency as seen in *Figure 13*.

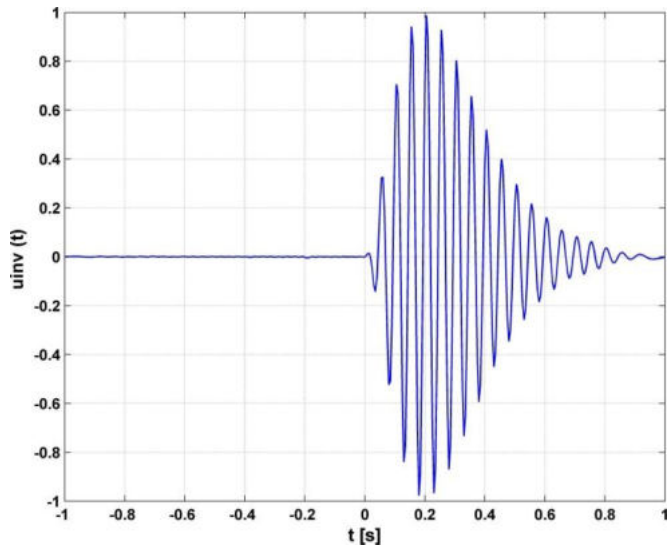

Figure 13. *The calculated wavelet of the randomly sampled datasets by IRLS-FT in the time domain.*

A close examination of the inverted time-domain signal shows a considerable similarity when compared to the signal generated for the 1D equidistant case *(Figure 2)* with numerical data distance value of 0.0029. Accordingly, it can be said that the 1D inversion-based Fourier transformation (IRLS-FT) method is effective, robust, and extremely applicable for processing 1D complete datasets gathered in non-equidistant or random walk measurement configurations.

3.3. TESTING THE INCOMPLETE SAMPLING PROBLEM

In geosciences, most of the natural phenomena of the earth's interior are commonly analyzed through field datasets measuring rather than laboratory experiments. Depending on the conditions whereby the measurements are taken, the vast majority of these datasets are generally riddled with gaps. Therefore, challenges in a variety of disciplines, such as spectrum estimation and defining boundary conditions for numerical models can be caused by this incomplete or missing data (Kondrashov and Ghil, 2006). In a common geophysical modeling protocol, data processing can be provided by missing input variables of the measurements being disregarded, even if it's just one of the independent variables. But this situation does not guarantee the optimum solution where the model testing or construction is enhanced in the absence of these incomplete or missing data raws, and therefore, a lot of most likely information could, unfortunately, be lost (Gill et al, 2007). All of these factors are compelling enough to make our interest mainly directed to thinking about how we can treat this kind of missing dataset. In that regard, we develop a new MATLAB code that applies the inversion-based Fourier transformation (IRLS-FT) approach to incomplete synthetic datasets sampled both equidistantly and non-equidistantly in one dimension. In reality, this research point is highly promising especially in geophysics since the majority of geophysical datasets are inhabited by different degrees of incomplete or missing values.

3.3.1. INCOMPLETE DATASET WITH REGULAR SAMPLING

In this section, the newly developed algorithm of the inversion-based 1D Fourier transformation method is implemented to test its effectiveness when some observing points in the regularly sampled datasets are neglected or cancelled. To do so, the same example of regular interval sampling over 401 measuring points is used as shown by a straight line in *Figure 14a* resulting in a time-domain Morlet wavelet in the [-1 to +1] time range *(Figure 14b)*.

In order to prove the stability and efficacy of the inversion method, we present here three different examples of missing data strategies. In each case, we randomly cancel some percentages of the sampling datasets from the entire data points taken over a regular grid. The objective is to make sure at which level of missing data percentage can the inversion-based 1D Fourier transformation method still effectively work without any distortion.

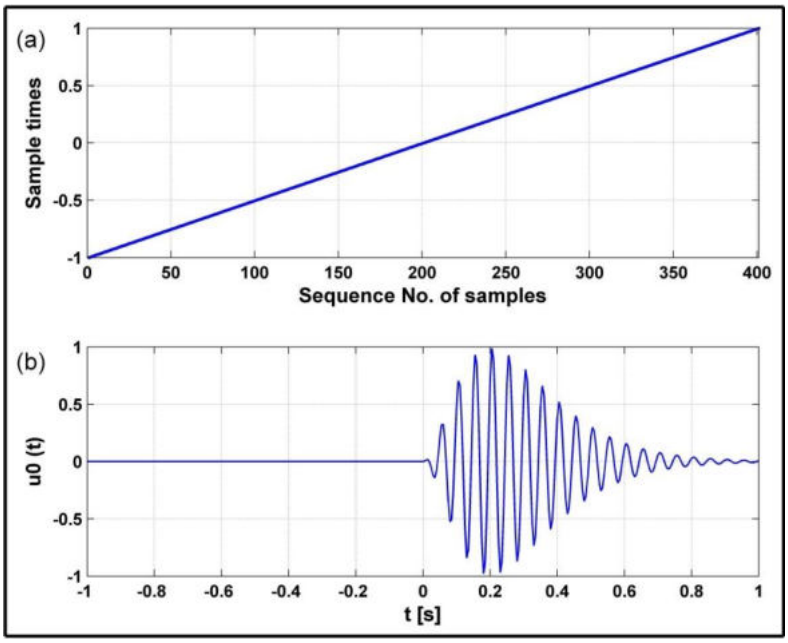

***Figure 14**. a) Plot of equidistantly sampled measuring points, b) Regular sampling wavelet in the time-domain.*

Now, let's start with an example where 15% of the measuring data points are randomly missed or cancelled, so in this case, a total of around 340 measurement points is used. The algorithm of the 1D inversion method is then applied to such missing datasets to give IRLS Fourier frequency spectrum as illustrated in *Figure 15a*. It is obvious that the spectrum obtained by the inversion approach is the same as the spectrum obtained by the DFT method in the case of complete datasets sampled regularly *(Figure 3a)*. Furthermore, the calculated wavelet in the time-domain is obtained by the inverse Fourier transformation as depicted in *Figure 15b*. A striking resemblance appears between the inverted signal when 15% missing data are utilized and the complete regular sampling wavelet in the time-domain *(Figure 14b)*. This is also achieved by the very low value of data distance of 0.0030 reflecting that the inversion method (IRLS-FT) is highly applicable when 15% of datasets are missing. However, the second example represents the missing 30% of the data. Hence, the inversion-based 1D Fourier transformation (IRLS-FT) algorithm is only implemented on about 282 observing data points. *Figure 16a* shows the 1D Fourier frequency spectrum by the IRLS inversion method which is identical to those obtained in both cases of 15% missing data *(Figure 15a)* and the complete datasets sampled regularly through the DFT method *(Figure 3a)*. In addition, the 1D inverse Fourier transformation of the IRLS spectrum-based 30% missing data gives a time-domain wavelet as seen in *Figure 16b*. This inverted wavelet has similar characteristics when compared to those found by the complete and 15% missing datasets *(Figure 14b* and *Figure 15b* respectively*)*. Moreover, the data distance in the time-domain between the inverted signal in

the case of 30% missing data and the complete regular sampling wavelet *(Figure 14b)* is numerically computed to be 0.0032.

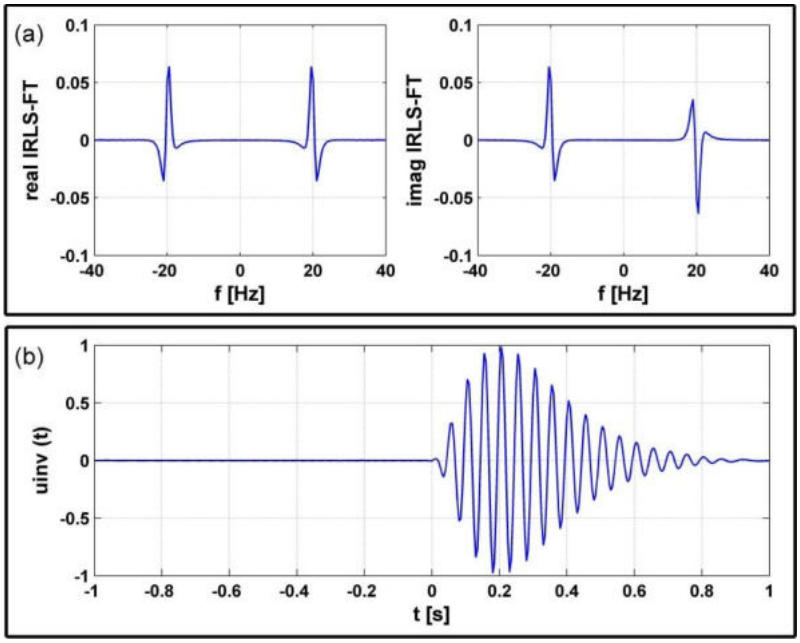

Figure 15. *a) The IRLS-FT spectrum when 15% of the regularly sampled datasets are missing in the frequency domain, b) The calculated wavelet using IRLS-FT when 15% of the regularly sampled datasets are missing in the time domain.*

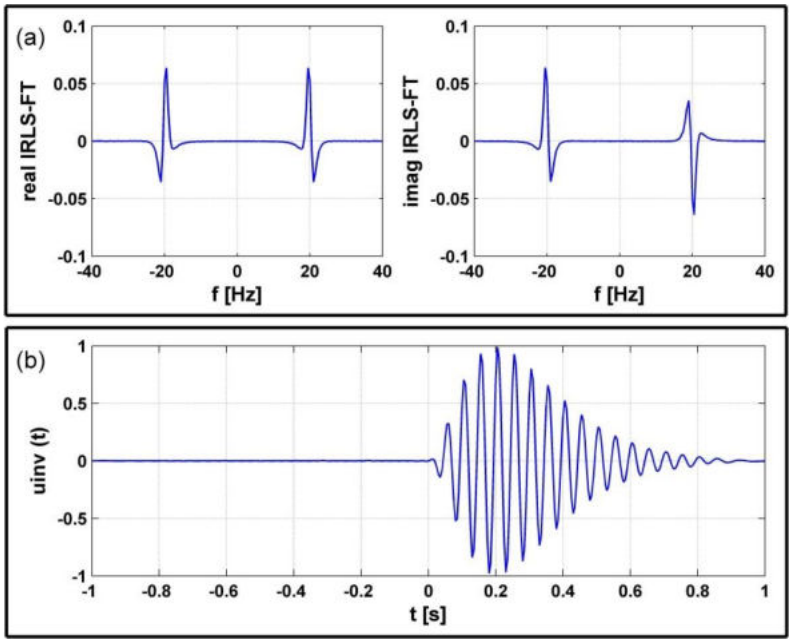

Figure 16. *a) The IRLS-FT spectrum when 30% of the regularly sampled datasets are missing in the frequency domain, b) The calculated wavelet using IRLS-FT when 30% of the regularly sampled datasets are missing in the time domain.*

This similarity and low value of data distance prove that our IRLS-FT inversion method is effective and still working when we deal with datasets that have 30% of their power missing or cancelled. Depending on the stability and efficiency of the newly developed inversion-based 1D Fourier transformation (IRLS) method in the above-mentioned cases of 15% and 30% missing datasets, we increase the percentage of data missed up to 50%. This means that only half of the measuring data points are subjected to the inversion method while neglecting the other half. In this situation, a total approximate number of the observing points 200 is used to produce a 1D Fourier frequency spectrum by the IRLS inversion method as depicted *in Figure 17a*. It is noticed that the resultant frequency spectrum is very similar in shape and amplitude to those spectra obtained by the DFT of regularly sampled complete datasets *(Figure 3a)*, IRLS-FT of 15% missing datasets *(Figure 15a)*, and IRLS-FT of 30% missing datasets *(Figure 16a)*. The accuracy is further demonstrated in *Figure 17b* by applying the 1D inverse Fourier transformation on the IRLS spectrum deduced from 50% missing datasets. Comparing this inverted wavelet to those found in the different cases, such as equidistantly sampled complete datasets *(Figure 14b)*, missing 15% of the data *(Figure 15b)*, missing 30% of the data *(Figure 16b)* shows great similarities. As in the previous cases, the time-domain wavelet driven from the missing 50% of the datasets is numerically characterized by a low data distance value to be 0.0033. This demonstrates that half of the measuring points are sufficient to provide effective results in both time and frequency domains.

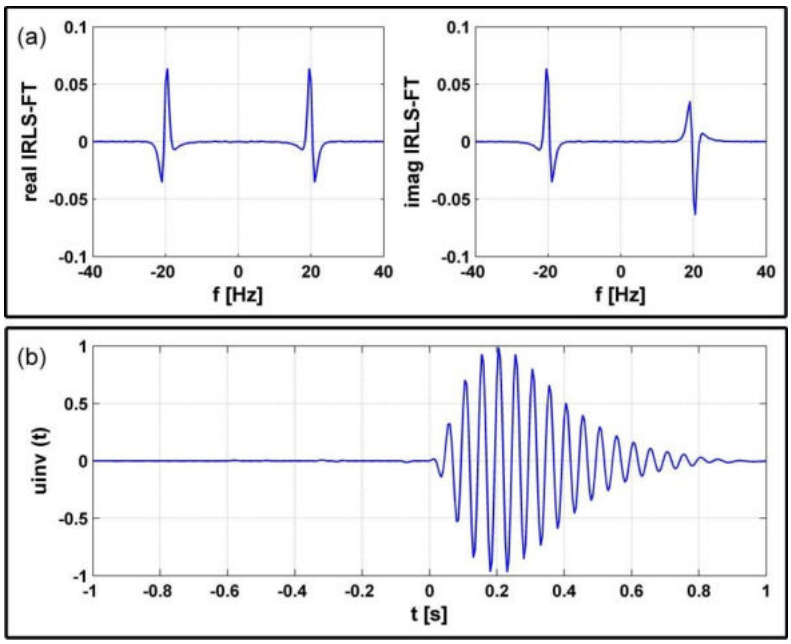

Figure 17. *a) The IRLS-FT spectrum when 50% of the regularly sampled datasets are missing in the frequency domain, b) The calculated wavelet using IRLS-FT when 50% of the regularly sampled datasets are missing in the time domain.*

From our experience, we can say that the newly developed MATLAB code of the inversion-based 1D Fourier transformation (IRLS) method is highly effective, robust, and applicable for processing regularly sampled incomplete datasets even when 50% of the datasets are missing. Above this percentage, the inversion algorithm becomes unstable and does not perform effectively. *Figure 18* shows a relationship between the percentage of missing data and the data distance values. It can be seen that the data distances record relatively low values for all missing data below roughly 50% and gradually increase after this missing data point, indicating that our inversion with incomplete datasets sampled regularly has made considerable improvements. Because of these enhancements and the efficient inversion findings, the algorithm is improved to be evaluated on non-regularly sampled incomplete datasets as described in the following section.

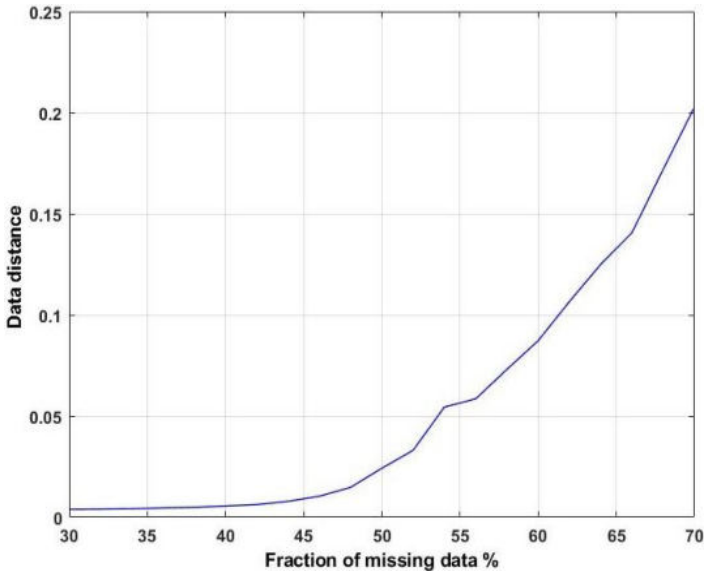

Figure 18. *A plot of missing data fraction against data distance for regularly sampled incomplete datasets using the IRLS-FT method.*

3.3.2. INCOMPLETE DATASET WITH NON-REGULAR MEASUREMENT LOCATIONS

In the geophysical field survey, we occasionally face more challenging hurdles. One of the concerns that should be addressed is non-equidistant measurement positions in conjunction with incomplete datasets. This sampling problem piques our curiosity, thus the newly developed algorithm of the inversion-based 1D Fourier transformation (IRLS) method is applied to incomplete datasets collected on a non-regular grid. To deal with this measurement configuration, the same initially created 1D complete equidistant datasets over 401 measuring points are randomized as shown in *Figure 19a*. The resultant Morlet wavelet in the time-domain is introduced in *Figure 19b* using an interval of [-1 to +1]. It is observed that the time-domain

signal is highly disturbed and deformed due to the random selection procedure of the measuring data points.

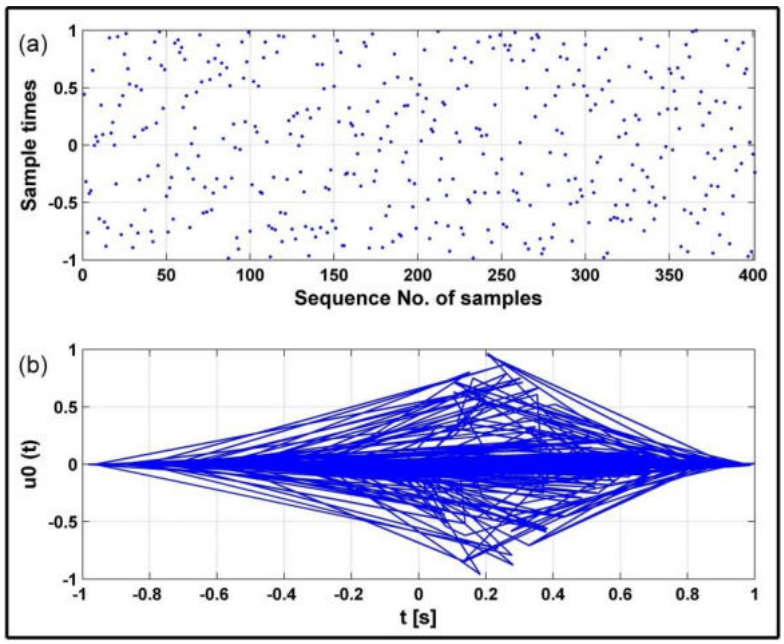

Figure 19. *a) The randomized measuring points to produce non-equidistant intervals, b) Non-regular sampling wavelet in the time domain.*

To compare the results to those obtained by the above-stated incomplete regularly sampled data case, the IRLS inversion method is applied to 1D non-equidistantly sampled data with a missing 15% of the measurements. *Figure 20a* and *Figure 20b* represent the 1D Fourier frequency spectrum and the time-domain wavelet obtained by 1D inverse Fourier transformation respectively. The inverted time-domain signal *(Figure 20b)* is effectively improved and enhanced as indicated by a very low data distance value of 0.0026. The inversion process, as can be seen, returns the signal's original shape, which is comparable to the regularly sampled wavelet *(Figure 14b)*. In addition, the same results are produced when compared to the 15% missing data associated with the regularly sampled spectrum and wavelet shown in *Figure 15a* and *Figure 15b* respectively. These commonalities in the time and frequency domains demonstrate the inversion approach's reliability in the non-regular case with 15% missing datasets. Furthermore, 30% of the measuring data points are randomly cancelled or removed from the whole dataset and supposed to the IRLS inversion method for evaluation. The 1D frequency spectrum is estimated in *Figure 21a* while the inversed Morlet wavelet in a time domain is illustrated in *Figure 21b*. As previously estimated, both the Fourier spectrum and inverted time-domain signal are identical when compared to those obtained in the case of the 30% missing data sampled regularly *(Figure 16a* and *Figure 16b* respectively*)*. The data

distance is numerically computed as 0.0027 demonstrating the higher applicability of our inversion when 30% of the non-equidistantly sampled datasets are missing.

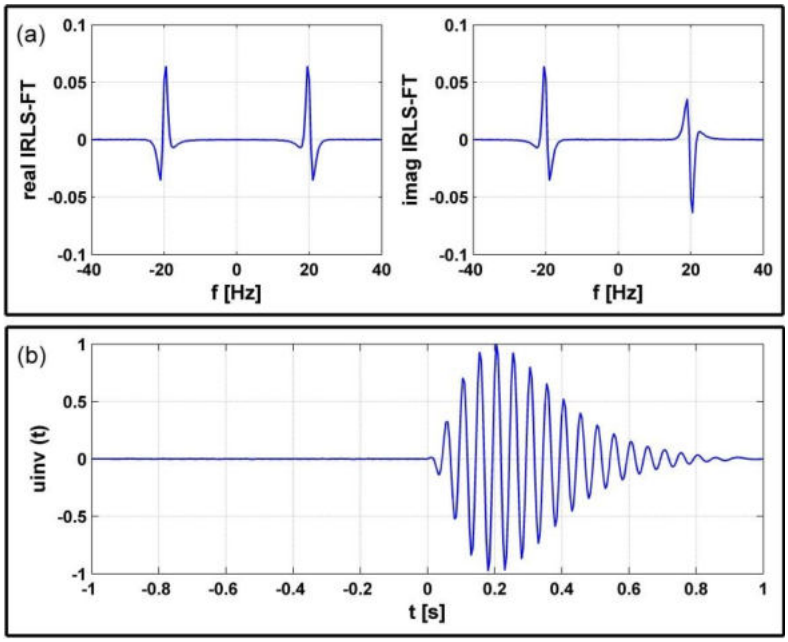

Figure 20. *a) The IRLS-FT spectrum when 15% of the non-regularly sampled datasets are missing in the frequency domain, b) The calculated wavelet using IRLS-FT when 15% of the non-regularly sampled datasets are missing in the time domain.*

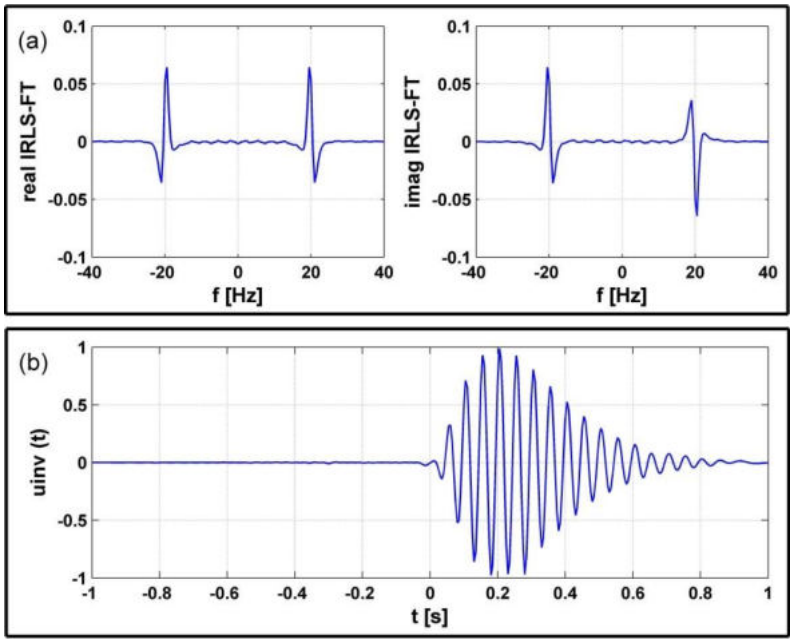

Figure 21. *a) The IRLS-FT spectrum when 30% of the non-regularly sampled datasets are missing in the frequency domain, b) The calculated wavelet using IRLS-FT when 30% of the non-regularly sampled datasets are missing in the time domain.*

To demonstrate the efficiency of the newly designed IRLS inversion method concerning the non-regular sampling measures, we merely implement the proposed procedure to half of the datasets and cancel the other half. Therefore, the inversion algorithm is performed on 1D non-equidistant datasets that are missing 50% of their total power. *Figure 22a* shows the Fourier frequency spectrum which is nearly similar to that obtained by the regular 50% missing data case *(Figure 17a)*. In that case, the data distance is calculated as 0.0029. In addition, the wavelet calculated by the inverse Fourier transformation in the time-domain *(Figure 22b)* is improved to be nearly similar to those estimated by both regularly sampled complete *(Figure 14b)* and 50% missing *(Figure 17b)* datasets. This means that our inversion method is applicable and effective when fifty percent of the observing datasets are only available.

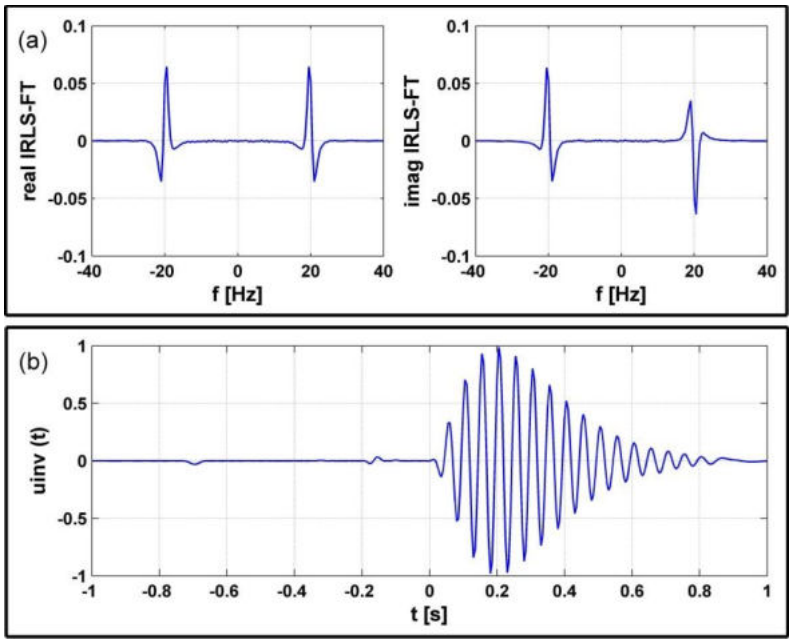

Figure 22. *a) The IRLS-FT spectrum when 50% of the non-regularly sampled datasets are missing in the frequency domain, b) The calculated wavelet using IRLS-FT when 50% of the non-regularly sampled datasets are missing in the time domain.*

According to these considerable and efficient results in both time and frequency domains, we can honestly conclude that the newly developed MATLAB code of the inversion-based 1D Fourier transformation (IRLS) method can proficiently work with the non-regularly sampled incomplete datasets as one would measure along with regular interval grids. Moreover, it can be recommended for any investigated area either with regular or non-regular incomplete datasets, even if half of the observation sites (50%) (as a maximum) are missing. This is also clearly proven when we involve a significant relationship between the missing data percentage and data distances as illustrated in *Figure 23*. Treating the non-regularly sampled incomplete datasets indicated that for all missing data under 50%, the newly developed inversion approach

exhibit relatively low data distance values and steadily grows following this missing data point reflecting the higher inversion efficiency.

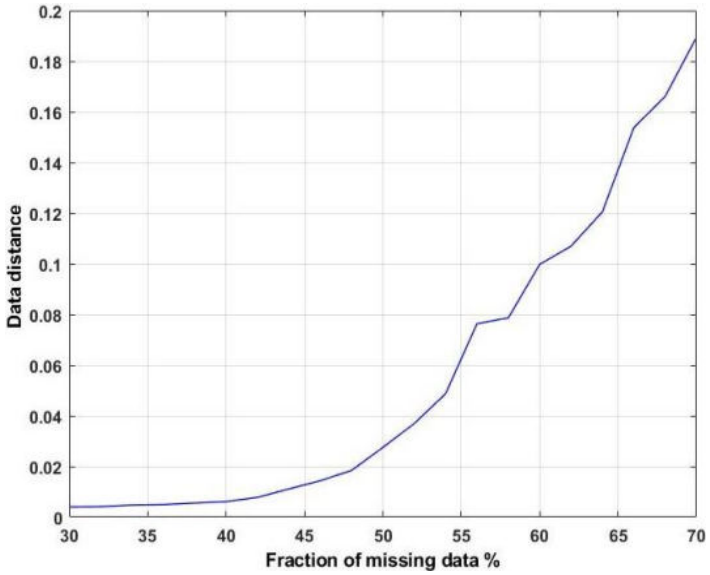

Figure 23. *A plot of missing data fraction against data distance for non-regularly sampled incomplete datasets using the IRLS-FT method.*

3.4. TESTING THE BLOCK INCOMPLETE SAMPLING PROBLEM

The effectiveness and quality of the newly developed algorithm of the inversion-based 1D Fourier transformation (IRLS) method applied to incomplete datasets sampled both regularly and non-regularly paid our attention to more complicated geophysical survey designs. Due to the presence of various variables beyond our control, such as lakes, caves, mountains, and so on, geophysical field datasets taken along profiles are occasionally corrupted. These natural features can be carefully ignored and progressively continue the field data acquisition along with the same profile directions. This problem may result in a large gap or in other words a wide area with no information laying in between the other two informational parts of the measured profile, and hence, a complex problem of so-called block-incomplete sampling is provided. The block-incomplete designs have been thoroughly detailed in some literature (See et al, 1997; Kaur and Garg, 2020; Sangpara et al, 2021). In that regard, the newly developed reconstruction algorithm has been modified to see how effective it is with this type of survey design. The basic concept is to calculate the Fourier spectrum of the incomplete dataset by using the inversion-based Fourier transformation method and to use the inverse Fourier transformation of this spectrum in calculating the time domain data even in the time points in which there are no input data. It is assumed that – to a limit- the data missing doesn't modify the spectrum significantly, so the inverse FT can give a reliable approximation. In this section, the same time-domain Morlet wavelet described in all of our prior analyses is utilized through

401 measuring points in a tested area covering a time range of [-1 to +1] as shown in *Figure 24*.

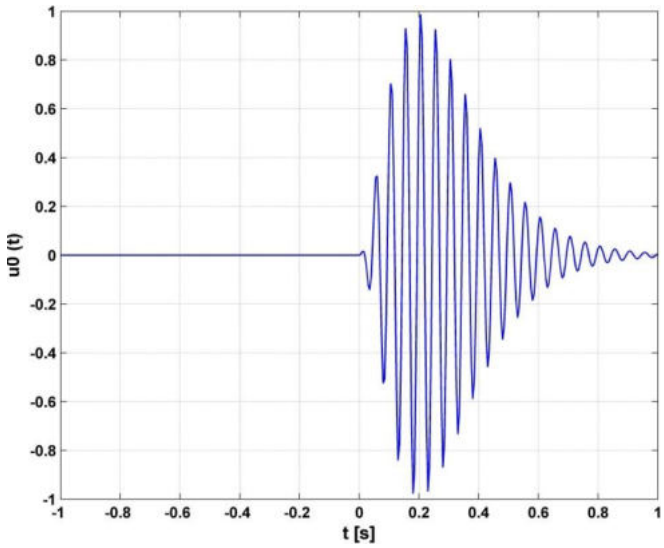

***Figure 24**. The generated Morlet wavelet in the time domain.*

To test our one-dimensional inversion (IRLS) method on block-incomplete sampling datasets, a block with an interval of 0.2 sec to 0.23 sec is specified concerning the above-mentioned wavelet in the time-domain *(Figure 24)*. All the data contents of this block are randomly removed or cancelled for designing such block-incomplete sampling as demonstrated in *Figure 25*. In this case, the 1D inversion procedure is approximately implemented only on 395 total data points.

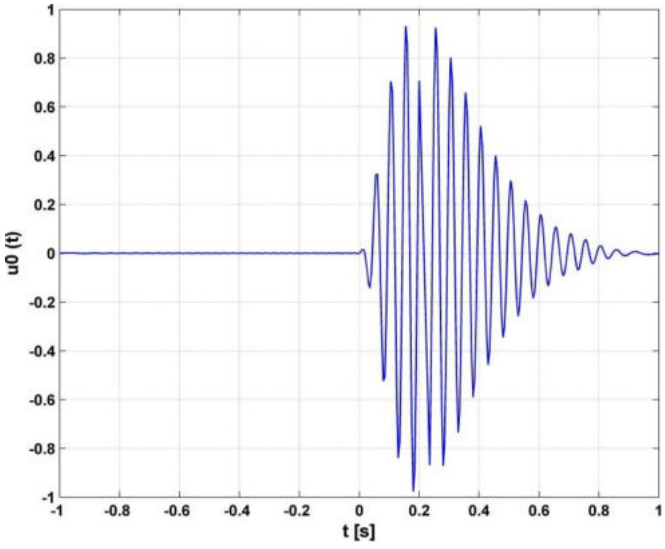

***Figure 25**. The block-incomplete sampling wavelet with an interval [0.2–0.23 sec] in the time domain.*

The algorithm of the inversion-based 1D Fourier transformation is then applied to this block-incomplete sampling wavelet resulting in a Fourier frequency spectrum as shown in

Figure 26a. It is seen that a great similarity in both shapes and amplitude is proven when a comparison is performed between the frequency spectrum of the block-incomplete sampling datasets *(Figure 26a)* and the spectra deduced for the regularly sampled complete datasets by the DFT approach *(Figure 3a)* as well as the regularly and non-regularly sampled incomplete datasets by IRLS inversion method (for example *Figure 15a* and *Figure 20a* respectively). Moreover, the calculated wavelet by the 1D inverse Fourier transformation is shown in *Figure 26b*. For easier evaluation, both the inverted wavelet and the principle-generated wavelet are displayed on the same plot with red and blue colors respectively. The inverted wavelet is shown to be the same as that generated in the complete dataset case *(Figure 24)*, with very minor variations, as numerically proven by a very low data distance value of 0.0034. In addition, compared to those estimated by the IRLS of incomplete datasets sampled both regularly and non-regularly (for example *Figure 15b* and *Figure 20b* respectively), sufficient agreements are highly evidenced.

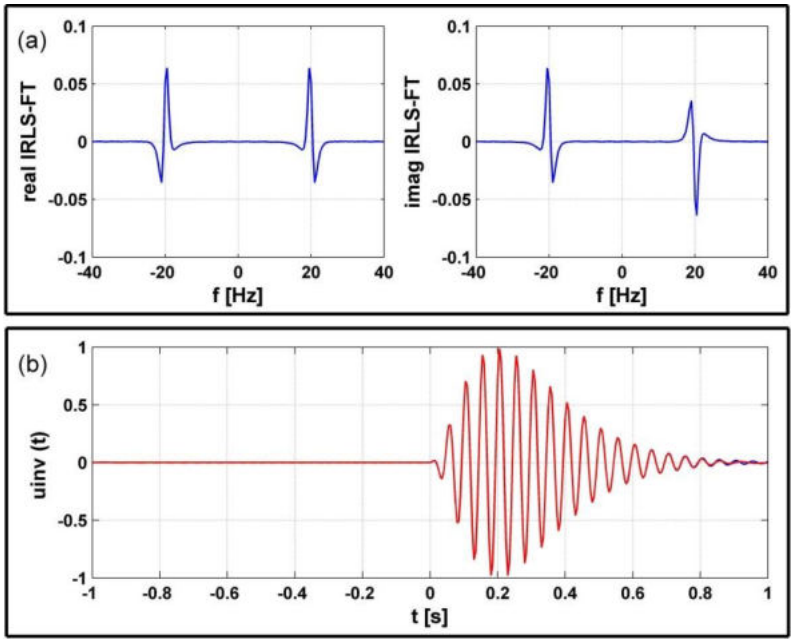

Figure 26. *a) The IRLS-FT spectrum of the block-incomplete sampling signal with an interval [0.2–0.23 sec] in the frequency domain, b) The calculated wavelet of the block-incomplete sampling datasets with an interval [0.2–0.23 sec] by IRLS-FT in the time-domain.*

Furthermore, another example of the block-incomplete sampling datasets is introduced in this study for the sake of our 1D inversion evaluation. In this example, the limits of the block are further expanded to be set from 0.2 sec to 0.26 sec as clearly demonstrated by a wide gap with no information *(Figure 27)*. This means that a higher number of the observing data points

is missing, so about 390 points are only subjected to the inversion-based Fourier transformation (IRLS) method.

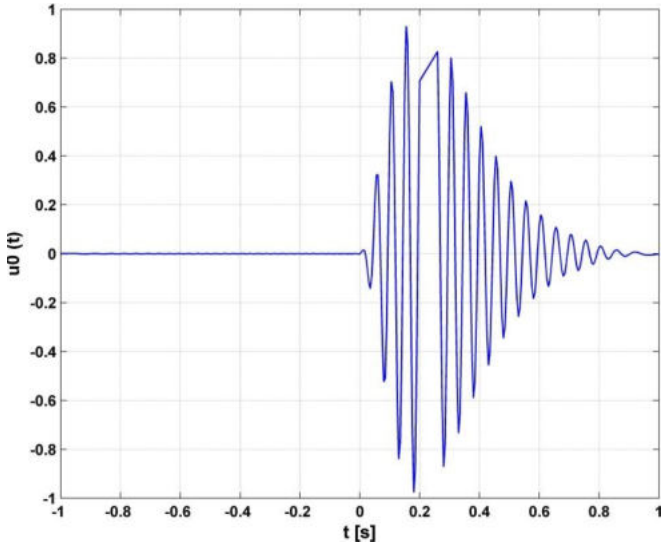

***Figure 27**. The block-incomplete sampling wavelet with an interval [0.2–0.26 sec] in the time domain.*

Following this block sampling interval, the 1D Fourier frequency spectrum estimated by the IRLS inversion method *(Figure 28a)* is very close to those given by the DFT of the regularly sampled complete datasets *(Figure 3a)* and the IRLS of both regularly and non-regularly sampled incomplete datasets (for example *Figure 15a* and *Figure 20a* respectively). However, the 1D inverse Fourier transformation is performed resulting in a time-domain wavelet of block-incomplete sampling datasets as shown in *Figure 28b*. With the same data distance value 0.0034 obtained above in the [0.2 - 0.23] block limits case, it is evident that only minor differences exist between the IRLS inverted wavelet (red color) and the principal one (blue color). This means that the inverted wavelet by our inversion method reconstructed its original shape as if we were operating with complete datasets. Furthermore, the inverted wavelet by the inversion algorithm *(Figure 28b)* is nearly identical to those driven by the IRLS of both regularly and non-regularly sampled incomplete datasets (for example *Figure 15b* and *Figure 20b* respectively). According to these findings, the newly developed inversion-based 1D Fourier transformation algorithm is extremely effective, stable, robust, and applicable at this level of block-incomplete sampling datasets, and it may thus be suggested for geophysical field datasets gathered in incomplete block designs.

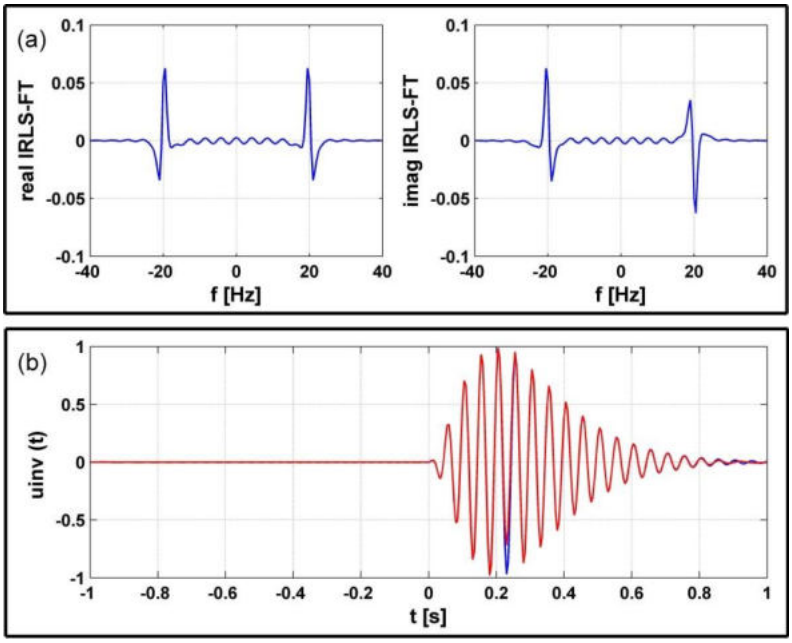

Figure 28. *a) The IRLS-FT spectrum of the block-incomplete sampling signal with an interval [0.2–0.26 sec] in the frequency domain, b) The calculated wavelet of the block-incomplete sampling datasets with an interval [0.2–0.26 sec] by IRLS-FT in the time-domain.*

These procedures are continued at higher windows of block-incomplete datasets to detect the block limits at which the 1D inversion method is fundamentally broken or not satisfied. To do so, the block distance range is raised to [0.2–0.3 sec], and hence, the IRLS inversion algorithm is only executed on roughly 381 total data points as shown in *Figure 29*.

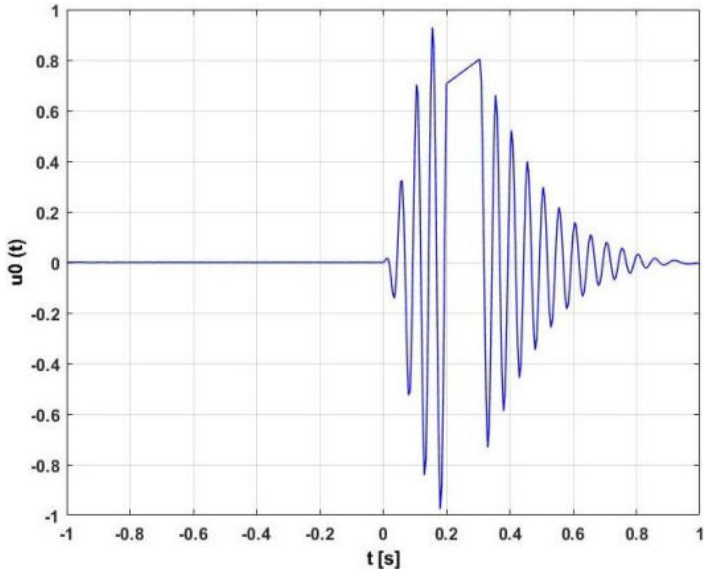

Figure 29*. The block-incomplete sampling wavelet with an interval [0.2–0.3 sec] in the time domain.*

The used data points are subjected to the inversion resulting in a 1D Fourier frequency spectrum as presented in *Figure 30a*. A distinct distortion in both real and imaginary parts of

the IRLS spectrum with this block-incomplete interval is revealed. This deformation appears in shapes and amplitudes when compared to those of regularly sampled complete datasets *(Figure 3a)* by the DFT method as well as the regularly and non-regularly sampled incomplete datasets by the IRLS inversion procedure (for example *Figure 15a* and *Figure 20a* respectively). Furthermore, *Figure 30b* shows the 1D inverse Fourier transformation of the mentioned spectrum producing a block-incomplete sampling wavelet in the time domain. In a comparison with the signal obtained with complete datasets case *(Figure 24)* and those found by the IRLS of incomplete datasets sampled regularly and non-regularly (for example *Figure 15b* and *Figure 20b* respectively), the inverted signal using this block interval [0.2–0.3 sec] is highly distorted in both shapes and amplitudes. This means that the stability, efficiency, and applicability of the newly developed inversion-based 1D Fourier transformation in solving the block-incomplete sampling problems is well satisfied for specific block intervals; beyond this interval, the inversion algorithm is not effective enough and stops working.

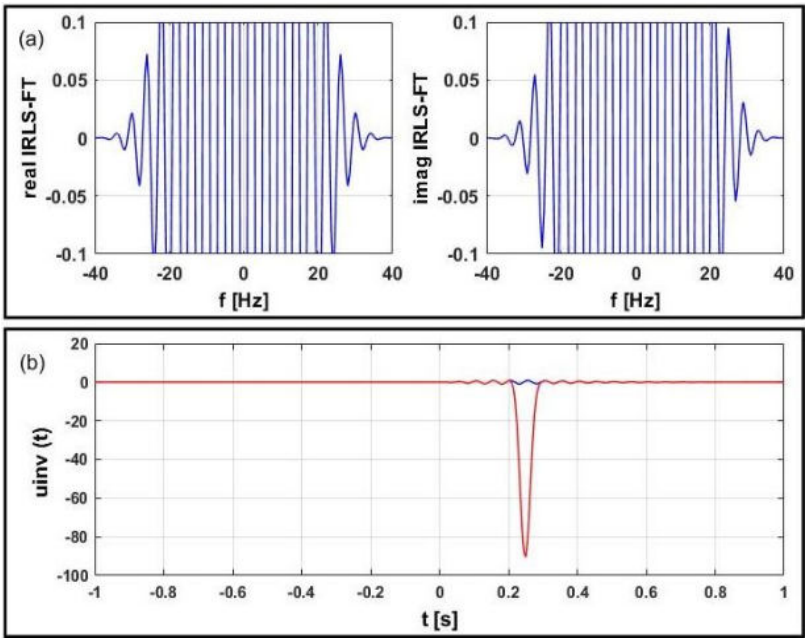

Figure 30. *a) The IRLS-FT spectrum of the block-incomplete sampling signal with an interval [0.2–0.3 sec] in the frequency domain, b) The calculated wavelet of the block-incomplete sampling datasets with an interval [0.2–0.3 sec] by IRLS-FT in the time-domain.*

According to the results demonstrated above, I declare thesis statement 1 as follows:

Thesis 1

I gave a comprehensive analysis of the inversion-based Fourier transformation algorithm applied to 1D synthetic wavelet to reduce the outlier sensitivity as well as to deal with the issues

of irregular sampling and incomplete missing data. To achieve such aims, I used a time-domain signal created over 401 measuring points which are then contaminated with both Gaussian and Cauchy distributed noises.

Dealing with the regularly and non-regularly sampled incomplete datasets, I found that the inversion is highly effective, robust, and applicable even when half of the measurements (50%) are missing. Far beyond this percentage, the inversion algorithm becomes unstable and does not perform effectively where the inverse problem is marginally over-determined. In my analysis of the block-incomplete sampling problem, I found that for certain block intervals, the accuracy of the 1D datasets in both frequency and time domains is well satisfactory which is highly distorted as I increase the missing block intervals because of the marginal over-determination rate.

Chapter 4
MAGNETIC DATA PROCESSING USING THE 2D IRLS-FT METHOD

4.1. AN OVERVIEW OF THE MAGNETIC SURVEYING METHOD

It is generally believed that the magnetic method is one of the oldest and the most widely used geophysical techniques for subsurface geology investigation. The rapidity of the coverage procedures and low cost for areas being surveyed are just some of the many advantages of the technique. The magnetic method is primarily based on measuring variations in the magnetic field of the earth, whose intensity is related to the location of the observing points in the magnetic field at the surface of the earth, as well as the distribution of subsurface magnetized sources with varying magnetic susceptibilities, i.e. induced and/or remanent magnetization (Hubbard and Linde, 2011). The properties of some magnetic minerals (such as magnetite) found in the earth's interior rock bodies were first analyzed by the Chinese, while Sweden began exploring local anomalies of iron-ores in the direction of the magnetic field in 1640 and continued until the end of the century. Even though the majority of rock-forming minerals are practically non-magnetic, certain rock types include enough magnetic minerals to cause considerable magnetic anomalies. Magnetic anomalies can also be produced by man-made ferrous artifacts. In general, most of the magnetic effects are created from the igneous and metamorphic rocks which are characterized by higher magnetic susceptibilities, and low magnetic susceptibility is usually associated with sedimentary rocks. The magnetic susceptibilities of various rocks and ore minerals are depicted in further depth (Dobrin, 1976; Milsom, 2003).

Moreover, a wide range of applications using the geophysical magnetic method is also provided. This can be performed either for small-scale purposes such as environment, engineering, and archaeology to determine the buried features of interest or in large-scale designs for regional exploration of subsurface structures. In fact, one of the most important applications of the magnetic method in water, gas, and oil explorations is mapping the basement surface, and therefore, the thickness of the overlying sedimentary strata, which acts as potential zones, can be predicted in significant amounts. Several applications of the geophysical magnetic method have been introduced by many authors to solve a variety of considerable issues such as (Studt, 1964; Hochsten and Hunt, 1970; Mekhemer et al, 2007; Rabeh and Miranda, 2008; Araffa, 2013; Abdelaziz, 2019; Araffa et al, 2019; Montsion et al, 2021). Using different types of magnetic instruments (for example Fluxgate magnetometer), the magnetic surveys can be

executed on land, at sea, and in the air. Magnetic data processing is subsequently improved to correct the measured datasets that have been corrupted by the effects of the natural magnetic field of the earth. The magnetic data that has been enhanced by proposed processing procedures can be displayed in the forms of profiles or contour maps as a function of total intensity, relative intensity, and vertical or horizontal gradient anomalies. More details about the processing and interpretations approaches of magnetic datasets can be found in a wide variety of geophysical literature (Dobrin, 1981; Reynolds, 1997; Sharma, 1997; Kearey et al, 2002; Milsom, 2003; Hinze et al, 2013).

4.2. REDUCTION TO POLE OF THE EARTH'S MAGNETIC FIELD

Reduction to the pole (RTP) is considered to be one of the most common procedures in magnetic data processing, particularly for large-scale mapping. Since the Earth's geomagnetic field is characterized by its dipolar nature, any magnetic body buried beneath the ground, especially located at the intermediate latitudes, can produce an anomaly consisting of two parts (positive and negative). The exact location of this causative body lies between these two parts (Dobrin, 1981). This property changes the shape of the magnetic field due to the inclination and orientation of the induced magnetization vector from the magnetic poles to the equator and causes some difficulties in detecting the exact shapes and locations of the magnetized causative sources. This problem can be only solved when the total magnetic field is reduced to the northern or southern Earth's poles or equator to avoid the inclination and polarity effect (Baranov and Naudy, 1964; Blakely, 1996). In that regard, the reduction to the pole (RTP) provides the conversion of data carried out in the inclined Earth's magnetic field to that as one would measure at the magnetic pole. As a result, the anomaly can be located directly above the center of the causative body with the assumption that remanent magnetism is minor in comparison to induced magnetization. It means that the reduction to pole enables us to detect the anomaly source position more accurately. Moreover, it is generally observed that at low latitudes such as the geomagnetic equator, the processing of the measured datasets concerning the reduction to the pole is suffering from a complicated instability problem and in this case, the reduction to the equator is highly suggested (Luo Yao et al, 2010). In this chapter, the applicability of the two-dimensional inversion-based Fourier transformation in the field of reduction to the earth's magnetic pole is evaluated on 2D synthetic magnetic datasets sampled both equidistantly and non-equidistantly as well as the incomplete sampling design of missing data.

4.3. TESTING THE NOISE REDUCTION CAPABILITIES IN 2D

As described above for the processing of the 1D synthetic datasets in chapter 3, the noise rejection capacities of both the conventional Fourier transformation (2D DFT) and the inversion-based 2D Fourier transformation (2D IRLS-FT) approaches are tested on equidistant synthetic magnetic datasets sampled in two-dimension. The previously discussed 2D algorithm of the inversion method in chapter 2 was developed by the University of Miskolc, Department of Geophysics. Several articles have been already introduced concerning the application of the two noise filtering techniques to 2D magnetic datasets sampled both equidistantly and non-equidistantly (Dobróka et al, 2017; Nuamah and Dobróka, 2019; Abdelaziz and Dobróka, 2020; Nuamah et al, 2021). Therefore, in this PhD dissertation, 2D magnetic data processing is briefly provided through a comparison between the two proposed methods in the framework of reduction to the earth's magnetic pole.

4.3.1. SYNTHETIC 2D MAGNETIC DATA GENERATION

For magnetic data processing, the method constructed by Kunaratnam (1981) was applied to calculate two-dimensional synthetic magnetic datasets. In this investigation, the synthetic data are generated above an FT-shaped magnetic anomaly with a surface ranging from -100 m to +100 m in both x and y coordinates. The proceeded magnetic anomaly is created using a total magnetic field of 200 nT. To do so, the values of the magnetic inclination and declination, which are selected depending mainly on the geographical locations, are 63° and 3° respectively (for a hypothetical Hungarian location). Furthermore, the measuring data points are equidistantly sampled with an interval of 5 m spacing in both x and y directions (grid cell size is 5m). It means that 1681 total measurements are subjected to the traditional Fourier transformation (2D DFT) and the inversion-based 2D Fourier transformation (2D IRLS-FT) method for achieving such an aim of magnetic data filtering. *Figure 31* shows a total intensity magnetic anomaly map of the noise-free datasets. A close examination of this map demonstrates that the FT-shaped magnetic body is characterized by the presence of relatively high and low magnetic anomalies of different magnitudes varying from -100 nT to 250 nT.

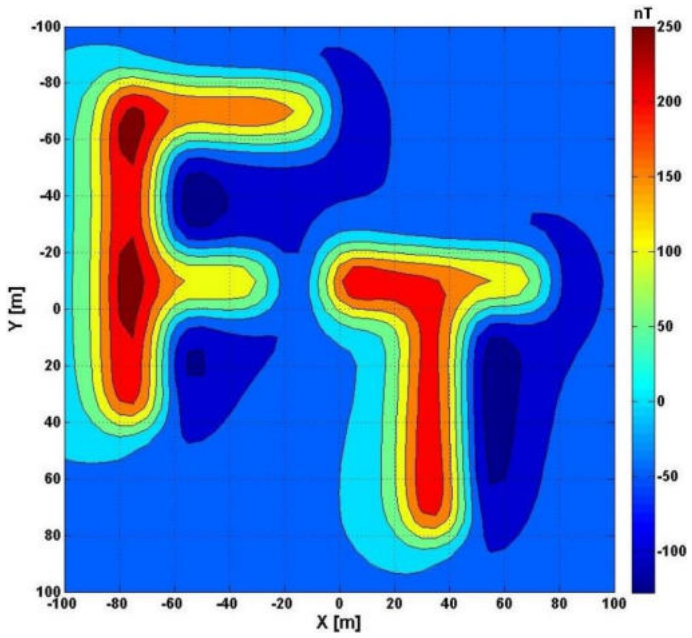

Figure 31. *2D noise-free synthetic magnetic datasets for the regularly sampled grid.*

4.3.2. REDUCING OUTLIER SENSITIVITY IN 2D IRLS-FT

For signal processing and testing how effective the inversion-based 2D Fourier transformation (2D IRLS-FT) method is for outlier sensitivity reduction, the above-mentioned noise-free magnetic datasets *(Figure 31)* are initially sampled equidistantly through the distribution of 41 x 41 surface observation locations as demonstrated in *Figure 32* by the regular sampling interval (5 m in *x* and *y* coordinates).

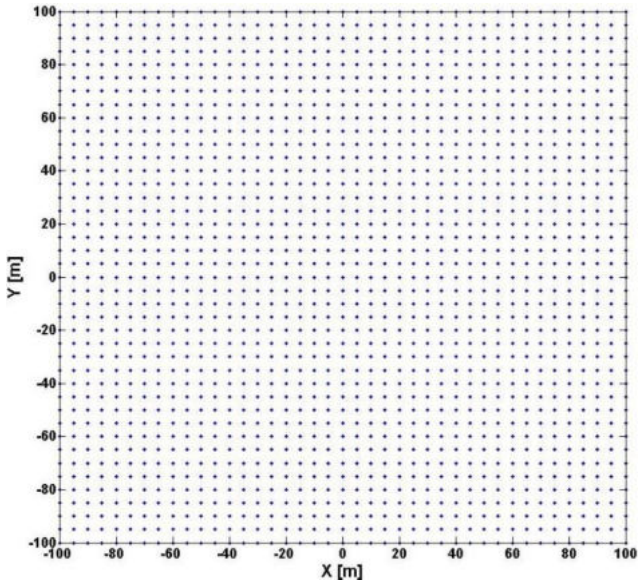

Figure 32. *The regular grid of the measuring magnetic stations.*

The 2D algorithm of the traditional Fourier transformation (2D DFT) is used to convert generated magnetic datasets *(Figure 31)* from the space to the frequency domain, resulting in a

2D amplitude spectrum of noise-free data as illustrated in *Figure 33*. To obtain a suitable scale in the wavenumber domain, the limits of x and y coordinates were simplified to be from – 0.1 to 0.1 (the Nyquist interval of the spectrum).

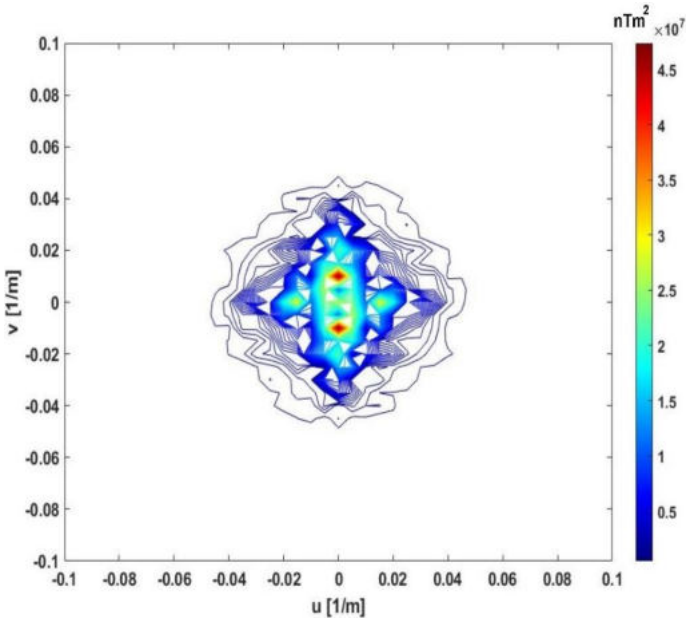

Figure 33. *The 2D amplitude spectrum of the regularly sampled noise-free magnetic data using DFT.*

The magnetic data interpretation can be improved by reducing the observations to the pole ($I = 90°$) in the frequency domain using the following formula:

$$R(u, v) = T(u, v)\, S(u, v), \tag{91}$$

where $T(u, v)$ is the 2D Fourier transform of the data and $S(u, v)$ is the frequency domain operator of the pole reduction (Blakely, 1996). The calculated data in the form of an RTP (magnetic map reduced to the pole) *(Figure 34)* can be then obtained utilizing the 2D inverse Fourier transformation of the spectrum given in Eq. 91 containing the IRLS FT spectrum of the dataset. It is obvious that the negative anomalies are cancelled in the RTP map, its data distance from the original (non-reduced to pole) map is 0.0109.

To test the noise reduction capability of the inversion method and also for simulating the real datasets, the generated synthetic magnetic data are contaminated with random noise of Cauchy distribution using a scale parameter of 0.03 as seen in *Figure 35*. It is observed that the Cauchy noise-contaminated data is distorted and displays more spikes compared to the noise-free datasets *(Figure 31)*, resulting in quite higher magnetic values (from -300 nT to 400 nT).

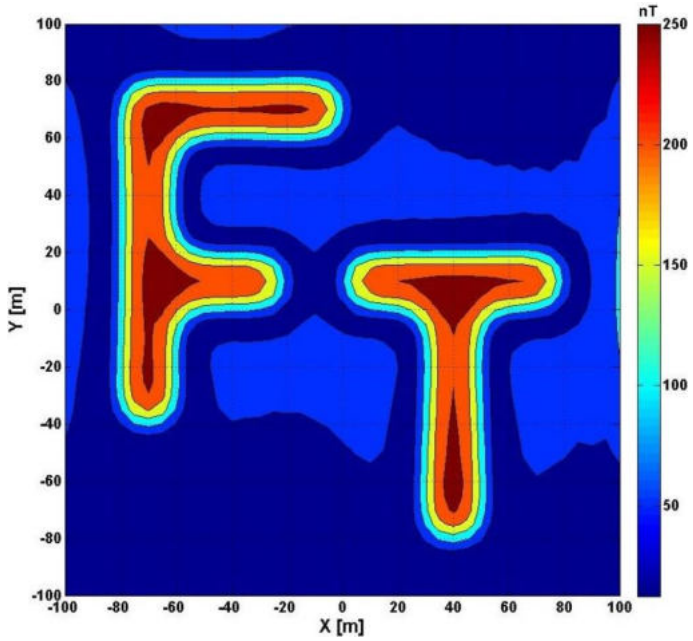

Figure 34. *The reduced-to-pole of the regularly sampled noise-free magnetic data sets using DFT.*

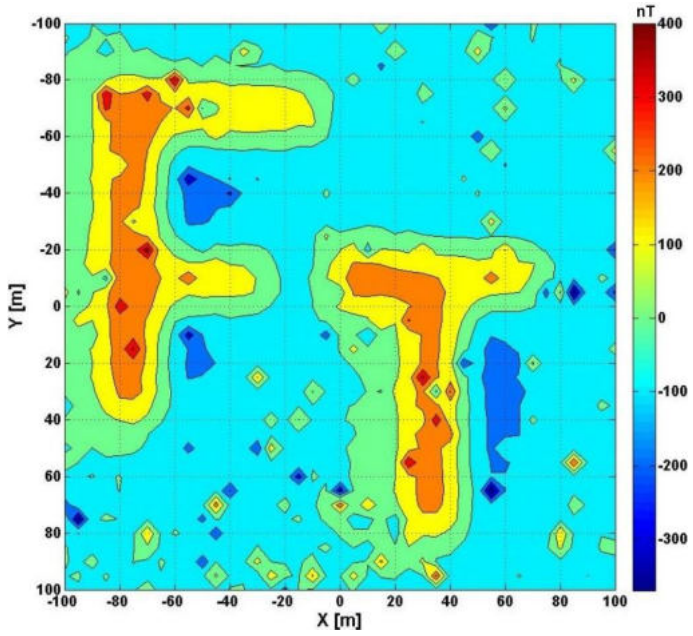

Figure 35. *2D noisy synthetic magnetic datasets for the regularly sampled grid.*

Similarly, the traditional Fourier transformation (2D DFT) is applied to the regularly sampled noisy magnetic datasets to obtain the 2D Fourier spectrum as presented in *Figure 36*. The number of samples in data space and frequency spaces are the same, so the inhomogeneous linear algebraic set of equations is well-determined. This results in a direct projection of the noises from the data space to the frequency space. Compared to the 2D spectrum of the noise-free data *(Figure 33)*, the noisy amplitude spectrum estimated by the DFT is deformed, reflecting its quite high sensitivity to the added noise.

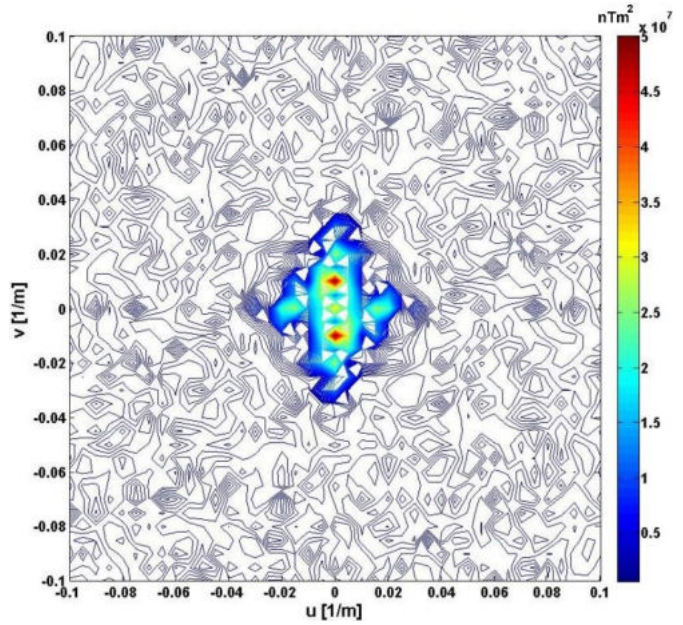

Figure 36. *The 2D amplitude spectrum of the regularly sampled Cauchy noise-contaminated data using DFT.*

Using the DFT method, the 2D inverse Fourier transform is applied to the noisy data showing the reduced-to-pole space domain data (RTP) in *Figure 37*. The calculated pole reduction of the noisy magnetic datasets is extremely distorted which is clearly evidenced by the remains of the added Cauchy noise simulating the outlier effect as well as the higher magnetic values (from 50 nT to 450 nT) compared to the generated noisy magnetic datasets *(Figure 35)*. In such a case, the mean data distance records an extremely high number of 0.4831, revealing the low noise rejection capability of the 2D DFT, especially in the Cauchy noise case.

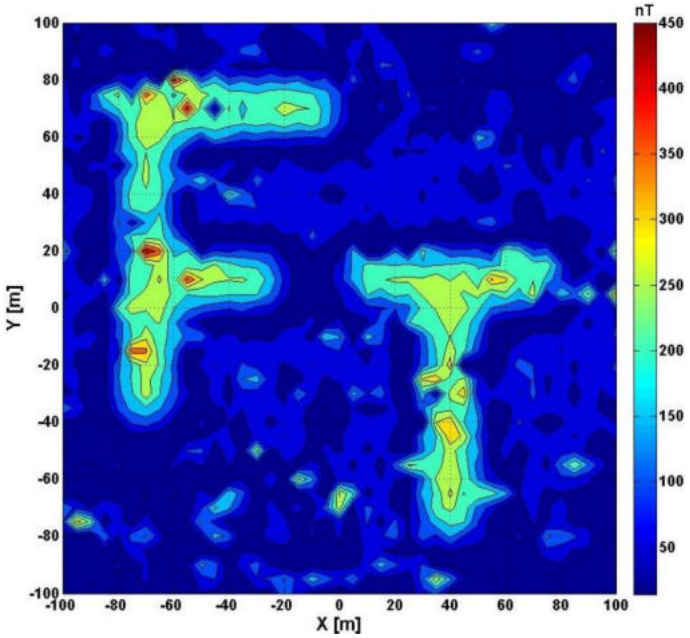

Figure 37. *The reduced-to-pole of the regularly sampled Cauchy noise-contaminated data using DFT.*

At this stage the 2D algorithm of the inversion-based Fourier transformation method (2D IRLS-FT) is implemented on the same generated noisy data *(Figure. 35)*, resulting in the 2D amplitude spectrum as shown in *Figure 38*. This spectrum is estimated using Hermite functions of order M_x=M_y=25 as an inversion parameter to achieve both inversion accuracy and stability. The total number of the unknown expansion coefficients is 625, and the number of the data is 1681, so the inverse problem is highly overdetermined resulting in a reduced level of projecting of the noises from the data space to the frequency space. The results of inversion prove sufficient improvements where most of the defects that appeared with DFT *(Figure. 36)* have been removed to be nearly similar to that of the noise-free spectrum *(Figure 33)*.

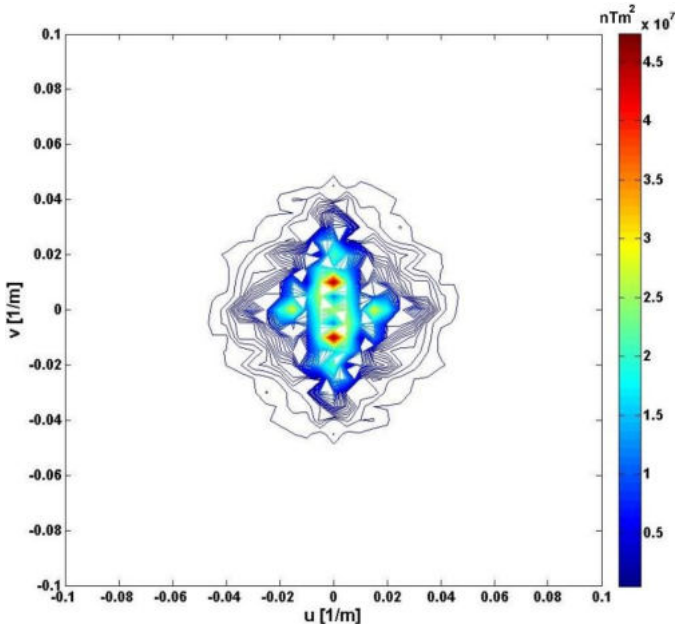

Figure 38. *The 2D amplitude spectrum of the regularly sampled Cauchy noise-contaminated data using IRLS-FT.*

To prove the inversion success, the reduction to the pole of the same noisy magnetic datasets is calculated using the 2D Fourier transformation of the IRLS, as depicted in *Figure 39*. It is indicated that the reduced-to-pole space domain data by the IRLS-FT is extremely improved and enhanced compared to the outliers problem associated with the noisy RTP magnetic map of the DFT method *(Figure. 37)*. The RTP map of *Figure 39* is almost identical to that of the noise-free reduction to pole space domain data (*Figure 34*) in both anomaly shape and magnetic amplitudes. The improvement is also evidenced by a relatively low data distance value of 0.0132 compared to the noise-free DFT used RTP map. According to the results, the iteratively reweighted least-squares Fourier transformation (2D IRLS-FT) method is outlier-resistant and quite robust, demonstrating great success in the framework of noise rejection capability for processing magnetic data measurements.

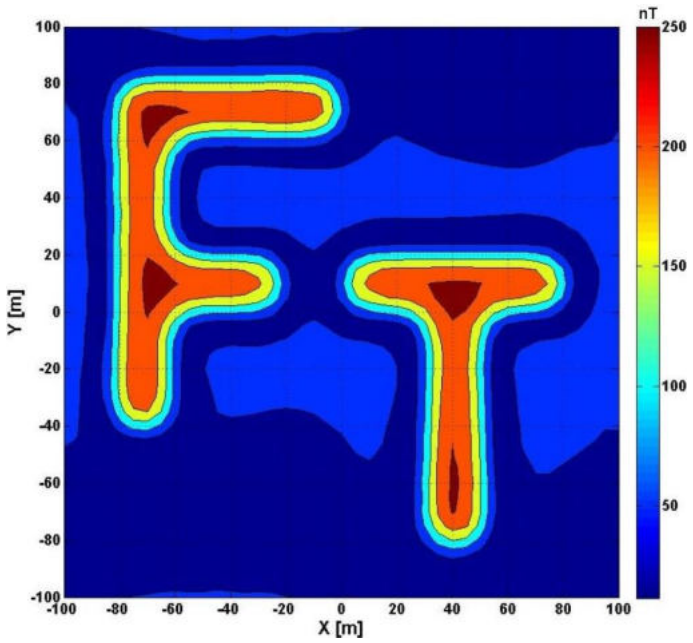

Figure 39. *The reduced-to-pole of the regularly sampled Cauchy noise-contaminated data using IRLS-FT.*

4.4. TESTING THE NON-REGULAR SAMPLING PROBLEM IN 2D

In geophysical surveys such as magnetic data acquisition, one gets large data sets ranging from 10,000 to 1 million measuring points with highly irregular sampling patterns. The challenge is to efficiently interpolate or estimate the measurements into the grid. Treating such sampling configurations necessitates the use of rapid and reliable algorithms as well as the consideration of extra information such as statistical features or the underlying physical process. In that regard, the two-dimensional algorithm of the inversion-based Fourier transformation (2D IRLS-FT) was developed to enhance the data quality and solve some of the problems accompanied by non-regular sampling. With non-equidistant specimen magnetic datasets, Nuamah and Dobróka (2019) proved the applicability and efficiency of the 2D inversion approach for pole reduction of CL-shaped magnetic bodies. Here, we use the above-mentioned example of the FT-shaped magnetic body to show briefly the stability and accuracy of the IRLS inversion approach with the non-regular sampling procedures.

4.4.1. COMPLETE MAGNETIC DATASET WITH RANDOM MEASUREMENT LOCATIONS

In this section, the same number of magnetic data points (1681 measurements) are randomized along with a non-regular grid in x and y directions. This sampling strategy of the random measurement positioning is performed through a random shifting process of observing

data points around the exact positions by half sampling interval in both coordinates. The non-equidistant distributions of measured data points are represented in *Figure 40*.

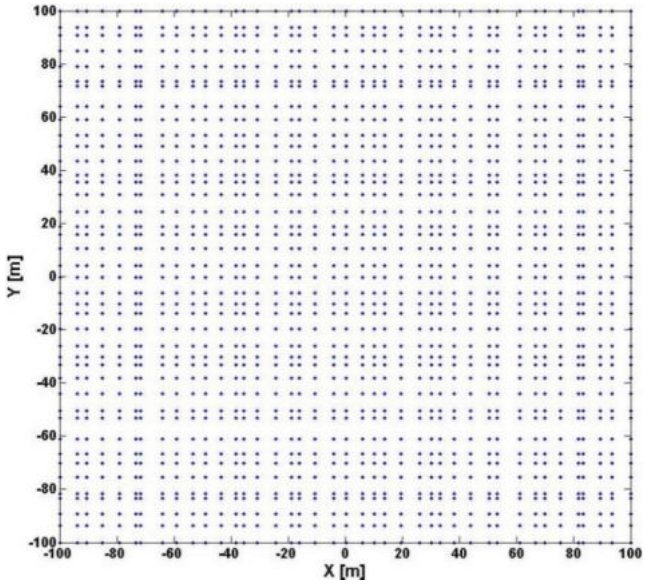

Figure 40. *A plot of non-equidistantly sampled magnetic measuring points.*

The non-regular noise-free sampling data points are subjected to the inversion-based FT algorithm utilizing the same order of Hermite functions (M=25 in both *x* and *y* directions) to establish the 2D Fourier frequency spectrum as illustrated in *Figure 41*. It is demonstrated that the non-regular IRLS spectrum is nearly similar to those of the regularly sampled noise-free spectra by the DFT and inversion-based FT (*Figure 33*).

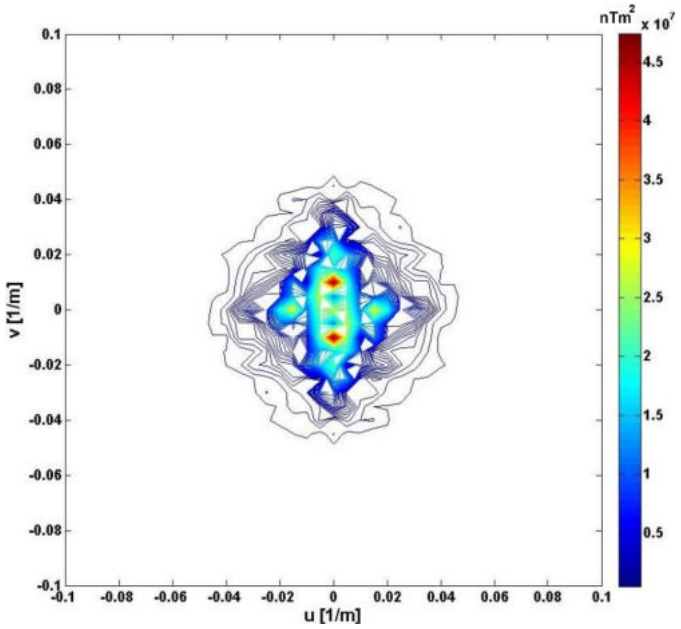

Figure 41. *The 2D amplitude spectrum of the non-regularly sampled magnetic data using IRLS-FT.*

In the same way, the reduced-to-pole magnetic data in the space domain is obtained using the 2D inverse Fourier transformation of the non-regular spectrum in Eq. 91 as shown in *Figure 42*. Comparing this pole reduction model to those found in the noise-free regular case by the DFT *(Figure 34)* as well as the IRLS inversion of the noisy data *(Figure 39)*, great similarities are established. This can also be proved quantitatively by its low mean distance value of 0.0147. It dedicates the higher applicability and effectiveness of the 2D inversion-based Fourier transformation (2D IRLS-FT) algorithm when we deal with complete datasets sampled non-equidistantly.

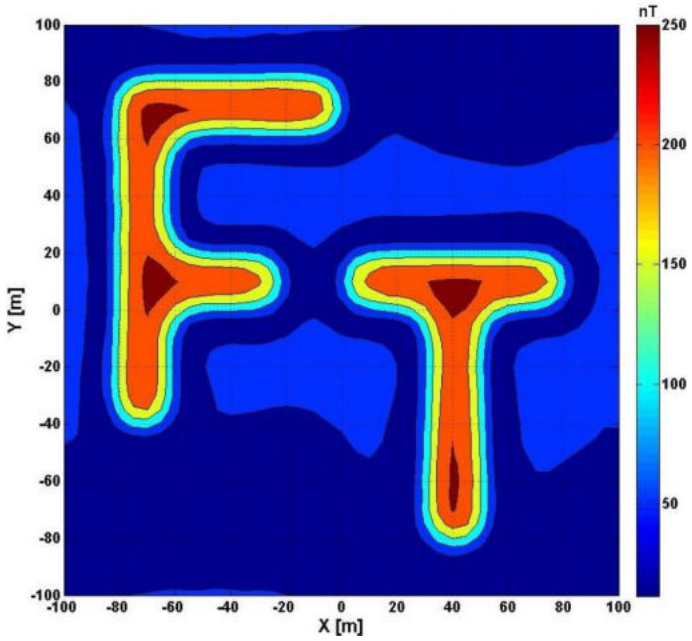

Figure 42. *The reduced-to-pole of the non-regularly sampled magnetic data using IRLS-FT.*

4.4.2. COMPLETE MAGNETIC DATASET WITH RANDOM WALK MEASUREMENT POSITIONS

4.4.2.1. The Concept of Random Walk Measuring

The geophysical survey of random walk measurements procedure that is carried out over non-equidistant grids is provided due to the restrictions associated with the unpaved areas under investigation or advances in the survey's equipment. Although for geophysical field observations, regular sampling across the explored areas has traditionally been the norm, most geophysical datasets, particularly magnetic and gravitational measurements, are frequently carried out along with irregularly sampled grids, making survey design a vital step in both obtaining enough data at a low cost and avoiding difficulties associated to survey spacing. This can also help to speed up the data collection procedures in the field (Dimri and Srivastava, 2007). The design of the survey network plays an important role in maximizing the number of

data points and improving data quality. The objective of the survey, the size and shape of the anomalies of interest, the amount of noise in the magnetic field, the size of the area being explored, and available instrumentation are all factors to consider while designing the survey. Since the magnetic and/or gravitational data acquisition is occasionally carried out in unpaved investigation areas, which may be the result of natural or man-made causes, we are sometimes forced to measure the data irregularly (randomly selected measuring points). Problems with this kind of data sampling are rarely rectified during processing, hence they must be resolved only during the survey design phase. Therefore, in our investigation, we employ the above-mentioned inversion-based 2D Fourier transformation (IRLS-FT) method to random walking magnetic measurements to assess its processing capability in the framework of the earth's magnetic pole reduction of 2D synthetic magnetic datasets.

4.4.2.2. The 2D Application

To verify the random walk sampling of 2D magnetic datasets, the regular measurements (1681 observing points) stated above are totally randomized over the tested area in both *x* and *y* directions. *Figure 43* shows the random spread of the magnetic measuring stations. It is seen that all the observing points are separated from each other by irregular or random spacing intervals.

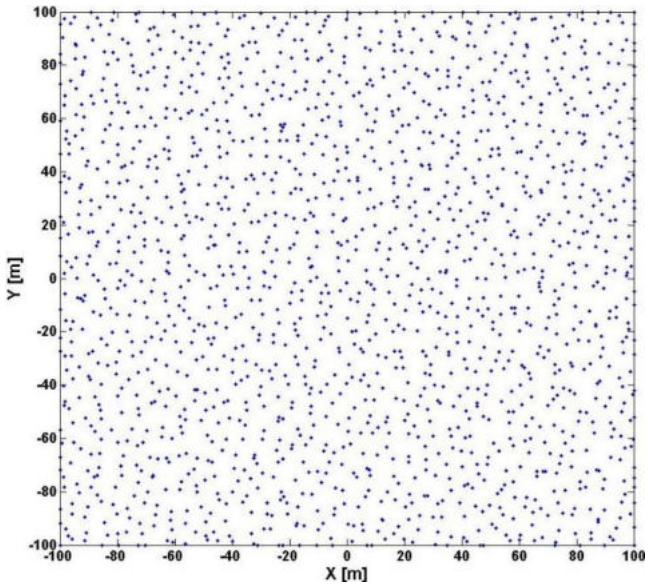

Figure 43. *A plot of the randomly sampled magnetic measuring points.*

The total intensity magnetic anomaly map produced by the randomly sampled magnetic measurements is presented in *Figure 44*. The irregularity of the measuring intervals causes the magnetic contour lines to be sharper with an irregular surface that does not cover the whole tested area when compared to the magnetic map of the regular surface *(Figure 31)*.

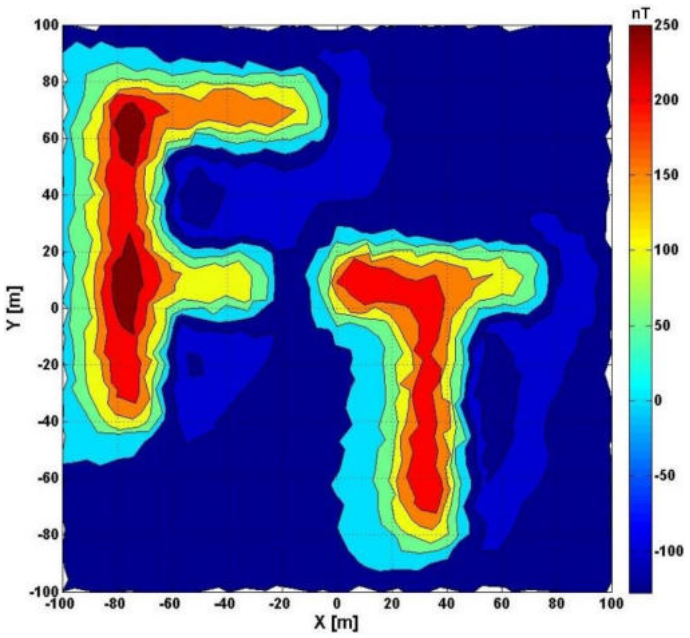

Figure 44. *2D synthetic magnetic datasets for the randomly sampled grid.*

Employing the 2D inversion algorithm to the randomly sampled magnetic datasets using (M_x=M_y=25) provides a 2D Fourier frequency spectrum as displayed in *Figure 45*. In this case of randomly sampled datasets, the calculated spectrum indicates a great similarity in both anomaly shapes and amplitudes when compared to those of the regular sampling found by the DFT and IRLS inversion approaches (*Figure 33* and *Figure 38* respectively) as well as the non-equidistantly estimated IRLS spectrum *(Figure 41)*.

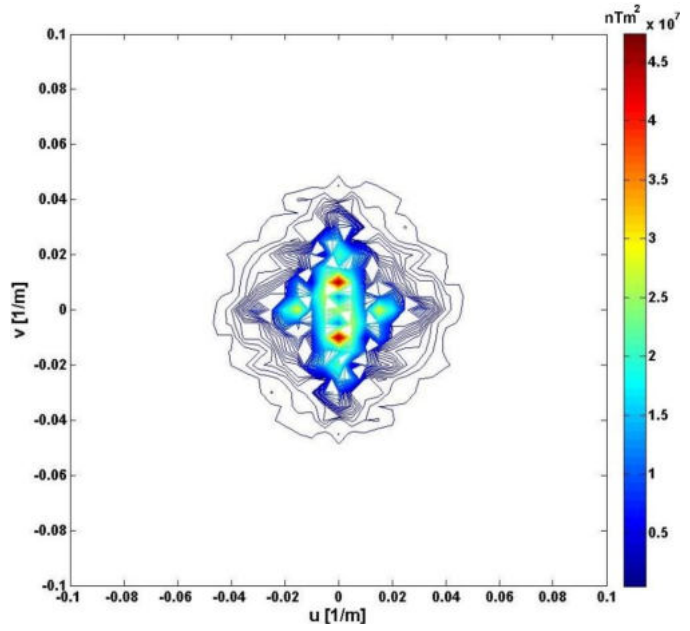

Figure 45. *The 2D amplitude spectrum of the randomly sampled magnetic data using IRLS-FT.*

The 2D inverse Fourier transformation is subsequently applied to the spectrum given in Eq. 91 (containing the IRLS-FT of the randomly sampled magnetic data) to establish a reduced to pole magnetic data in random sampling geometry as illustrated in *Figure 46*. It is seen that the RTP magnetic map obtained by the inversion-based FT algorithm with random walk measurements returned to its original form as we would measure along with an equidistantly or regularly sampled grid. It is identical to those found by both the DFT and IRLS of regularly sampled datasets (*Figure 34* and *Figure 39* respectively) and that estimated by the IRLS inversion method in the non-equidistantly sampling case *(Figure 42)*. Additionally, the inversion method's effectiveness and reliability are numerically verified by a relatively low value of the mean distance of 0.0153. The developments and quality of the proceeded magnetic datasets sampled in a random walk configuration are quite enough to prove the stability, effectiveness, and higher applicability of the inversion-based 2D Fourier transformation, and therefore, it can be suggested to deal with other geophysical tools such as the gravitational method.

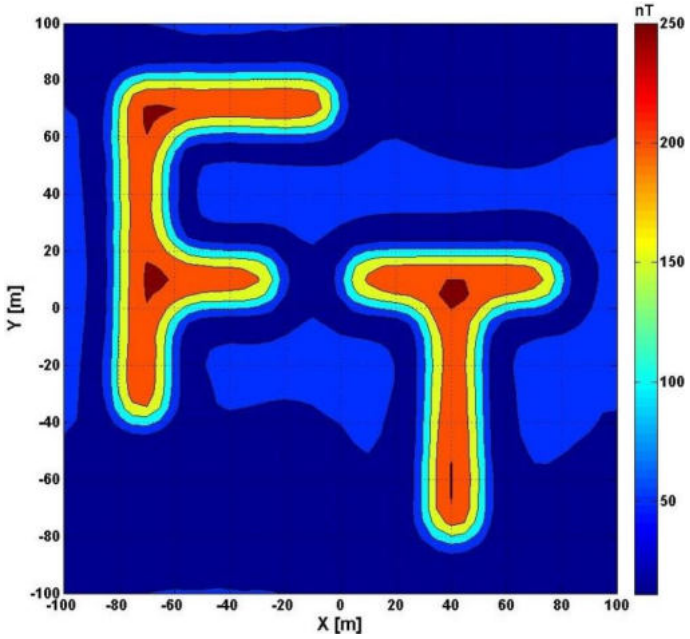

Figure 46. *The reduced-to-pole of the randomly sampled magnetic data using IRLS-FT.*

4.5. TESTING THE INCOMPLETE SAMPLING PROBLEM IN 2D

It is generally believed that the incomplete sampling problem is of great importance since the magnetic field data measurements are sometimes missing or lost due to several natural and/or artificial conditions as discussed before. Therefore a newly developed inversion-based 2D Fourier transformation MATLAB code is created to achieve such an aim of magnetic data processing. In that regard, regularly, non-equidistantly, and randomly sampled magnetic

datasets in 2D are all subjected to the inversion algorithm using different missing percentages. Both the noise-free and noisy datasets are presented for evaluating the inversion efficiency and stability. This new technique enables us to solve the unlikely problems associated with incomplete or missing data sets during the field acquisition procedures.

4.5.1. INCOMPLETE MAGNETIC DATASET WITH REGULAR SAMPLING

The initially regularly sampled complete magnetic data over 1681 measurements *(Figure 32)* are used to evaluate the inversion efficiency in different scenarios of the missing datasets. We start with a first example when 10 percent of the 2D magnetic datasets are randomly missing or cancelled from the entire data. *Figure 47a* shows the distribution of nearly 1512 magnetic observations after randomly removing 10% of the data in *x* and *y* directions. In this case, the newly developed inversion algorithm is implemented using Hermite functions of order (M_x=M_y=25) to establish the 2D Fourier frequency spectrum as illustrated in *Figure 47b*. The 2D inverse RTP map of the incomplete data set is presented in *Figure 47c*.

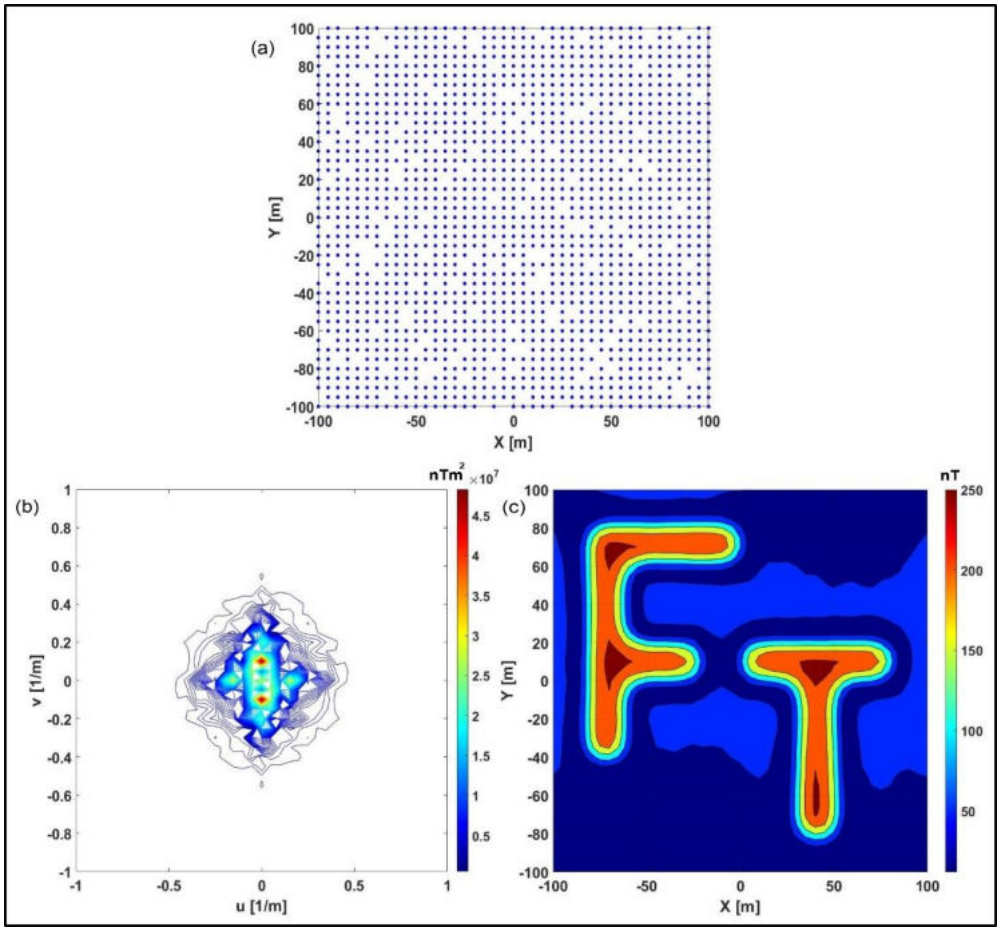

Figure 47. *a) A plot of regular sampling magnetic observations when 10% of the datasets are missing in the x and y-directions, b) The 2D IRLS-FT spectrum when 10% of the regularly sampled magnetic datasets are missing, c) The reduced-to-pole map using IRLS-FT when 10% of the regularly sampled magnetic datasets are missing.*

It is observed that both the estimated 2D spectrum and the RTP map are similar to those found by the traditional DFT method of the regularly sampled complete datasets (*Figure 33* and *Figure 34* respectively) as well as those obtained with the IRLS inversion (*Figure 38* and *Figure 39* respectively). In this case, the value of the mean distance is numerically computed to be 0.0073 reflecting the inversion effectiveness with a 10% missing date case. Furthermore, to demonstrate the stability and accuracy of the inversion approach, a higher percentage of missing datasets are subjected to the newly developed algorithm. *Figures 48a, b, c,* and *d* illustrate the 2D amplitude spectra by the IRLS method at 15%, 25%, 35%, and 45% missing datasets respectively. The total number of datasets used in 15% missing case is about 1417, 1271 for 25% while 1094 and 922 for the cases of 35%, and 45% missing data respectively. All the calculated IRLS spectra prove high similarity when compared to those of the regularly sampled complete datasets by the DFT and IRLS-FT (*Figure 33* and *Figure 38* respectively).

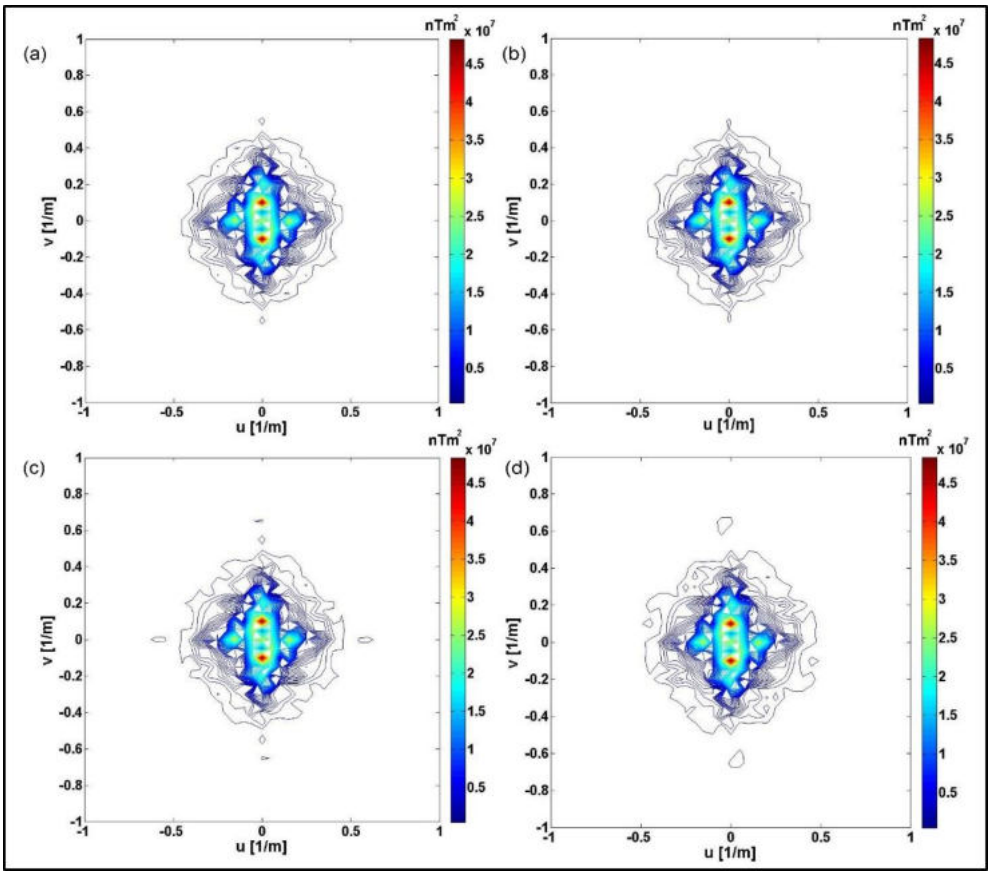

Figure 48. *The 2D IRLS-FT spectra when a) 15% of the regularly sampled magnetic datasets are missing, b) 25% of the regularly sampled magnetic datasets are missing, c) 35% of the regularly sampled magnetic datasets are missing, d) 45% of the regularly sampled magnetic datasets are missing.*

Similarly, the above-mentioned spectra are used in Eq.91 to produce the space domain pole reduction of 15%, 25%, 35%, and 45% missing datasets as shown in *Figures 49a, b, c,* and

d respectively. It is seen that when compared to the reduced-to-pole datasets obtained by the DFT and IRLS of the complete regular sampling case (*Figure 34* and *Figure 39* respectively), the RTP magnetic anomaly maps of the above missing data percentage demonstrate a great similarity. This qualification can be numerically proven by the mean distance values which are estimated as 0.0104, 0.0123, 0.0230, and 0.0487 for 15%, 25%, 35%, and 45% missing datasets respectively. In these cases of the missing datasets, the 2D inversion algorithm is effective and applicable because the inverse problem is still over-determined as indicated by the higher number of data points when compared to the number of the model parameters (M_x M_y=625).

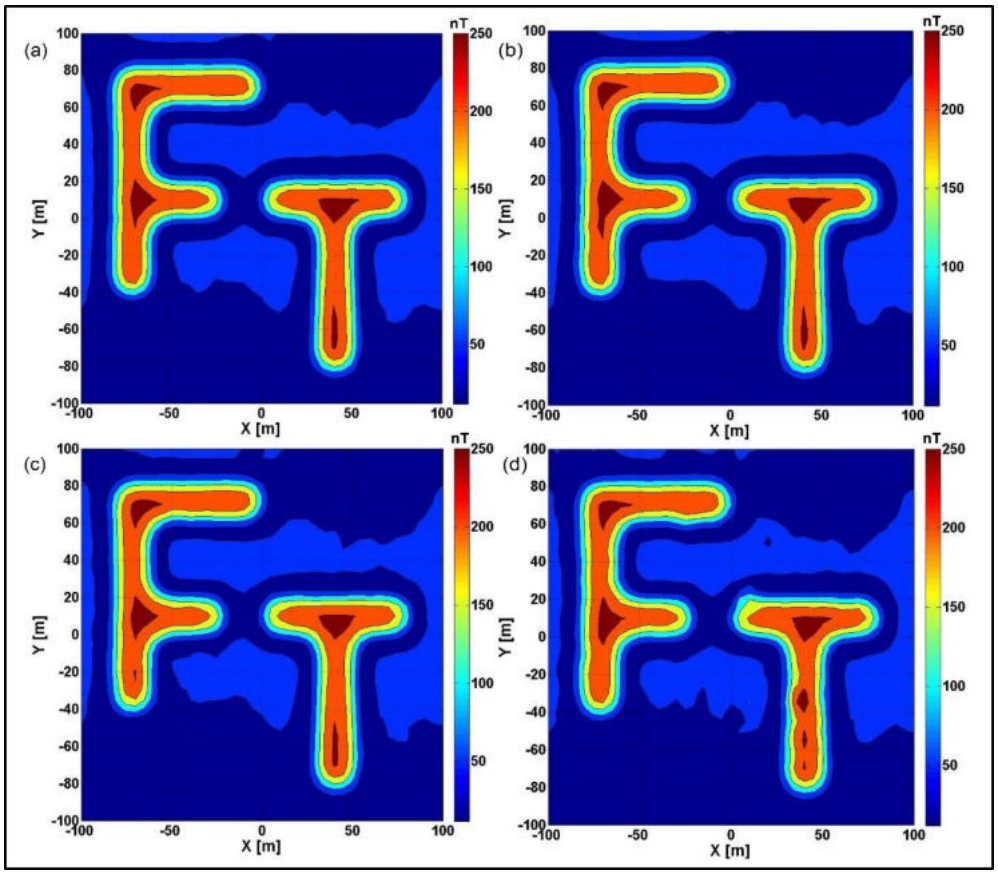

Figure 49. *The reduced-to-pole maps using IRLS-FT when a) 15% of the regularly sampled magnetic datasets are missing, b) 25% of the regularly sampled magnetic datasets are missing, c) 35% of the regularly sampled magnetic datasets are missing, d) 45% of the regularly sampled magnetic datasets are missing.*

On the other hand, *Figures 50a* and *b* depict the 2D spectrum and the RTP map respectively when 50% of the data are missing (usage of about 839 data points) while those estimated in 60% missing data case (usage of about 657 data points) are represented in *Figures 50c* and *d* respectively. It is observed that when 50% of the data are missing, the inversion still works but with a much higher mean data distance value (0.0743), almost seven times higher than in the 15% missing datasets. In contrast, both the 2D spectrum and the RTP magnetic map established

with the missing 60% of the datasets are distorted with a very high value of the mean distance of 0.4764. It is that by increasing the percentage of the missing datasets, the inverse problem becomes less and less over-determined where the model parameters (M_x M_y=625) are close to the number of data points (marginally overdetermined inversion). Accordingly, the newly developed inversion-based 2D Fourier transformation is effective and robust when processing regular sampling datasets even with 50% missing points as a maximum (The over-determination rate is 839/625=1.34). These investigations show a new (and economically meaningful) possibility in planning field measurements: the number of measurement points can be reduced before reaching an acceptable inaccuracy in determining the RTP map.

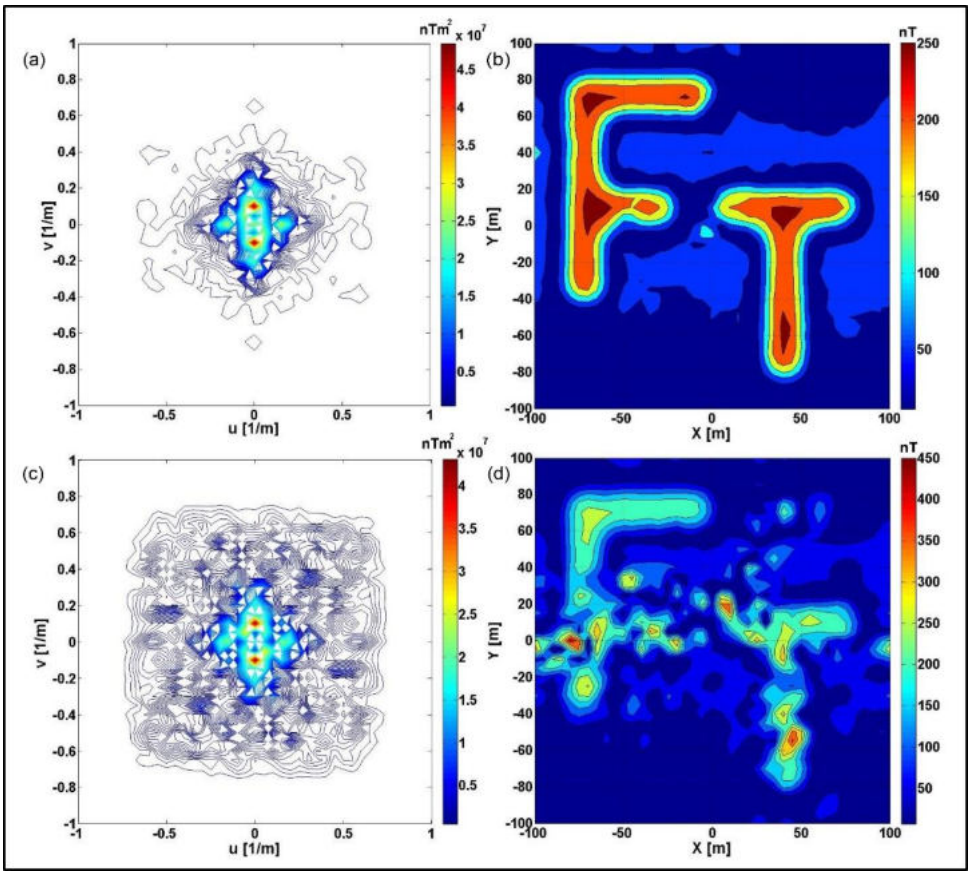

Figure 50. *a) The 2D IRLS-FT spectrum when 50% of the regularly sampled magnetic datasets are missing, b) The reduced-to-pole map using IRLS-FT when 50% of the regularly sampled magnetic datasets are missing, c) The 2D IRLS-FT spectrum when 60% of the regularly sampled magnetic datasets are missing, d) The reduced-to-pole map using IRLS-FT when 60% of the regularly sampled magnetic datasets are missing.*

4.5.2. INCOMPLETE MAGNETIC DATASET WITH NON-EQUIDISTANT SAMPLING

Similar to that of the regularly sampled datasets, the inversion-based FT method is applied to the non-equidistantly sampled incomplete magnetic datasets. This is to confirm the quality and applicability of the inversion approach in tackling this incomplete sample problem. The

same non-equidistant sampling magnetic datasets established previously for 1681 measurements with 5m random shifting of the observation points in *x* and *y* directions are utilized *(Figure 40)*. To compare the results to those obtained in regular sampling geometry, the non-equidistantly sampled incomplete datasets are subjected to the IRLS inversion algorithm at the same percentages of the missing magnetic observations. *Figures 51a, b, c,* and *d* demonstrate the 2D amplitude-frequency spectra at 15%, 25%, 35%, and 45% missing datasets. Hermite functions of order (M_x=M_y=25) are also selected to present such Fourier spectra. The random sampling procedure leads to about 1435 observing points in the 15% missing data case, 1274 for 25%, whereas the total number of the measuring stations in the 35% and 45% missing data circumstances is 1090 and 947 respectively. Hence, all the 2D non-equidistantly estimated spectra at the mentioned missing data percentages are nearly similar in anomaly shapes and amplitudes when compared to those found by the IRLS inversion method of the regularly sampled incomplete magnetic datasets at the same missing proportions *(Figure 48)*. These findings prove that the inversion procedure can give acceptable results even in such an incomplete sampling configuration.

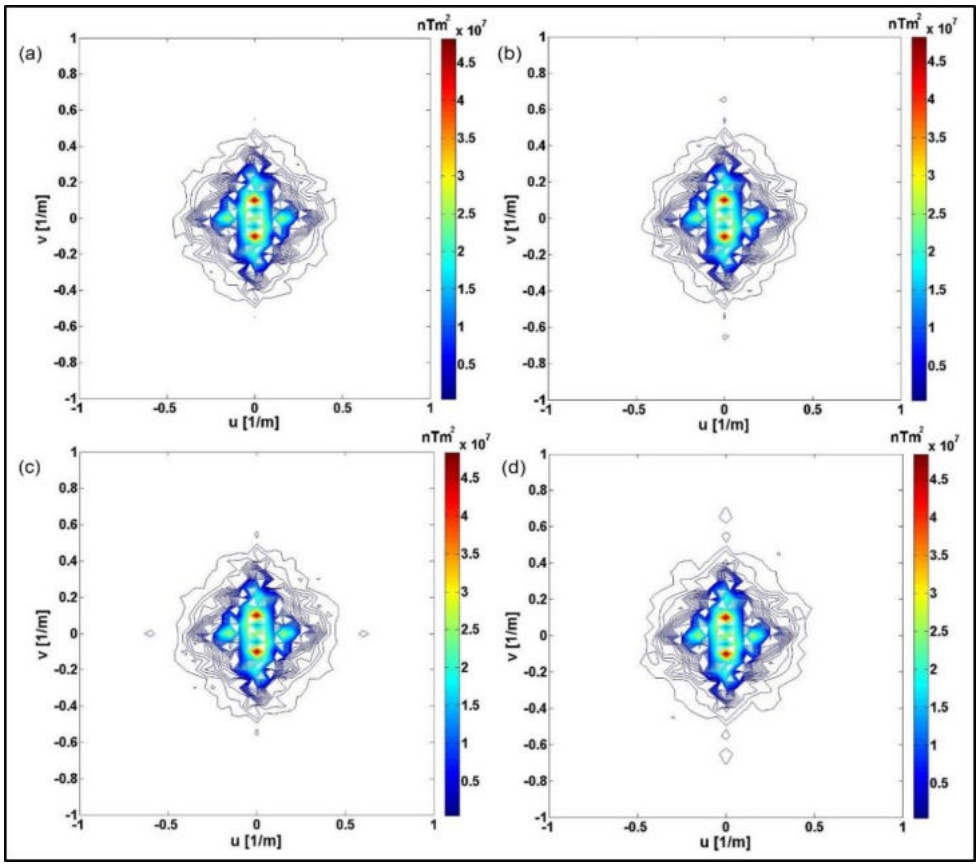

Figure 51. *The 2D IRLS-FT spectra when a) 15% of the non-equidistantly sampled magnetic datasets are missing, b) 25% of the non-equidistantly sampled magnetic datasets are missing, c) 35% of the non-equidistantly sampled magnetic datasets are missing, d) 45% of the non-equidistantly sampled magnetic datasets are missing.*

The 2D Fourier transformations of the above amplitude spectra are used in Eq. 91 giving the pole reduction maps of the non-equidistant magnetic data sampled incompletely with missing 15%, 25%, 35%, and 45% of measuring points respectively (the RTP maps in *Figures 52a, b, c,* and *d* respectively). Compared to the reduced-to-pole magnetic anomaly maps estimated in the incomplete magnetic data sampled regularly *(Figure 49)*, great similarities are provided. At this stage, the over-determined inverse problem is dominant where the number of the data points is still more than that of the model parameters (M_x M_y=625). The inversion efficiency is also quantitatively proven through the mean distance calculations. A value of 0.0107 is estimated in the 15% missing data case and 0.0135 for 25% missing case whilst the 35% and 45% missing data cases are represented by mean distance values of 0.0297 and 0.0435 respectively.

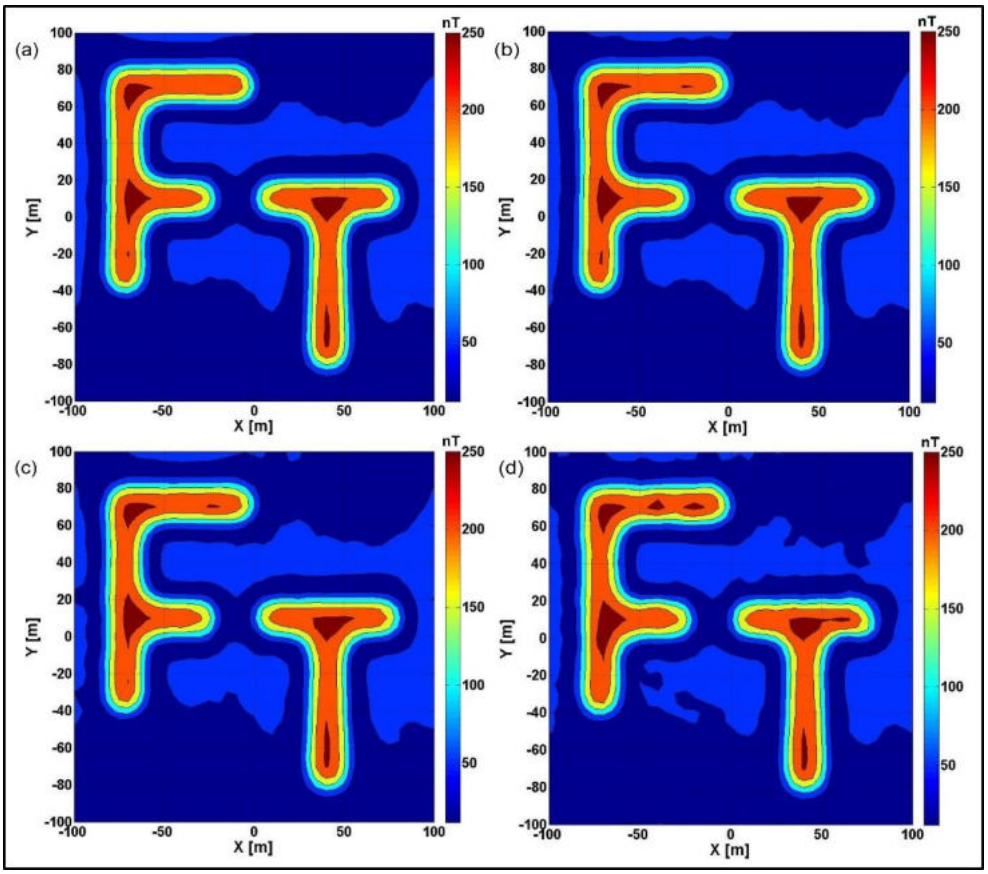

Figure 52. *The reduced-to-pole maps using IRLS-FT when a) 15% of the non-equidistantly sampled magnetic datasets are missing, b) 25% of the non-equidistantly sampled magnetic datasets are missing, c) 35% of the non-equidistantly sampled magnetic datasets are missing, d) 45% of the non-equidistantly sampled magnetic datasets are missing.*

To evaluate the inversion applicability at higher levels of missing data points, half of the magnetic measurements (50%) are randomly cancelled, resulting in a total of roughly 825 observing points on which the IRLS inversion algorithm is implemented. *Figures 53a* and *b* respectively demonstrate the 2D amplitude spectrum and the RTP magnetic anomaly map of the non-equidistant incomplete data that is missing 50% of its total power. It is estimated that both the spectrum and pole reduction data are roughly similar to those deduced by the IRLS inversion of the 50% missing datasets sampled regularly *(Figures 50a* and *b)*. In this case, the estimated value of the mean distance is 0.0709 which is approximately seven times higher than that of the non-equidistant 15% missing datasets.

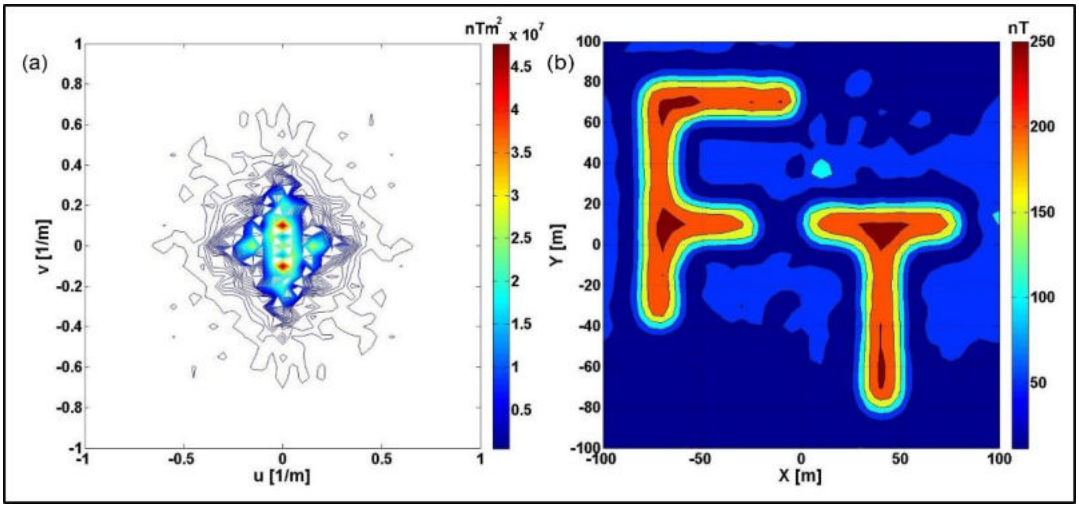

Figure 53. *a) The 2D IRLS-FT spectrum when 50% of the non-equidistantly sampled magnetic datasets are missing, b) The reduced-to-pole map using IRLS-FT when 50% of the non-equidistantly sampled magnetic datasets are missing.*

The missing data percentage is subsequently raised to 60%, and hence the 2D IRLS inversion approach is only performed on 641 measuring data points. It means that our inverse problem is sounded to be marginally over-determined where the observing data points used are nearly close to the number of the model parameters (M_x M_y=625). This is the primary cause of the inversion procedure's inefficiency and instability, which is exacerbated by larger rates of missing data proportions, as in our instance (60%).

Moreover, the 2D amplitude spectrum and the RTP magnetic anomaly map are shown in *Figures 54a* and *b* respectively. The high distortion appears in both anomaly shapes and amplitudes providing a very high value of the mean data distance of 0.3468. As concluded in the regularly sampled incomplete datasets, the reconstructed inversion method is satisfactory and highly applicable to the incomplete magnetic datasets sampled non-equidistantly up to 50% missing points as a maximum. This is because the inverse problem is constrained by the over-

determination value which in our case is calculated as 825/625=1.32. Accordingly, our inversion-based FT can be recommended to deal with the incomplete random walk measurement positioning.

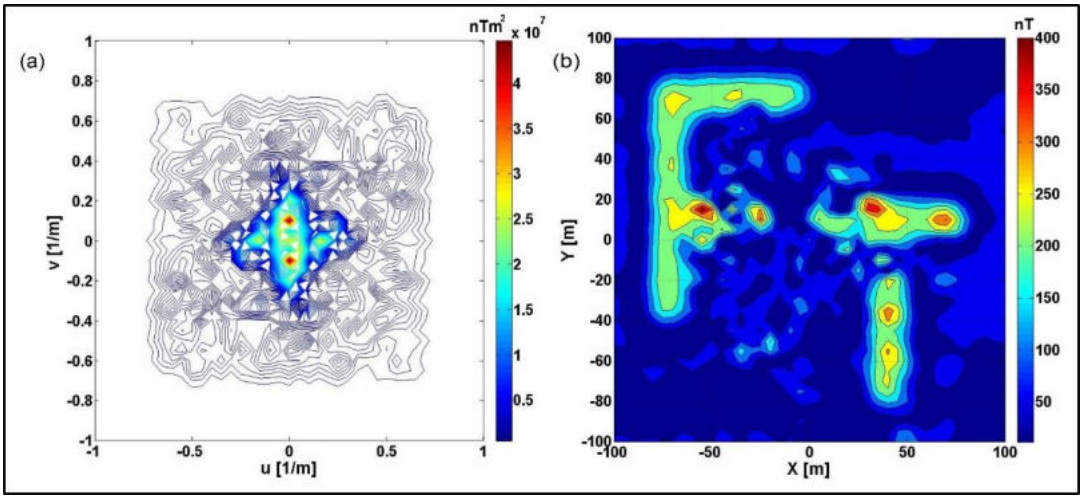

Figure 54. *a) The 2D IRLS-FT spectrum when 60% of the non-equidistantly sampled magnetic datasets are missing, b) The reduced-to-pole map using IRLS-FT when 60% of the non-equidistantly sampled magnetic datasets are missing.*

4.5.3. INCOMPLETE MAGNETIC DATASET WITH RANDOM WALK SAMPLING

The inversion efficiency and applicability to the incomplete magnetic datasets sampled both regularly and non-equidistantly motivated us to evaluate the 2D algorithm on the incomplete random walk measurement positions. The randomly distributed magnetic datasets taken over 1681 observations *(Figure 43)* are used to achieve such an aim.

For comparison, the 2D amplitude-frequency spectra estimated by the IRLS inversion method with missing 15%, 25%, 35%, and 45% of the data are shown in *Figures 55a, b, c,* and *d* respectively. The same order of the Hermite functions (M=25 in both *x* and *y*- directions) is chosen. 1426 measuring points are utilized for the 15% missing data case, while 1267, 1073, and 941 for the 25%, 35%, and 45% cases respectively. The results demonstrate great similarity when we compare the 2D spectra of the random walk measurements to those found by the inversion approach in both the regularly *(Figure 48)* and the non-equidistantly sampled incomplete magnetic datasets *(Figure 51).*

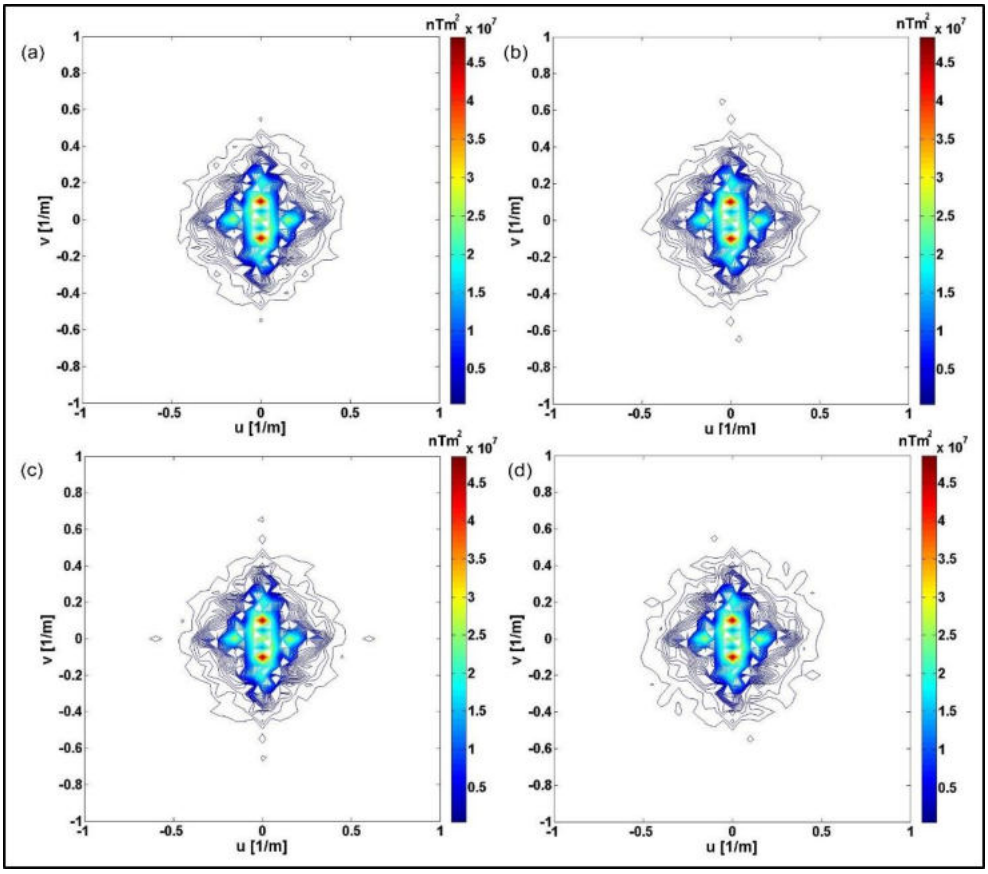

Figure 55. *The 2D IRLS-FT spectra when a) 15% of the randomly sampled magnetic datasets are missing, b) 25% of the randomly sampled magnetic datasets are missing, c) 35% of the randomly sampled magnetic datasets are missing, d) 45% of the randomly sampled magnetic datasets are missing.*

The reduced to pole magnetic anomaly maps calculated at 15%, 25%, 35%, and 45% missing data are illustrated in *Figures 56a, b, c,* and *d* respectively. It is noticed that the pole reduction findings are nearly similar to those provided with the incomplete magnetic datasets sampled regularly *(Figure 49)* and non-equidistantly *(Figure 52)* in the anomaly shapes and amplitudes. In the case of 15% missing data, the mean distance value is calculated as 0.0297 while the data that missing 25%, 35%, and 45% from the entire measurements are characterized by distances 0.0350, 0.0444, and 0.0603 respectively, which are generally a little bit higher than those calculated in both the regular and non-equidistant sampling configurations. These percentages of the missing data reflect the effectiveness of the inversion method where the inverse problem is over-determined. Similarly, the IRLS inversion method is applied to half of the magnetic measurements through missing 50% of the data, and hence 852 observing points are used. *Figures 57a* and *b* show the calculated 2D frequency spectrum and the reduced to pole map respectively.

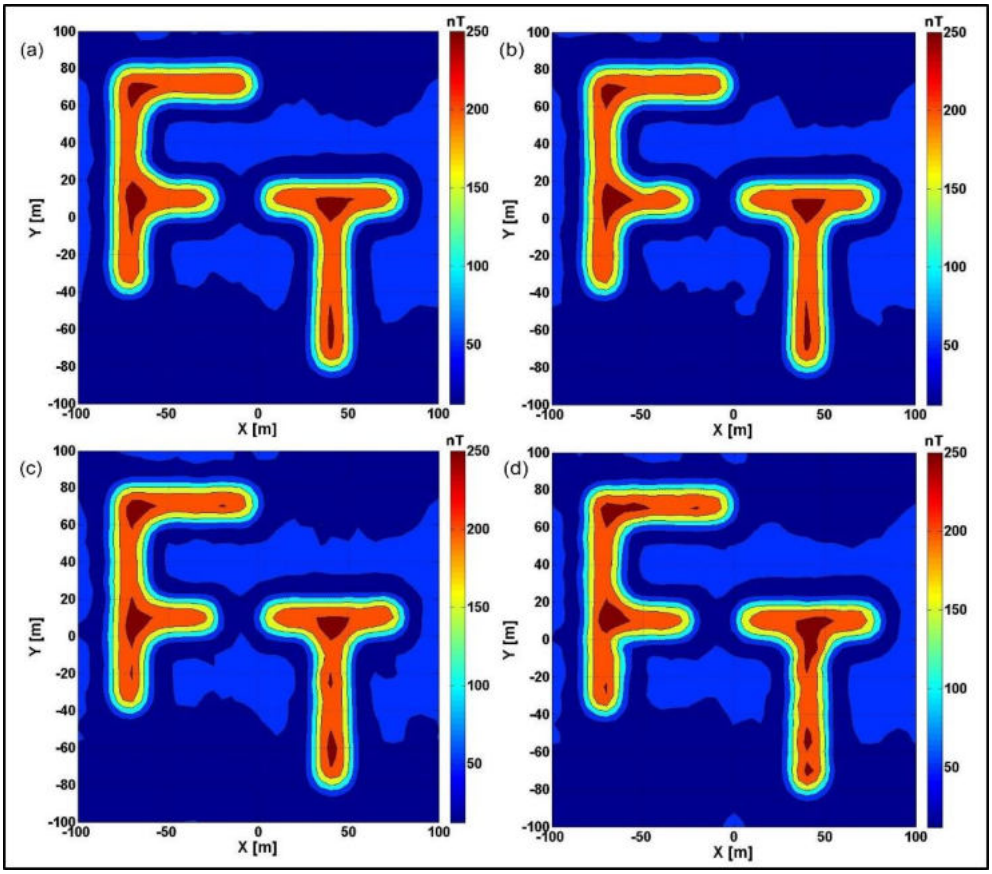

Figure 56. *The reduced-to-pole maps using IRLS-FT when a) 15% of the randomly sampled magnetic datasets are missing, b) 25% of the randomly sampled magnetic datasets are missing, c) 35% of the randomly sampled magnetic datasets are missing, d) 45% of the randomly sampled magnetic datasets are missing.*

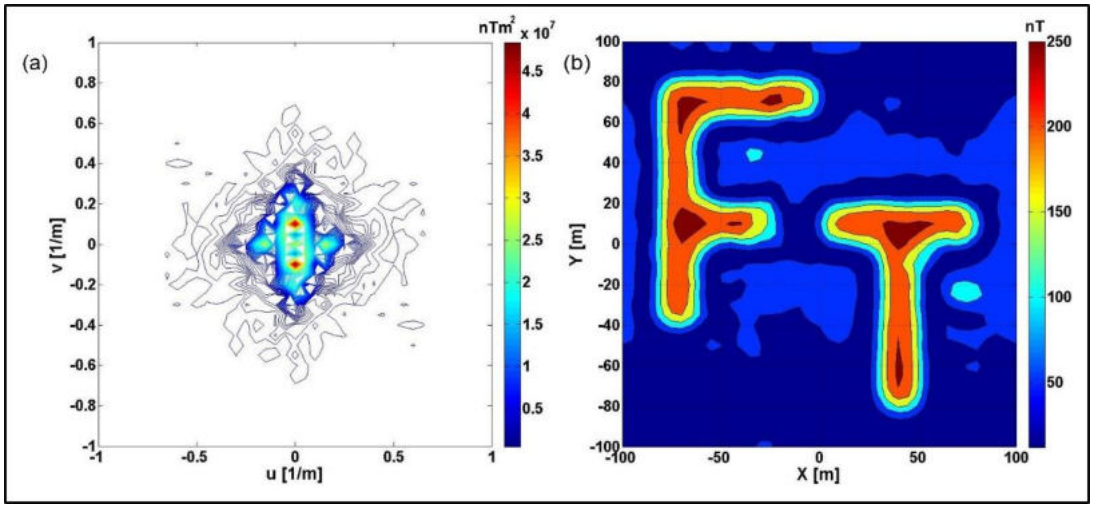

Figure 57. *a) The 2D IRLS-FT spectrum when 50% of the randomly sampled magnetic datasets are missing, b) The reduced-to-pole map using IRLS-FT when 50% of the randomly sampled magnetic datasets are missing.*

Sufficient improvements are proven when comparing both the spectrum and RTP map of 50% missing data to those obtained in both cases of the regularly *(Figures 50a* and *b*

respectively*)* and the non-equidistantly sampled incomplete magnetic datasets *(Figures 53a* and *b* respectively*)*. The data distance is numerically computed as 0.0827 which is almost seven times higher than in the 15% missing datasets. This is the maximum missing data percentage at which the inverse problem is completely overdetermined, as estimated by the over-determination ratio of 852/625=1.3632.

Furthermore, the incomplete datasets with missing 60% of their total power are subjected to the inversion method resulting in the 2D amplitude spectrum and the pole reduction map as seen in *Figures 58a* and *b* respectively. In this case, the number of the data points employed is roughly 638, which is nearly close to the number of the model parameters (M_x M_y=625), making our inverse problem at a marginal level of over-determination. As quantified by a very high distance value of 0.3160, it causes extreme distortion in the anomaly patterns and amplitudes of both the spectrum and the RTP magnetic map. According to these results, it may be stated that the inversion method's reconstructed 2D algorithm is effective, stable, and satisfactory for preprocessing incomplete magnetic datasets sampled at random and that it can be suggested to treat diverse data taken via other geophysical tools.

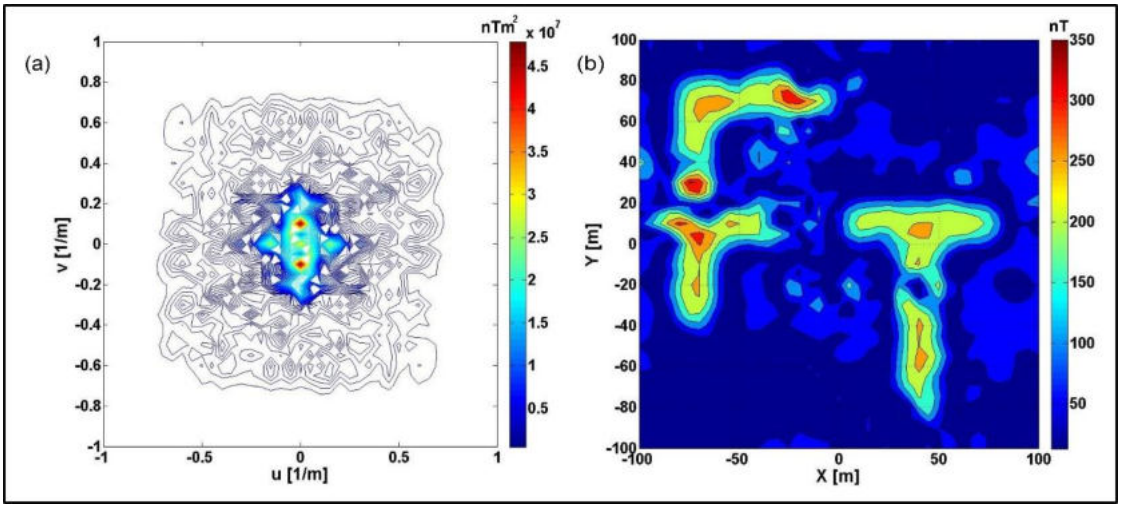

Figure 58. *a) The 2D IRLS-FT spectrum when 60% of the randomly sampled magnetic datasets are missing, b) The reduced-to-pole map using IRLS-FT when 60% of the randomly sampled magnetic datasets are missing.*

4.5.4. INCOMPLETE SAMPLING PROBLEM STUDIED ON NOISY MAGNETIC DATASETS

The incomplete sampling problem is examined on the same noisy magnetic datasets shown in *Figure 35* to assess the efficacy and stability of the newly proposed 2D IRLS-FT inversion technique. For comparison, the same missing data percentages (15%, 25%, 35%, 45%, 50%, and 60%) are chosen.

4.5.4.1. Study on Regular Grid

Firstly, the inversion is performed on the noisy datasets sampled over a regular grid. *Figures 59a, b, c, d, e,* and *f* show the 2D amplitude-frequency spectra at the above-mentioned percentages of the missing data respectively.

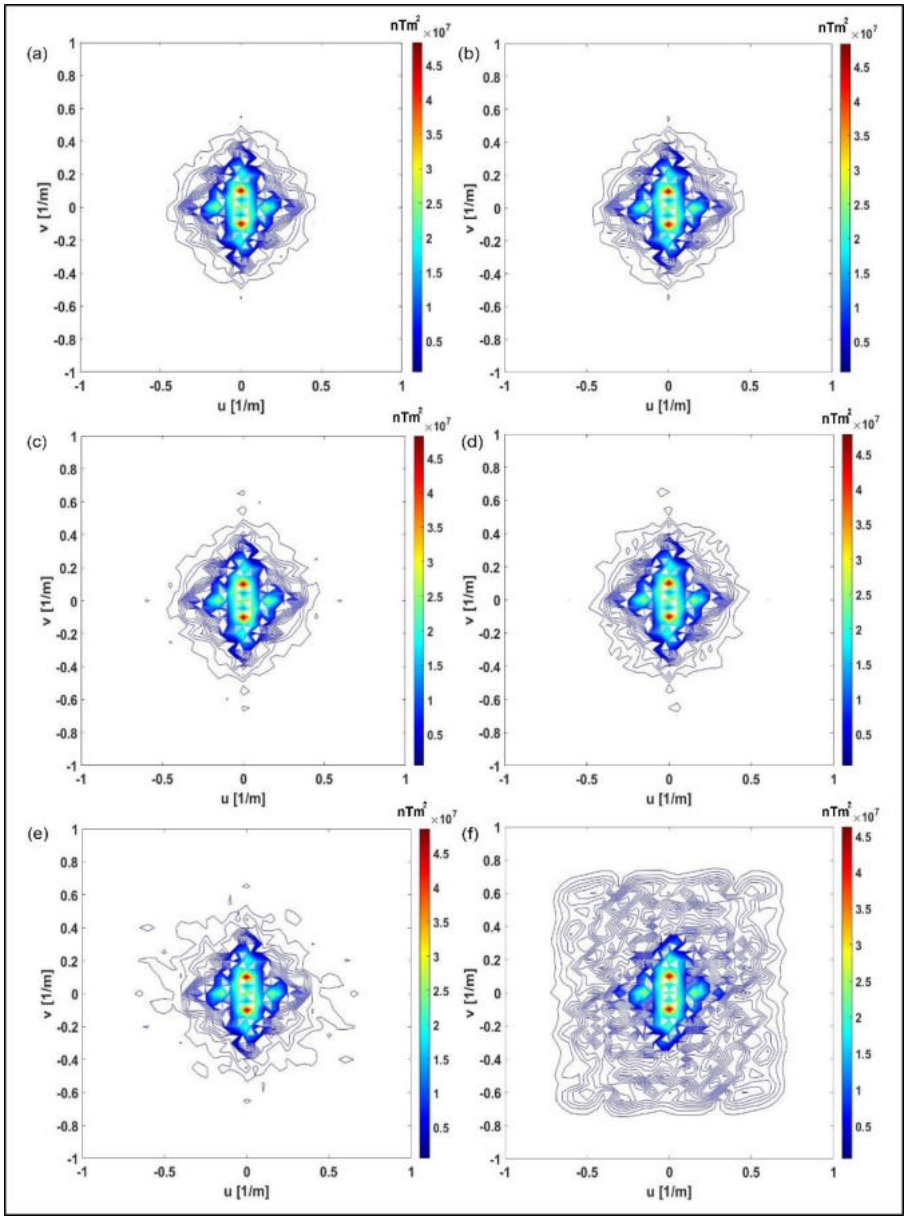

Figure 59. *The 2D IRLS-FT spectra when a) 15% of the regularly sampled noisy magnetic datasets are missing, b) 25% of the regularly sampled noisy magnetic datasets are missing, c) 35% of the regularly sampled noisy magnetic datasets are missing, d) 45% of the regularly sampled noisy magnetic datasets are missing, e) 50% of the regularly sampled noisy magnetic datasets are missing, f) 60% of the regularly sampled noisy magnetic datasets are missing.*

In 15%, 25%, 35%, and 45% missing data cases, the total number of the measurements are 1410, 1281, 1097, and 933 respectively, whereas 821 and 670 for 50% and 60% missing percentages are utilized respectively. The same order of Hermite function (M_x=M_y=25) is selected. The amplitude spectra derived when we randomly cancel 15%, 25%, 35%, and 45% of the observing points are extremely comparable to those obtained in the regularly sampled noise-free situation *(Figures 48a, b, c,* and *d)*, verifying the inversion effectiveness. In 50% of missing data, the spectrum is satisfactory and substantially identical to that calculated in the noise-free datasets *(Figure 50a)*. The noisy datasets illustrate the inversion limitations as displayed in the distorted spectrum due to the marginal over-determination ratio, similar to that produced with the noiseless data while missing 60% of the readings *(Figure 50c)*.

The estimated 2D pole reduction anomaly maps using Eq. 91 are illustrated in *Figures 60a, b, c, d, e,* and *f*. The RTP magnetic anomaly maps nearly indicate the same results as established in the regularly sampled noise-free datasets *(Figures 49a, b, c, d, 50b* and *d* respectively*)*. Numerically, very low data distance values are calculated as 0.0110, 0.0125 0.0238, and 0.0437 for 15%, 25%, 35%, and 45% missing data respectively. In our case, the over-determination is dominant when 50% of the measuring points are cancelled, with its ratio calculated as 821/625=1.3136. The data distance of 0.0716 is roughly seven times higher than that of the 15% missing datasets but it is generally satisfactory. The extremely deformed RTP map while missing 60% of the observation points reveals that the inversion is not effective, as numerically indicated by the very high value of the data distance of 0.4133, similar to the noise-free sampling datasets.

4.5.4.2. Study on Random Walk Sampling

Furthermore, we introduce the inversion findings in the case of random-walk sampling of the 2D noisy magnetic datasets. In that regard, 1435, 1272, 1065, 930, 860, and 635 measuring points are randomly selected to determine the 2D amplitude spectra when missing 15%, 25%, 35%, 45%, 50%, and 60% *(Figures 61a, b, c, d, e,* and *f* respectively*)*. When compared to those detected in regularly sampled noise-free datasets, it is proven that all spectra are substantially identical *(Figures, 48a, b, c, d, 50a,* and *c* respectively*)*. In addition, *Figures 62a, b, c, d, e,* and *f* demonstrate similar results concerning the reduced-to-pole maps of randomly missing noisy magnetic datasets using IRLS-FT. When we compare the RTP maps estimated by the inversion method at 15%, 25%, 35%, and 45% missing data to those determined in the regular noise-free data *(Figure 49a, b, c,* and *d,* respectively*)*, effective results and great similarities in anomaly shapes and amplitudes are attained, proving low data distance values (0.0235, 0.0317, 0.0467,

and 0.0654 respectively). The inversion still gives us acceptable results when half of the measuring points (50%) are executed with a data distance value of 0.0813, while the less over-determination of the inverse problem when missing 60% of the observing points is kept unchanged as numerically evidenced by a higher data distance value of 0.3297.

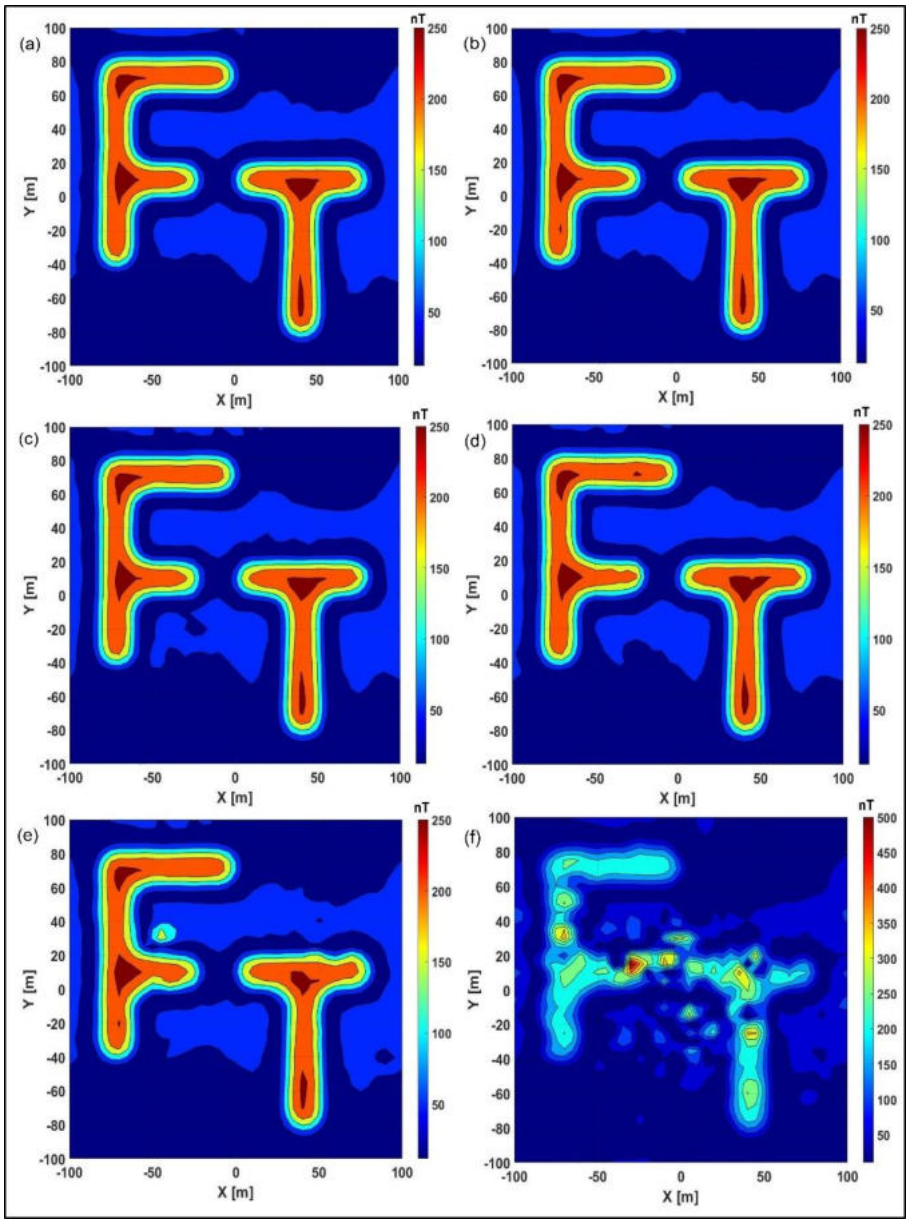

Figure 60. *The reduced-to-pole maps using IRLS-FT when a) 15% of the regularly sampled noisy magnetic datasets are missing, b) 25% of the regularly sampled noisy magnetic datasets are missing, c) 35% of the regularly sampled noisy magnetic datasets are missing, d) 45% of the regularly sampled noisy magnetic datasets are missing, e) 50% of the regularly sampled noisy magnetic datasets are missing, f) 60% of the regularly sampled noisy magnetic datasets are missing.*

According to these results, we can announce that the inversion-based FT is effective, stable, and robust in processing the noise-contaminated incomplete sampling datasets even when 50%

of the measuring points are neglected or lost as a maximum (The over-determination rate is 860/625=1.376). These investigations-based noise contamination show again the new (and economically meaningful) possibility in planning field measurements: the number of measurement points can be reduced before reaching an acceptable inaccuracy in determining the RTP map.

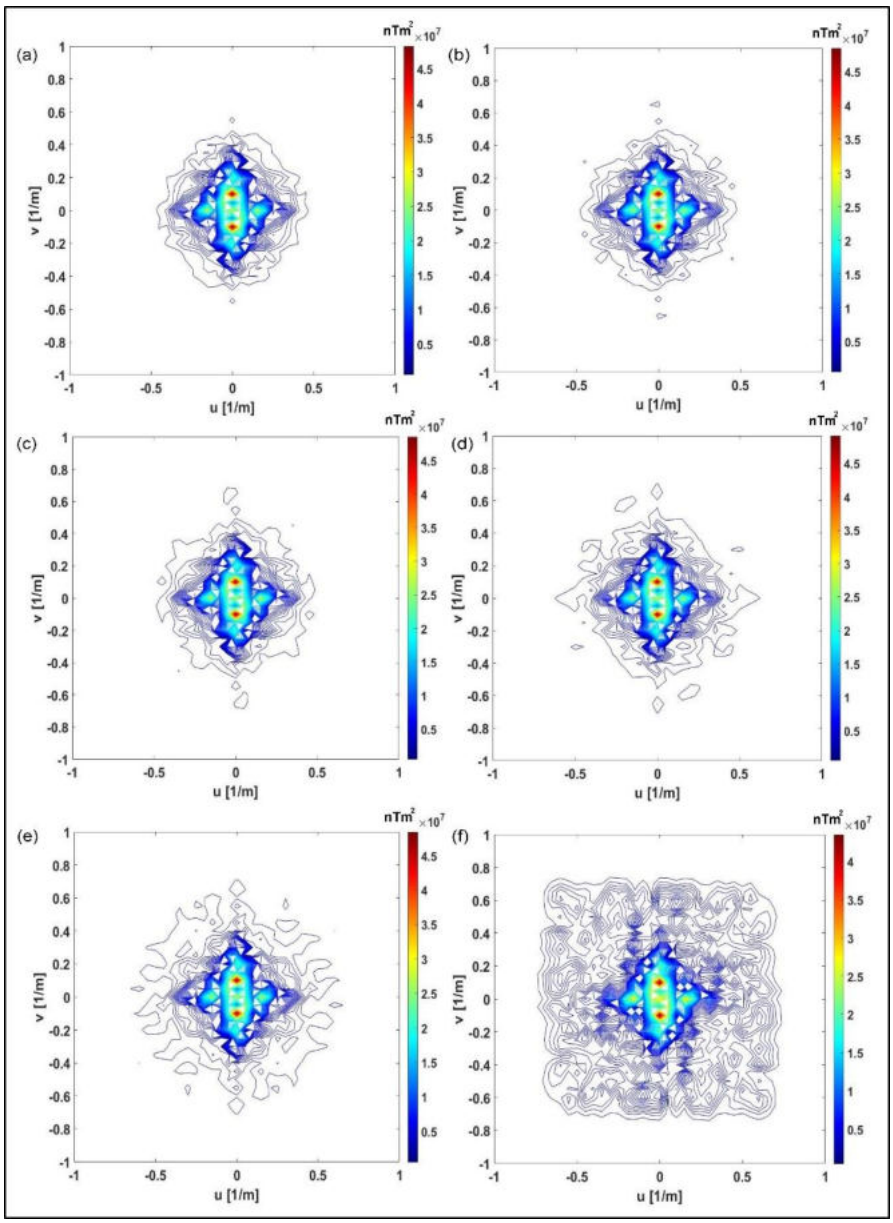

Figure 61. *The 2D IRLS-FT spectra when a) 15% of the randomly sampled noisy magnetic datasets are missing, b) 25% of the randomly sampled noisy magnetic datasets are missing, c) 35% of the randomly sampled noisy magnetic datasets are missing, d) 45% of the randomly sampled noisy magnetic datasets are missing, e) 50% of the randomly sampled noisy magnetic datasets are missing, f) 60% of the randomly sampled noisy magnetic datasets are missing.*

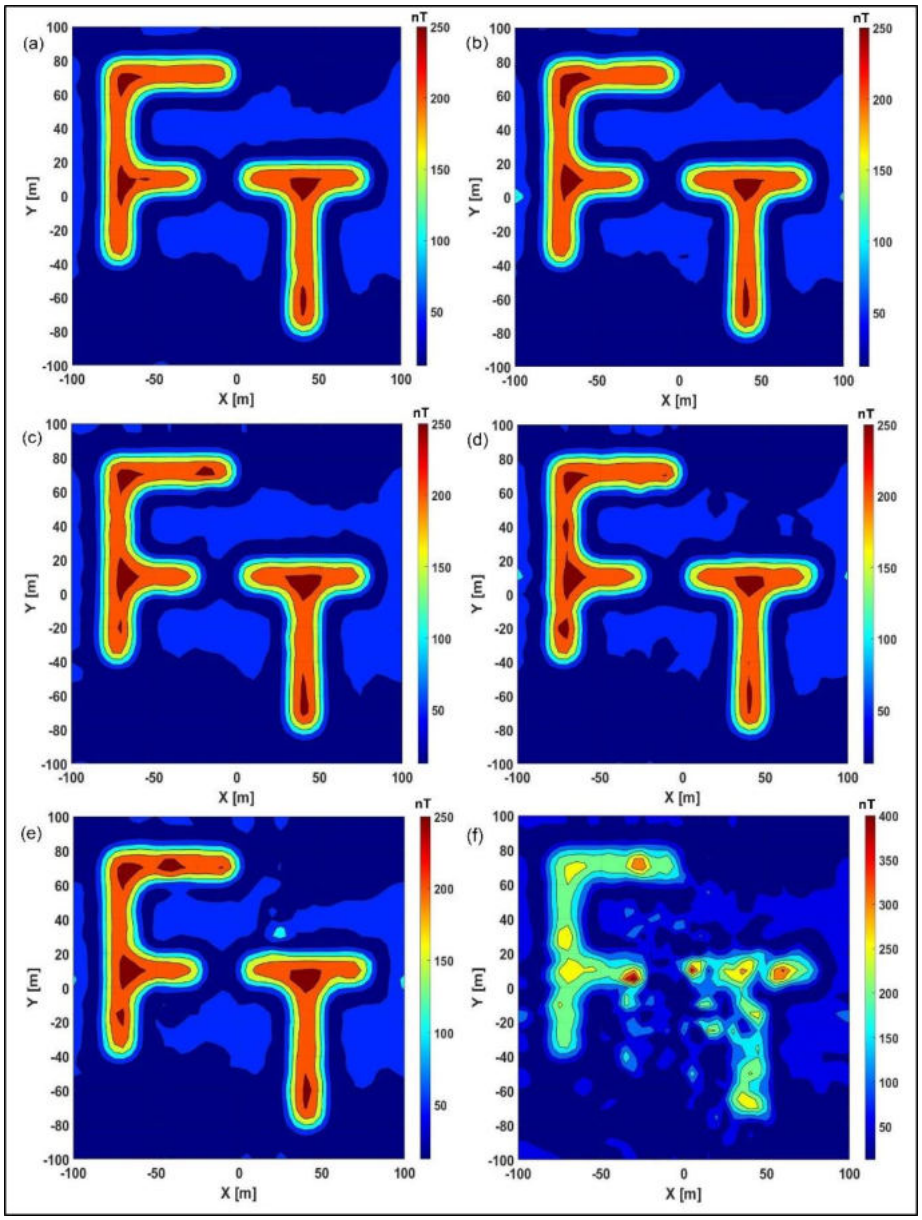

Figure 62. *The reduced-to-pole maps using IRLS-FT when a) 15% of the randomly sampled noisy magnetic datasets are missing, b) 25% of the randomly sampled noisy magnetic datasets are missing, c) 35% of the randomly sampled noisy magnetic datasets are missing, d) 45% of the randomly sampled noisy magnetic datasets are missing, e) 50% of the randomly sampled noisy magnetic datasets are missing, f) 60% of the randomly sampled noisy magnetic datasets are missing.*

According to the results demonstrated above, I declare thesis statement 2 as follows:

Thesis 2

In the field of pole reduction of magnetic data, I gave an extended analysis of the 2D IRLS-FT method with regard to the outlier sensitivity, the non-regular sampling, and the missing data

problems in a numerical example containing 41x41 measurement points and 25x25 unknown expansion coefficients.

Studying the missing data problem on noise-free and Cauchy noise-contaminated datasets (containing outliers), I found that in the case of regularly sampled, non-regularly sampled, and random-walk datasets, increasing the ratio of the missing data up to 50%, the accuracy of the pole reduced magnetic map is acceptable, while at higher ratios (when the over-determination rate is below 1.3) the results are highly distorted.

Chapter 5
GRAVITY DATA PROCESSING USING THE 2D IRLS-FT METHOD

5.1. AN OVERVIEW OF THE GRAVITY SURVEYING METHOD

The gravitational method is based on the measurement of gravity field changes due to horizontal and vertical density variations in the subsurface. The gravity method is preferred by many geophysicists due to several advantages where it is considered one of the cheapest geophysical tools and a way of remote sensing that is non-invasive and non-destructive. Besides, it is a passive geophysical exploration approach, in the sense that the phenomenon whose effect is measured is natural, so there is no need to produce it. Because of the large variation of densities among rock types, assumptions concerning stratum distribution can be made. For more details about the densities of several rocks and ore minerals, some geophysical books are available such as (Dobrin, 1976; Sharma, 1997; Milsom, 2003). Like other geophysical tools, the gravity method has a wide range of applications. For many issues involving surface mapping, it is an important technique and is the principal method for several specific geological study types. It can be used for regional petroleum exploration, mineral prospecting, geotechnical and archaeological investigations, environmental applications, and also to find out how deep the bedrock is, and the top of the source rock bodies (Roy, 1966; Nettleton, 1976; Ateya and Takemoto, 2002; Fedi, 2007; Lafehr and Nabighian, 2012; Hinze et al, 2013; Essa, 2014 and Luan Thanh Pham et al, 2021). In addition, it can be used in groundwater exploration and to detect structural trends controlling the regional geometry of the groundwater aquifers (Murty and Raghavan, 2002).

Furthermore, there are several types of gravity instrumentation (gravimeters) that can be involved to measure the changes in the vertical component of the gravity field. For accurate data collection, the gravimeters should be sensitive enough to be able to determine the earth's gravity field up to one part in a hundred million. Low mass, low volume, low power demand, safe operation, and ruggedness and reliability of the geophysical gravimeters are all requirements for human exploration of the solar system (Meyer et al., 1995; Hoffman, 1997; Budden, 1999). However, before the gravitational data interpretation, all the observing points have to be corrected to overcome the gravitational ambiguity problems associated with data interpretation. This can be done by removing all the unknown effects that are not connected to variations in subsurface density. Data processing includes corrections for drift, tide, free-air, Bouguer, latitude, and terrain. The processing and interpretation procedures of the gravity field

datasets have been found in several works of literature (Reynolds, 1997; Kearey et al, 2002; Milsom, 2003; Hinze et al, 2013). The gravity interpretation data can be then displayed as a map called the Bouguer anomaly contour map or along 2D profiles, illustrating the lateral variation in the underlying rock densities.

5.2. SPECTRAL LOW-PASS FILTERING TECHNIQUE

The gravitational data processing can be achieved by using numerous spectral filtering methods based on Fourier transform. Removing or suppressing the noise through the spectral data analysis is highly recommended compared to the digital data analysis in the space domain due to its flexibility, wider application, effectiveness for larger data sets, and lower edge effects (Hinze et al, 2013). These spectral filtering methods play a vital role in geophysical data processing because they help to isolate the responses of near-surface sources of interest that are embedded in deeper regional ones of widespread extension. Therefore, a simple low-pass filtering technique is used in our study as a geophysical application for gravity data processing. Generally, in the low-pass or high-cut filter, signals with frequencies lower than a selected cutoff frequency are passed (passband), while frequency signals over the cutoff value are suppressed (stopband). In gravity data processing, low-pass filtering is utilized to allow low wavenumber or long-wavelength components produced by deep gravitational sources (regional) to pass, while rejecting higher wavenumber or short-wavelength components associated with local shallow sources (residuals), resulting in a regional gravity anomaly map. This process can be performed depending on the cut-off frequency value accuracy selected from the 2D Fourier amplitude spectrum. Our present investigation analyzes the 2D synthetic noise-free and noisy gravitational datasets taken over equidistant and random sampling grids. In the framework of the low-pass filtering, a new practical geophysical applicability of both 2D DFT and 2D IRLS-FT approaches are implemented on such datasets to evaluate their noise reduction capabilities as well as to test the IRLS-based inversion efficiency in solving the incomplete sampling problems.

5.3. THE GRAVITATIONAL FIELD OF A RECTANGULAR PRISM

5.3.1. MATHEMATICAL SOLUTION OF THE PROBLEM

In this section, the mathematical expressions for estimating the vertical component of a right rectangular prism's gravitational attraction are presented. The right rectangular prisms assume that their edges are parallel to the coordinate system, unlike those of polygonal ones. The derivation of the gravitational field produced by a right rectangular prism has been explained

by many people (e.g. Nagy, 1966; Shuey and Pasquale, 1973; Rasmussen and Pedersen, 1979; Cady, 1980). Sir Isaac Newton (1687) derived two laws that serve as the foundation for the gravity method. The first namely Newton's universal law of gravitation states that the attraction force (F) between the two bodies is directly proportional to the product of their masses, and inversely proportional to the square of the distance between their centers. Thus, if we consider a given mass M (prismatic body) at a point (x, y, z) is acting on a proof mass m located at a measuring point $O = (x_0, y_0, z_0)$ as shown in *Figure 63*, the magnitude of the attraction force in the Cartesian coordinate system is defined by:

$$F = G\,\frac{M\,m}{r^2}, \tag{92}$$

where G is the universal gravitational constant, which has a value of 6.67 x 10^{-11} m^3 Kg^{-1} S^{-2} in the International System of Units (SI) and r is the distance between the two masses as follows:

$$r = [(x_0 - x)^2 + (y_0 - y)^2 + (z_0 - z)^2]^{\frac{1}{2}}. \tag{93}$$

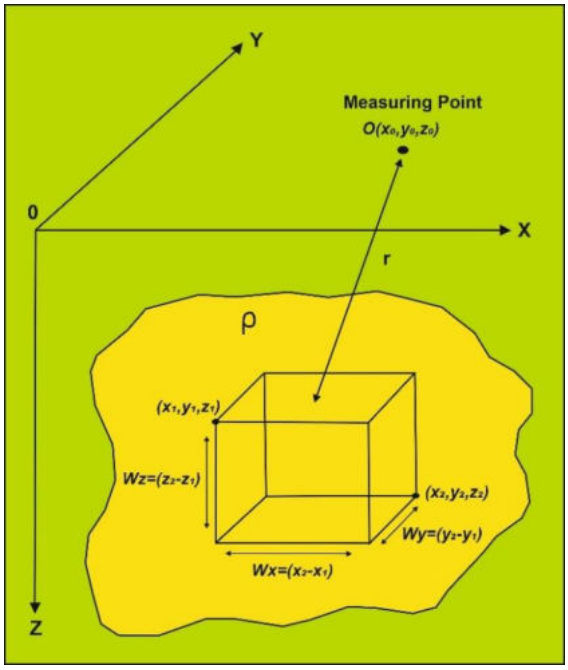

Figure 63. *Model for calculating the vertical component of a right rectangular prism's gravitational acceleration at a measuring point O (modified after Karcol and Pašteka, 2019).*

On the other hand, the Newton's Second Law of Motion states that the force (F) equals the product of acceleration and mass $(F = m.g)$, then the final product of the gravitational acceleration caused by a mass M on a proof mass m at the measuring point O is given by:

$$g(O) = -G\,\frac{M}{r^2}\hat{r}, \tag{94}$$

where $\hat{r}$ is the unit vector at observation point O that is oriented to mass M in the direction opposite to the field, and hence a negative sign is present. Following the Helmholtz theorem (Blakely, 1996), the potential field is irrotational, thus the gravitational acceleration is a conservative field which can be then indicated as the gradient of a scalar potential $g(O) = -\nabla U(O)$. Consequently, the gravitational potential at the point O can be written as follows:

$$U\,(O) = G\,\frac{M}{r}. \tag{95}$$

It is important to note that the gravitational potential follows the principle of superposition. This means that the gravitational potential of a group of masses is equal to the total of each mass's gravitational attraction. As a result, the net force acting on a proof mass is just the vector sum of all the forces produced by all the masses in the vicinity of the proof mass. By this principle, the value of U due to the body composed of N elements of mass ΔM_j can be estimated approximately according to the following formula:

$$U(O) \approx G\,\sum_{j=1}^{N}\frac{\Delta M_j}{\mathrm{r}_j}. \tag{96}$$

Moreover, ΔM can be replaced by dM which is a product of the distribution of the density in the body $\rho\,(x, y, z)$ and the body volume (dv). Using an integral instead of a sum in Eq. 96, the gravitational potential for distribution of constant density can be defined as:

$$U(O) = G\rho\,\int_V \frac{1}{r}\,dv. \tag{97}$$

The gradient $g(O) = -\nabla U(O)$ described above can be moved inside the integral in case of a continuous and finite function of the integrated value and its derivatives within the integration region. For example, the partial vertical derivative of U can be defined as:

$$\frac{\partial U(O)}{\partial z} = -G\rho\,\int_V \frac{\partial}{\partial z}\left(\frac{1}{r}\right)\,dv = G\rho\,\int_V \frac{(z_0 - z)}{r^3}\,dv, \tag{98}$$

with $z_0 = 0$ at the surface. For variables x and y, similar expressions are found. It is critical to note that when dealing with bodies of different geometries such in our case (prismatic body), the coordinate system has to be separated in order to convert the integral of volume to the surface one. This can be achieved by using the Cartesian coordinate system $(dv = dx\,dy\,dz)$. Therefore, the vertical component of the gravitational acceleration can be given as:

$$g_z = G\rho\,\int_{x_1}^{x_2}\int_{y_1}^{y_2}\int_{z_1}^{z_2}\frac{z\,dz\,dy\,dx}{r^3}, \tag{99}$$

where x_i, y_i, and z_i are the prism's corners as shown in *Figure 63*. Plouff (1976) presented a solution to this integral for prisms with polygonal horizontal sections, while in the present case of the rectangular prism or L-sided polygon, the integral's solution was detailed by (Nagy, 1966) and given by (Blakely, 1996) as the following:

$$g_z = G\rho \sum_{i=1}^{2}\sum_{j=1}^{2}\sum_{k=1}^{2}\mu_{ijk}\left[C_k \arctan\left(\frac{A_i B_j}{C_k R_{ijk}}\right) - A_i \ln\left(R_{ijk} + B_j\right) - B_j \ln\left(R_{ijk} + A_i\right)\right], \quad (100)$$

where $A_i = (x_0 - x_i)$, $B_j = (y_0 - y_j)$, and $C_k = (z_0 - z_k)$ is the distance from the measuring point O to each corner of the prismatic body, $R_{ijk} = \sqrt{A_i^2 + B_j^2 + C_k^2}$, and $\mu_{ijk} = (-1)^i (-1)^j (-1)^k$. This formula in Eq. 100 is used for the forward modeling and computing of the vertical component of the gravitational acceleration produced by a prism with arbitrary dimensions and density.

5.3.2. SYNTHETIC 2D GRAVITY DATA GENERATION

In order to test the noise reduction capability of both traditional Fourier transformation and the inversion-based Fourier transformation methods, synthetic datasets are constructed. A forward gravity modeling example produced by a right rectangular prism using Eq. 100 is presented in our investigation. The distances are in meters (m), the density is in g/cm^3 and the data of gravity are in mGal. For calculating an anomaly in a 2D density distribution, the subsurface is divided into rectangular cells with a side of the *x*-axis equal to the distance between two measuring stations. This example shows a prism inserted into the center of an area extending from -20 to +20 meters, which is represented in *x* and *y* directions by 40 by 40 rectangular cells spaced by 1 m as shown in *Figure 64a*. The prism can be specified by several particular parameters including width (*Wx*) in the *x*-direction, length (*Wy*) in the *y*-direction, thickness (*Wz*) in the *z*-direction, depth, centre coordinates, and density as described in *Figure 63*. In the present study, a prism with a centre is assumed to be 5 m deep, 10 m wide and long, and 5 m thick. To solve the edge effects problem, the prism is given a density of 1 while the background density is set to 0. The responses collected in *x* and *y* directions on the surface, spaced by 1 m, centred on the prism's top, produced 1681 points of observation, and then the forward matrix consisted of 41 columns and 41 rows. *Figure 64b* shows the 2D gravity anomaly of the rectangular prism mentioned. The data generated can then be randomly contaminated with noise for real data simulation.

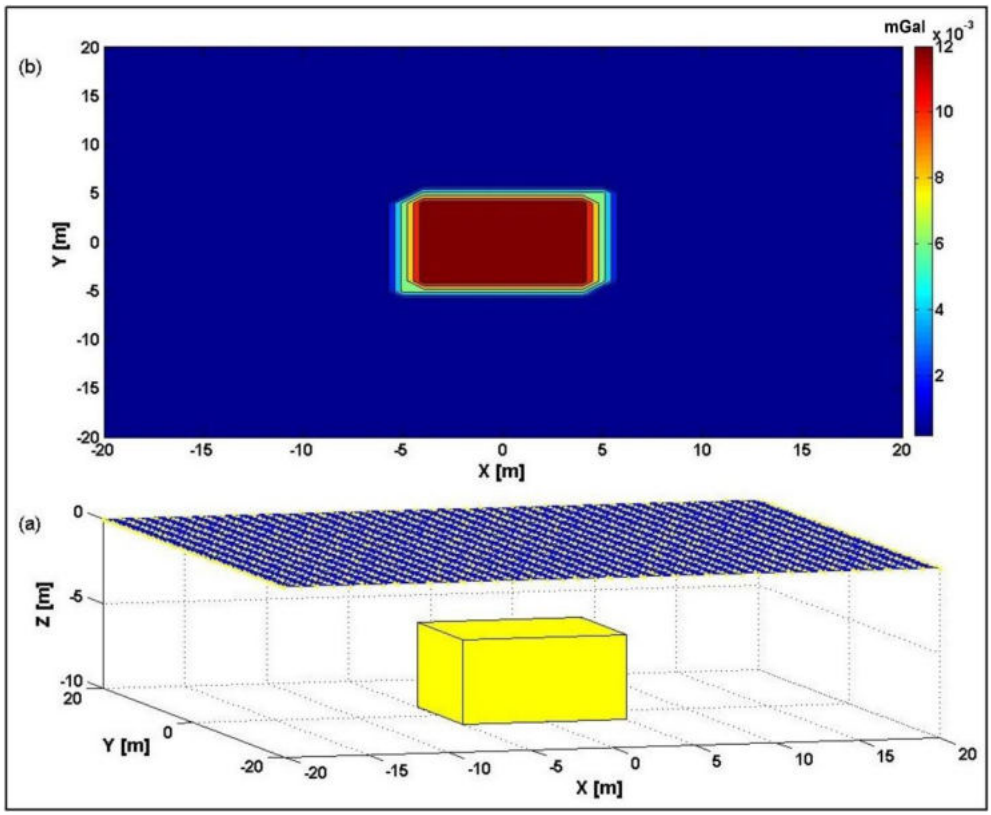

Figure 64. *a) Model showing the location of the rectangular prism in the subsurface and the distribution of the surface measuring stations, b) 2D gravity anomaly produced by the rectangular prism.*

5.4. REDUCING THE GRAVITATIONAL OUTLIER SENSITIVITY IN 2D IRLS-FT

For processing 2D magnetic data in the field of pole reduction, Dobróka et al. (2017) used both conventional Fourier transform (DFT) and inversion-based 2D Fourier transform (2D IRLD-FT) methods, as discussed in Chapter 4. Following up on the previous work, we deal with a new geophysical tool of gravity method to introduce the inversion efficiency and stability. Therefore, the new applicability of two approaches concerning gravitational datasets is presented. The outlier sensitivity is evaluated on the 2D synthetic gravity datasets mentioned above. The spectral low-pass filtering technique is used as a new practical branch in both cases of the equidistantly and non-equidistantly sampled gravity datasets.

5.4.1. APPLICABILITY TO THE EQUIDISTANTLY SAMPLED DATASETS

5.4.1.1. 2D Numerical Investigation

Noise-free and noisy equidistantly sampled gravity datasets are constructed to assess the efficiency and noise reduction capability of the two approaches. *Figure 65* depicts the surface observation locations (N=41x41) spread consistently over the investigated area with intervals of 1 m in the *x* and *y* directions.

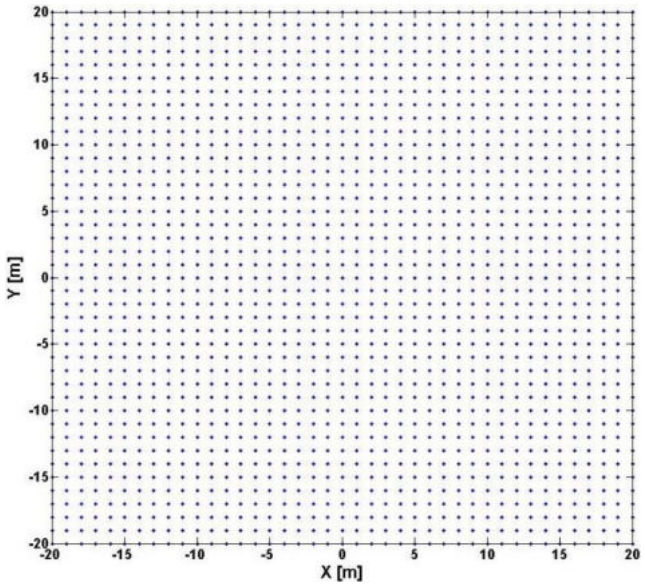

Figure 65. *The regular grid of the measuring gravity stations.*

The rectangular prism's 2D Bouguer gravity anomaly map is built by contouring all of the data gathered at these measuring stations *(Figure 64b)*, exhibiting comparatively high and low gravity anomalies. In *Figure 66*, we can see a 3D visualization of the measured two-dimensional noise-free gravity datasets that vary in magnitude from 0 to $12x10^{-3}$ mGal.

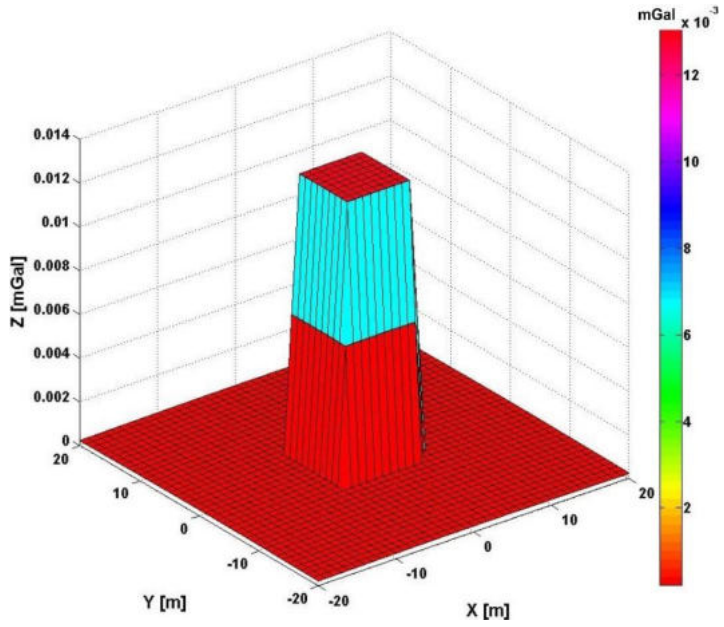

Figure 66. *3D view of the noise-free Bouguer gravity data for the regularly sampled grid.*

The noise-free gravity datasets are then converted from the space domain to the frequency domain using the Fourier transform algorithms. *Figure 67* demonstrates the 2D amplitude spectrum of the conventional DFT method, which is inserted in an area ranging from -0.5 to 0.5 in the *x* and *y* directions to obtain an appropriate scale in the wavenumber domain.

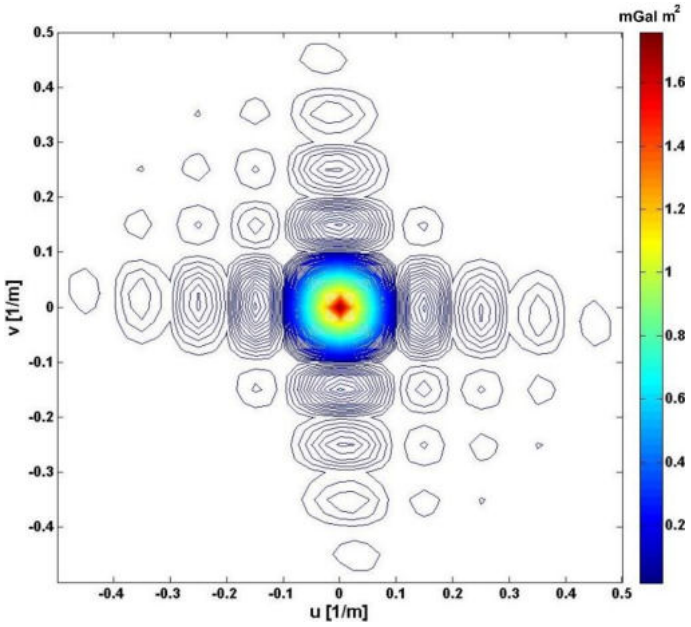

Figure 67. *The 2D amplitude spectrum of the noise-free gravity data using the DFT method.*

For our purposes of testing the noise suppression capacity of the two approaches, the above-mentioned 2D noise-free synthetic gravity data generated with an equidistantly sampled grid *(Figure 66)* is contaminated with random noise of the Cauchy distribution using a scale parameter of 0.02 as shown in *Figure 68* by a 3D view of the noisy Bouguer gravity field.

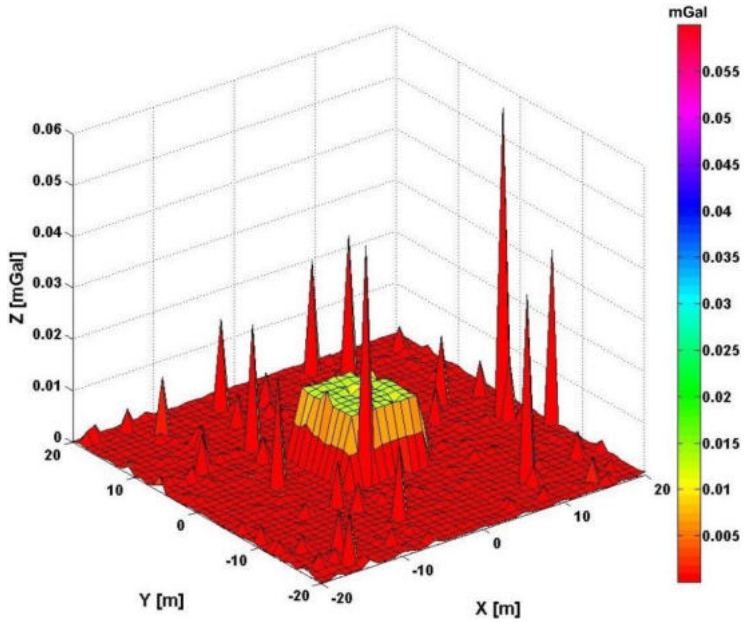

Figure 68. *3D view of the noisy Bouguer gravity data for the regularly sampled grid.*

It is seen that the noisy 2D gravity data is deformed and distorted in both anomaly shapes and amplitudes when compared to noise-free datasets *(Figure 66)*. The result is that gravity values appear to be quite random. Similarly, the noisy gravity datasets are treated with the traditional DFT technique, which produces a two-dimensional Fourier amplitude spectrum, as

illustrated in *Figure 69*. As can be seen, the spectrum of the noisy data produced by the traditional method is completely deformed when compared to that of the noise-free datasets *(Figure 67)*, illustrating that the DFT method is extremely sensitive to the added noise.

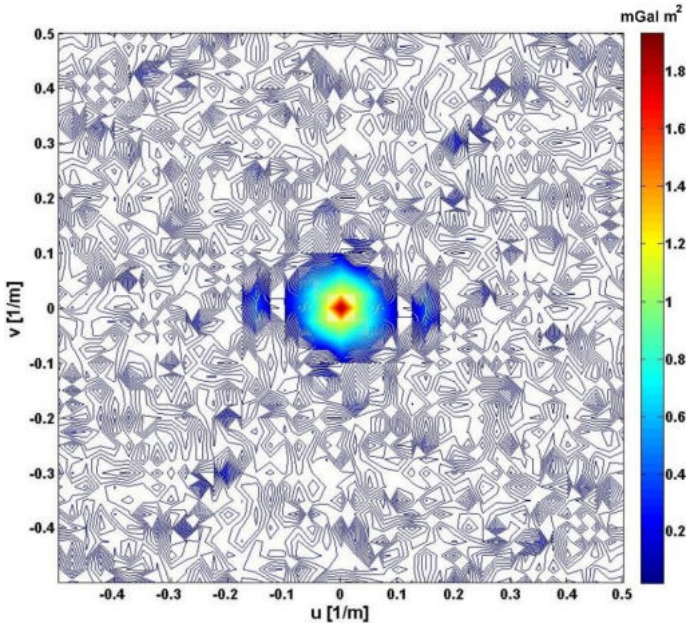

Figure 69. *The 2D amplitude spectrum of the noisy gravity data using the DFT method.*

Due to the limited noise reduction capability of the 2D DFT method, as demonstrated above by its Fourier spectrum, especially in Cauchy noise situations, the inversion-based Fourier transformation (2D IRLS-FT) method is used to compare and test how successful and effective the inversion procedure is on the same noisy datasets. The estimated two-dimensional amplitude Fourier spectrum utilizing the inversion-based method is introduced in *Figure 70*. The results show sufficient improvements, as evidenced by the spectrum of the inversion-based FT approach *(Figure 70)* as well as eliminating the problem of outliers effects associated with the DFT spectrum *(Figure 69)* to be nearly identical to that achieved for the noise-free gravity datasets *(Figure 67)*. Hence, Hermite functions of order M_x=M_y=25 were chosen as an inversion parameter utilizing series expansion to estimate the 2D IRLS-FT spectrum for a compromise between accuracy and stability of the inversion procedure.

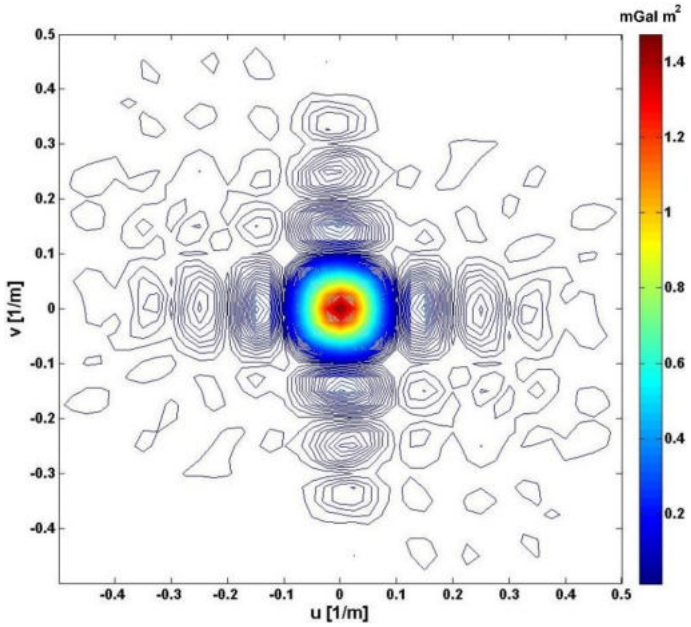

Figure 70. *The 2D amplitude spectrum of the noisy gravity data using the IRLS-FT method.*

To quantify the misfit between noisy and noise-free datasets, RMS computations are assessed. The following equation is used to calculate the distance between noisy and noiseless data in the space domain:

$$d = \sqrt{\frac{1}{N}\sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\left[u^{noiseless}(x_i, y_j) - u^{noisy}(x_i, y_j)\right]^2}, \tag{101}$$

where $N = N_x N_y$, N_x and N_y are the associated numbers of the 2D examination area's space points. As an alternative, the frequency domain model distance between spectra can be estimated using the following equation:

$$D = \left[\frac{1}{M}\sum_{i=1}^{M_x}\sum_{j=1}^{M_y}\left(Re\left[U^{noisy}(\omega_{xi}, \omega_{yj})\right] - Re\left[U^{noiseless}(\omega_{xi}, \omega_{yj})\right]\right)^2 + \right.$$

$$\left.\frac{1}{M}\sum_{i=1}^{M_x}\sum_{j=1}^{M_y}\left(Im\left[U^{noisy}(\omega_{xi}, \omega_{yj})\right] - Im\left[U^{noiseless}(\omega_{xi}, \omega_{yj})\right]\right)^2\right]^{\frac{1}{2}}, \tag{102}$$

where $M = M_x M_y$, M_x and M_y are the 2D test area's frequency points. In that regard, the calculated distance between Cauchy noise-contaminated data and noise-free data is d=0.0027. Moreover, the noiseless spectrum *(Figure 67)* and the noisy spectrum obtained by the DFT method *(Figure 69)* are separated by a model distance of D= 0.0995, whereas the estimated value for the noisy spectrum produced by the inversion (IRLS-FT) method *(Figure 70)* has a model distance of D= 0.0133. These numerical distances demonstrate the applicability and the improved noise reduction capability of the inversion-based Fourier transformation algorithm.

The effectiveness and efficiency of the inversion method are further demonstrated by the results of the 2D inverse Fourier transformation. This phase converts the data from the frequency domain to the space domain to estimate the calculated data (inverted model). *Figures 71* and *72* show the 2D inverse Fourier transformation of the noisy amplitude spectra produced by DFT and 2D IRLS-FT approaches, respectively. A thorough analysis of these maps reveals that the inverted model of the equidistantly sampling noisy gravity data obtained by the DFT approach *(Figure 71)* is the same as the noisy measured data *(Figure 68)* in terms of anomaly shapes and amplitudes, as it is expected. The added Cauchy noise that simulates outliers remains unchanged without any advancements. This is also shown by the fact that the data distance used to calculate the misfit between the noise-free measured data *(Figure 66)* and the result of the inverse Fourier transform (IDFT) method *(Figure 71)* is d=0.0027. As a result of the lower resolution as well as the increment in data distance value, the DFT method is not optimal and has a limited noise reduction capability whenever noisy data that contains outliers are available. In contrast, the quality of the inversion-based Fourier transform of the noisy 2D IRLS-FT spectrum *(Figure 72)* has been considerably enhanced with satisfactory improvements and is close to the noise-free Bouguer gravity datasets *(Figure 66)*. This can also be proved quantitatively by its distance from the noise-free data (d= 0.00032) which is much less than the distance calculated previously between the noiseless and Cauchy noise-contaminated data.

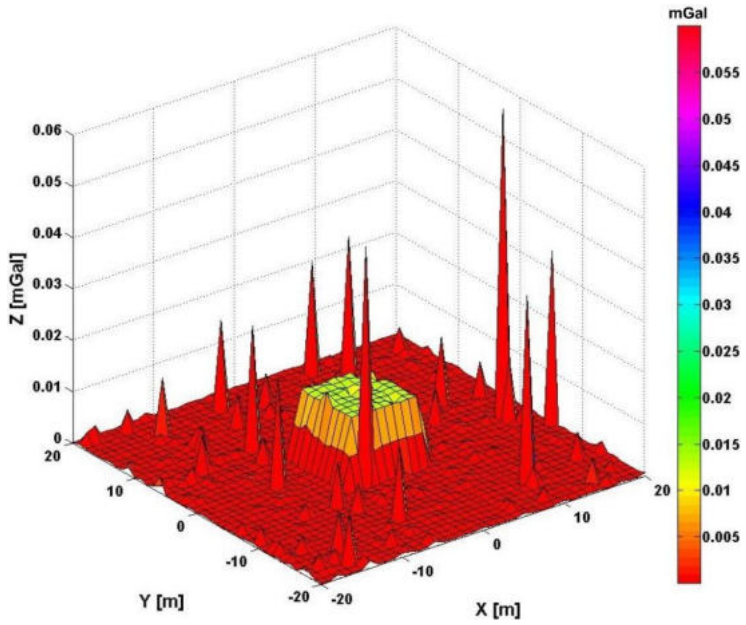

Figure 71. *3D view of the noisy calculated gravity data using the inverse DFT method.*

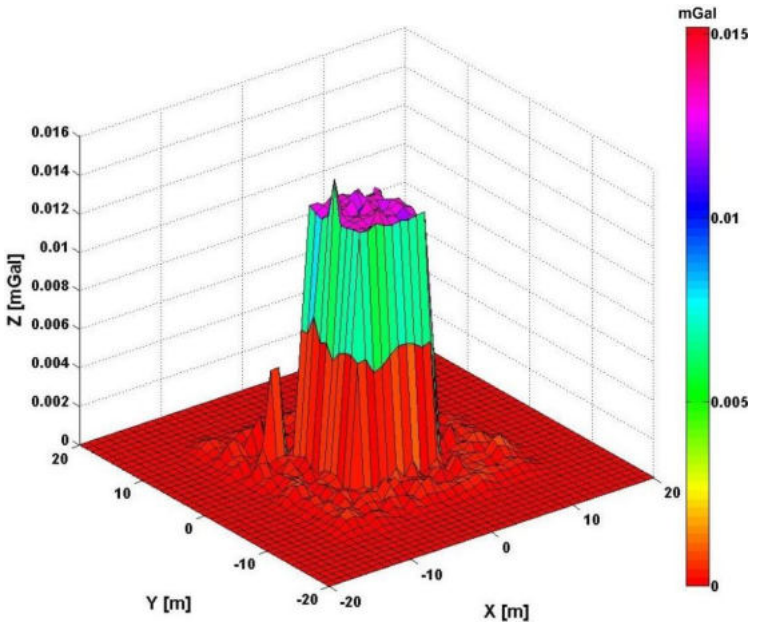

Figure 72. *3D view of the noisy calculated gravity data using the inverse IRLS-FT method.*

Based on these findings, it can be concluded that the inversion-based Fourier transform method applied in our study is highly effective, robust and provides roughly an order of magnitude better noise suppression capabilities in both terms of space and frequency domains than the DFT method, even when data is contaminated with complex noise such as outliers.

5.4.1.2. Low-pass Filtering Applicability

In this study, the low-pass filter is performed mathematically by multiplying the two-dimensional input data in the frequency domain by an appropriate filter function as introduced in the equation below:

$$F(u,v) = D(u,v)\,R(u,v), \tag{103}$$

where $F(u,v)$ is the 2D filtered data in the frequency domain, $D(u,v)$ is the 2D Fourier transformed input data set, and $R(u,v)$ is the filter function. In our present investigation, we use a Butterworth filter (Ellis, 2012; Thompson, 2013; Al Hinai, 2020)

$$R(u,v) = \left[(1+(u/u_c)^{2N})(1+(v/v_c)^{2N}) \right]^{-1/2}, \tag{104}$$

with the cut-off frequencies of $u_c = v_c = 0.125$ [1/m] and the filter order of N=10. The filtered data is subsequently returned to the space domain through the inverse Fourier transformation. To test and compare the noise reduction capabilities of both 2D DFT and 2D IRLS-FT approaches, the same noise-free and noisy data sets presented in *Figures 66 and 68* respectively

are subjected to low-pass filtering. *Figure 73* displays the low-pass filtered 3D noise-free regional gravity anomaly map by the conventional Fourier transformation (DFT) method.

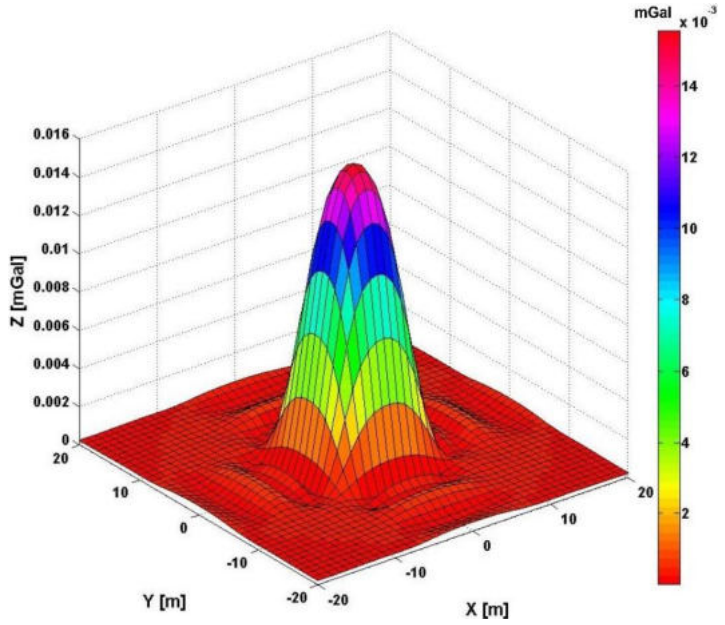

Figure 73. *3D view of the noise-free regional gravity anomaly data (low-pass filtered data) using the DFT method.*

The regional map with high amplitudes and low frequencies can be seen to represent a strong, broad, and great extension of gravity anomalies over the tested area. The low-pass filtering algorithm using the DFT method is then applied to the Cauchy noise-contaminated data *(Figure 68)* resulting in a 3D noisy regional gravity anomaly map as shown in *Figure 74*.

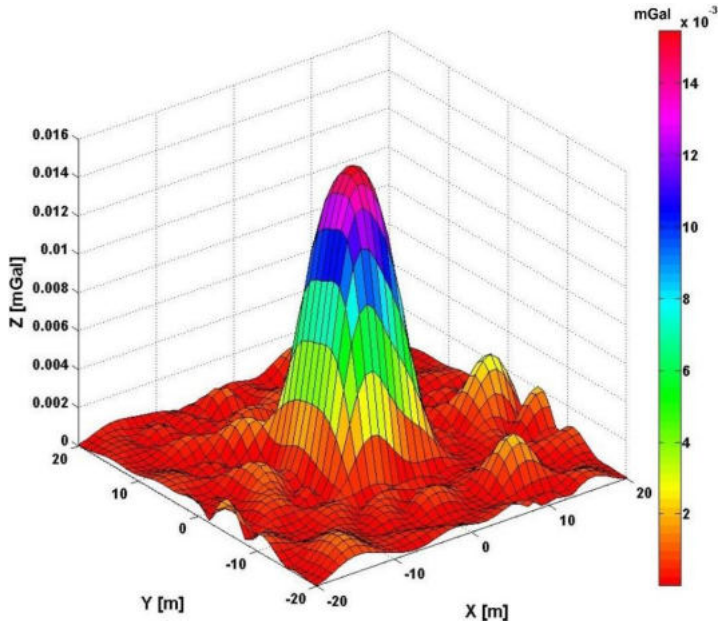

Figure 74. *3D view of the noisy regional gravity anomaly data (low-pass filtered data) using the DFT method.*

It is noticed that the noisy map for regional gravity anomalies is distorted in both anomaly shapes and amplitudes. The inability of handling the problem of outlier effects reflects the limited noise reduction capacity of the 2D DFT method, which is also proven by a higher value of the mean data distance of 0.3752 compared to that of the 3D noise-free regional anomaly map estimated by the DFT method *(Figure 73)*. On the other hand, the 3D noisy regional gravity anomaly map produced by the low-pass filtering based on the 2D IRLS-FT method shows sufficient improvements as seen in *Figure 75*. The result is close to that given by DFT in the case of filtering the noise-free data set showing a low data distance value of 0.0234.

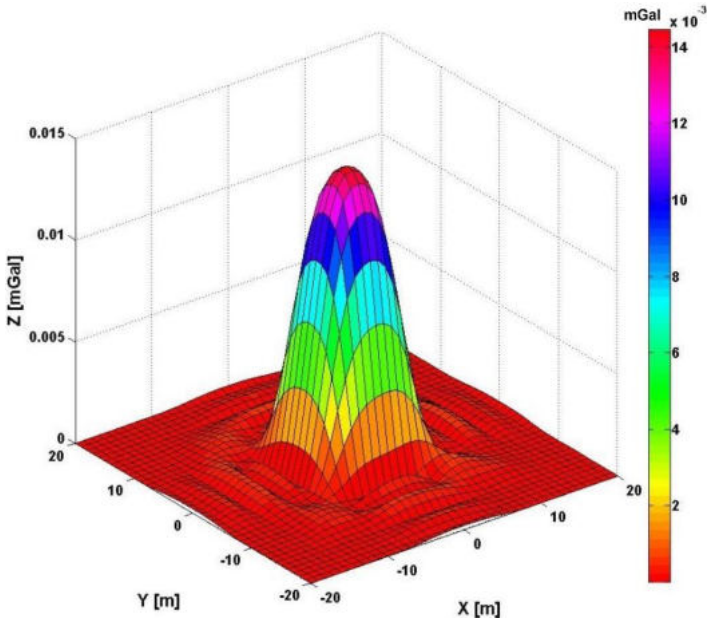

Figure 75. *3D view of the noisy regional gravity anomaly data (low-pass filtered data) using the IRLS-FT method.*

As proved above, when compared to the noisy regional gravity anomaly map produced by the traditional DFT method in *Figure 74*, the results obtained using the inversion-based Fourier transformation method are extremely satisfactory and demonstrate sufficient enhancements in both anomaly shapes and amplitudes. In addition, the inversion-based Fourier transformation method used in this study has solved the outlier effect problem to be somewhat similar to that of the noise-free regional gravity anomaly map *(Figure 73)*. These results are sufficient to demonstrate the inversion-based Fourier transformation 2D IRLS-FT method's applicability, efficiency, and noise reduction capability.

5.4.2. APPLICABILITY TO THE NON-EQUIDISTANTLY SAMPLED DATASETS

Since the 2D DFT and FFT methods are traditionally used for processing the geophysical data sampled equidistantly as well as their inaccuracy or non-applicability with the non-

equidistant grid intervals, we only employ the inversion-based filtering approach; the iteratively reweighted least-squares Fourier transformation (2D IRLS-FT), to random walking gravity measurements. The concept of the random walk measuring was detailed in chapter 4 for 2D magnetic data processing. To assess the noise rejection capacity, the non-equidistant sampling gravity measurements are then contaminated with Cauchy distributions' random noise simulating real data. The algorithm of the above-mentioned inversion-based Fourier transformation is applied to both non-equidistantly sampled noisy and noise-free 2D synthetic gravity datasets, providing practical effectiveness in the framework of the low-pass filtering.

5.4.2.1. 2D Numerical Investigation

To meet our main objective and to verify the performance of the previously indicated inversion method, the same model of the right rectangular prism indicated above in *Figure 64* is used to generate irregular or random gravity datasets. The model is initially applied along with a regular grid before being randomized to produce non-equidistantly sampled intervals. The observation points are positioned 1 meter apart in the *x* and *y* directions (1681 readings) to emulate the subsurface which is divided into 40 by 40 rectangular cells. *Figure 76a* shows the gravity anomaly in a 2D contour map made from a regular grid which is then used to obtain a non-equidistantly 2D gravity anomaly contour map, as seen in *Figure 76b*. In comparison to that obtained with equidistantly sampled data, the contour map of the 2D gravity anomaly in the case of non-equidistant intervals is considerably deformed in shape, revealing an irregular surface and does not cover the whole tested area.

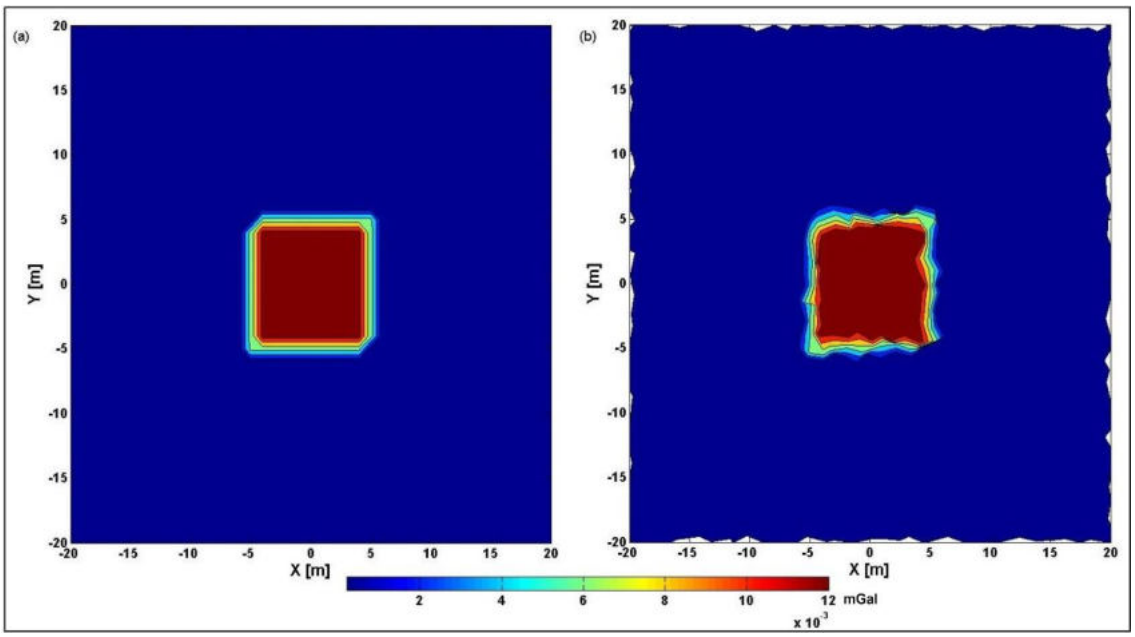

Figure 76. *a) 2D gravity anomaly produced by equidistantly sampled measurements, b) 2D gravity anomaly produced by non-equidistantly sampled measurements.*

In our case, the random walk measuring design can be constructed by arranging all *y* coordinates to all *x* locations. *Figure 77* depicts a non-equidistantly sampled grid of 41x41 surface measuring points generated by randomly chosen *x* and *y* positions.

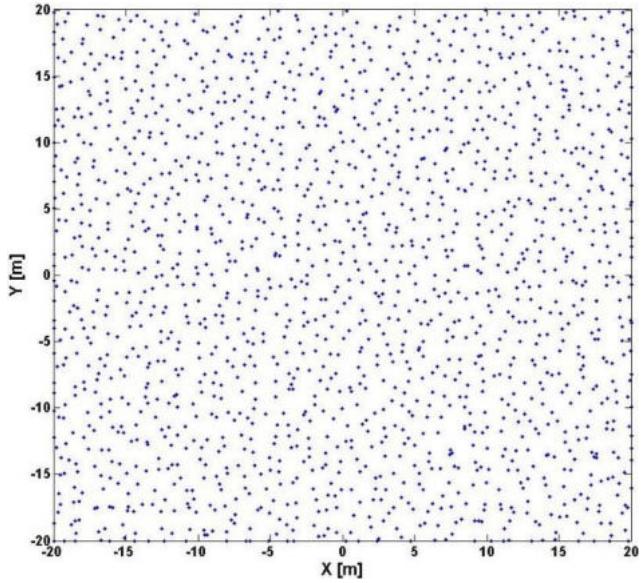

Figure 77. *The non-equidistant grid of the measuring gravity stations.*

First, the inversion-based Fourier transformation method is applied to the 2D noise-free synthetic gravitational datasets illustrated above in *Figure 76b*. To make interpretation processes easier, such datasets are displayed in a three-dimensional (3D) manner, as shown in *Figure 78*. It is observed that the base of the rectangular prism's 2D Bouguer gravity anomaly map is characterized by rectangular blocks of random distributions in *x* and *y* directions reflecting the non-equidistantly sampling.

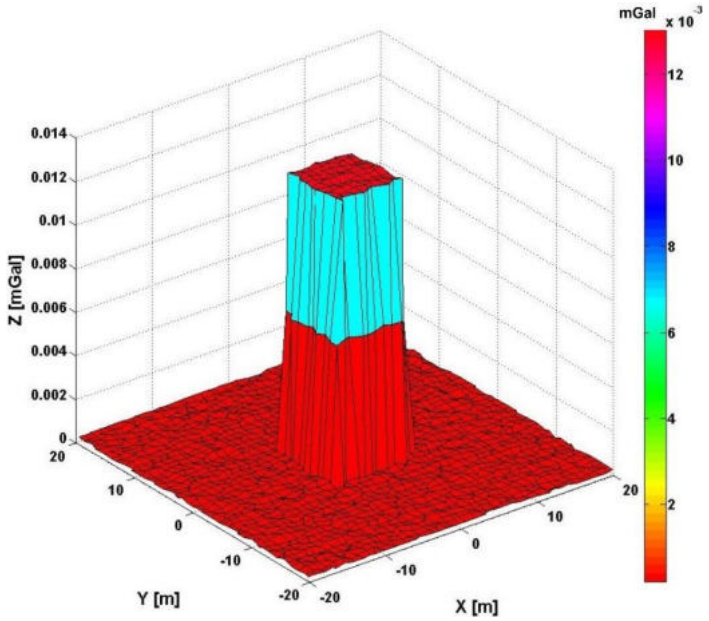

Figure 78. *3D view of the noise-free Bouguer gravity data sampled non-equidistantly.*

After converting the space domain noise-free Bouguer gravity data *(Figure 78)* to the space-frequency domain using our inversion-based Fourier transformation algorithm, we get a 2D amplitude Fourier spectrum of the non-equidistant noise-free datasets, as shown in *Figure 79*.

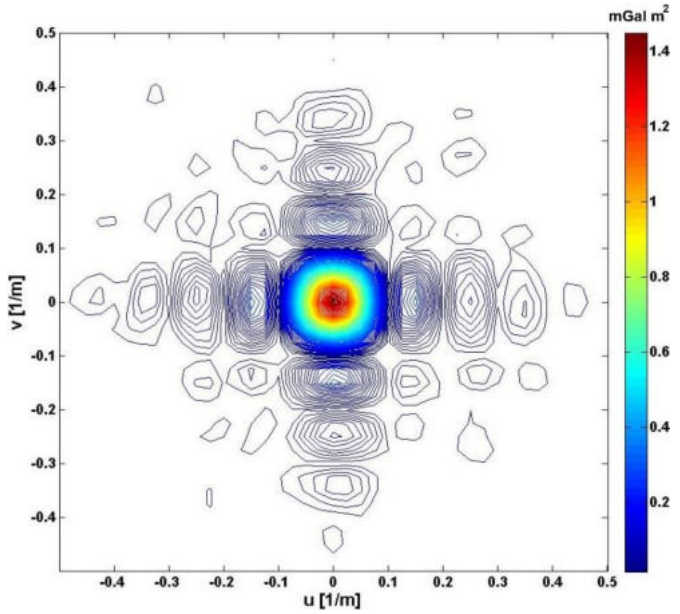

Figure 79. *2D IRLS-FT spectrum of the noise-free gravity data sampled non-equidistantly.*

Here, the 2D IRLS-FT spectrum is estimated using series expansion with Hermite functions of order M=25 in both *x* and *y*-direction. Two-dimensional noisy gravity datasets are then constructed to access the applicability of the proposed inversion-based Fourier transformation algorithm. To generate 2D Bouguer gravity data-based Cauchy noise, we utilize a scale parameter of 0.02 as presented in *Figure 80*.

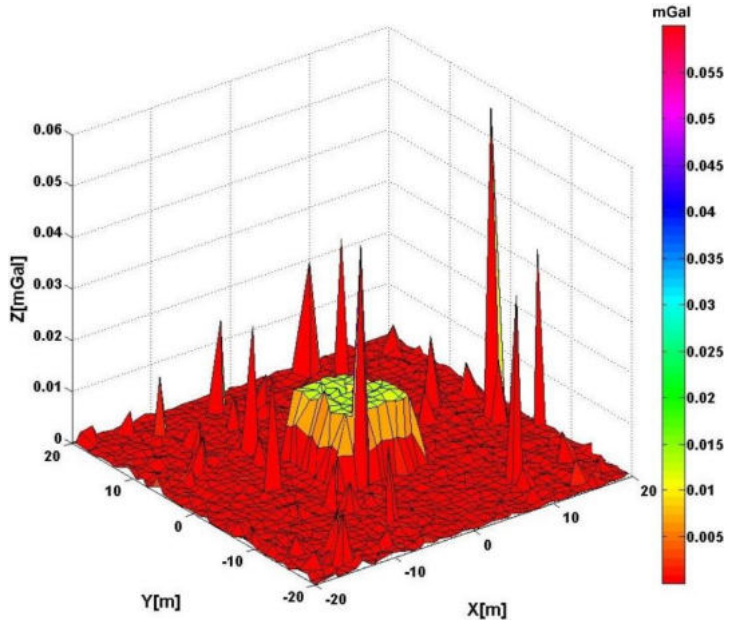

Figure 80. *3D view of the noisy Bouguer gravity data sampled non-equidistantly.*

It can be demonstrated that the outlier effects are readily visible in the noisy gravity data sets which is a crucial step in simulating real data in the field. For quantitative interpretation, the generated models can be evaluated mathematically by calculating the misfit between the noisy and noise-free datasets using Eq.101 indicated above. The outlier sensitivity is shown by the estimated value of the data distance between the noiseless *(Figure 78)* and the Cauchy noise-contaminated data *(Figure 80)* being d=0.0027. To compare the inversion method's applicability to that of noise-free gravity datasets, the noisy 2D gravity data *(Figure 80)* are subjected to the inversion algorithm, which produces a 2D amplitude Fourier spectrum of the non-equidistant noisy datasets as shown in *Figure 81*.

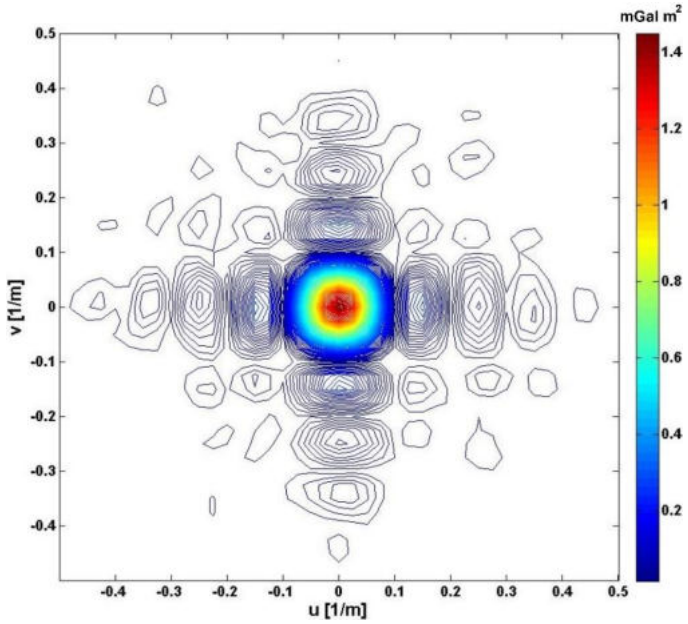

Figure 81. *2D IRLS-FT spectrum of the noisy gravity data sampled non-equidistantly.*

It is seen that the 2D amplitude Fourier spectrum produced by the inversion approach in the presence of Cauchy noise is nearly identical to that obtained in the noise-free situation. This is evident in both anomaly shapes and amplitudes exhibiting the same values to vary from 0.2 to 1.4. The same order of the Hermite functions M_x=M_y=25 was selected as an appropriate parameter in the inversion. In addition to the visual findings, the inversion method's efficiency may be numerically verified in the frequency domain by computing the model distance between spectra as explained earlier in Eq. 102. In this regard, D=0.0129 was calculated as the model distance between the 2D IRLS-FT spectra of noise-free *(Figure 79)* and noisy *(Figure 81)* gravity data. This low model distance value indicates the inversion-based Fourier transformation algorithm's efficiency as well as its superiority in reducing gravitational data noise. In contrast, the inverse Fourier transform is applied to convert the data from the space-frequency domain to the space domain. *Figure 82* shows the inverted space domain model

(calculated data) produced by the inverse Fourier transformation of the non-equidistantly sampling noisy spectrum presented above in *Figure 81*.

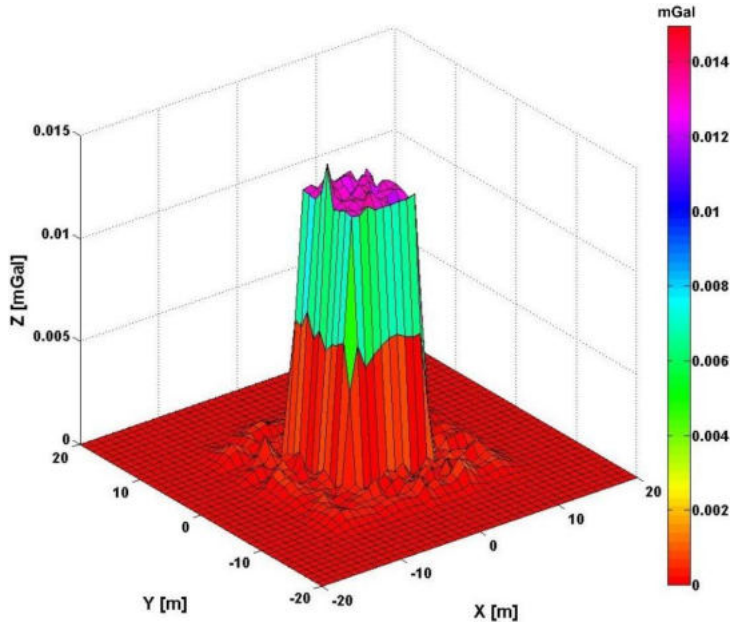

Figure 82. *3D view of the inverse IRLS-FT noisy calculated Bouguer gravity data.*

It is observed that the inverted model-based Cauchy noise *(Figure 82)* by the 2D IRLS-FT method is found to be virtually similar to the gravitational forward model of noise-free datasets *(Figure 78)*. One of the most noteworthy features of our inversion-based Fourier transformation approach, when applied to non-equidistantly sampled datasets, is that the inverted model returned as we would measure along with a regular sampling grid while maintaining the anomaly shapes and amplitudes unchanged. This is evident in the regular spacing of all rectangular blocks at the base of the inverted model in *x* and *y* directions to be nearly similar to that of the regularly sampled noise-free datasets *(Figure 66)*. In addition, the problem arising from outlier effects *(Figure 80)* was resolved, and all undesired noise signals were eliminated. Furthermore, the misfit between the 2D IRLS-FT noisy calculated Bouguer gravity data *(Figure 82)* and noise-free datasets *(Figure 78)* is d=0.00030 which is markedly less than the estimated value obtained for the noiseless versus Cauchy noise-contaminated data. This implies that the inversion-based Fourier transformation used in this study is highly effective in both the space and frequency domains, as well as in terms of application and robustness when dealing with noisy data like outliers collected on non-equidistant or irregular grids.

5.4.2.2. Low-pass Filtering Applicability

Similarly to that applied to the equidistantly sampled gravity datasets, the low-pass filtering technique is implemented on the non-equidistant measuring data points to demonstrate how

successful the proposed inversion is. Employing Eq. 103 and Eq. 104 described above, the regional gravity anomaly map is the cumulative outcome of this filtering, and it must be built using a proper cutoff frequency value selected from the 2D Fourier amplitude spectrum. Hence, we evaluated the noise reduction capacity of the inversion-based Fourier transformation approach by applying cutoff frequencies $u_c = v_c = 0.125$ [1/m] and a filter order of N=10 to both non-equidistantly noiseless and noisy gravitational datasets. *Figure 83* shows the regional gravity anomaly map in 3D after applying the low-pass filtering-based inversion to the noise-free Bouguer gravity data illustrated above in *Figure 78*. The obtained 3D regional gravity anomaly map is characterized by a low data distance value of 0.0268 when compared to that calculated by the DFT approach in the noise-free equidistantly sampled datasets *(Figure 73)*. Intensive, and widespread extension of gravity anomalies over the tested area may be noticed on the regional map with high amplitudes and low frequencies.

The Cauchy noise-contaminated data distributed over a non-equidistant grid *(Figure 80)* are then treated with low-pass filtering based on the 2D IRLS-FT method. The three-dimensional (3D) noisy regional gravity anomaly map is presented in *Figure 84*. It is seen that in the case of Cauchy noisy data, the filtered data of the regional gravity anomalies produced by the inversion method *(Figure 84)* is the same as the non-equidistantly noiseless regional gravity anomaly map *(Figure 83)* in anomaly shapes and amplitudes, overcoming the outlier effect problem with a low data distance value of 0.0287.

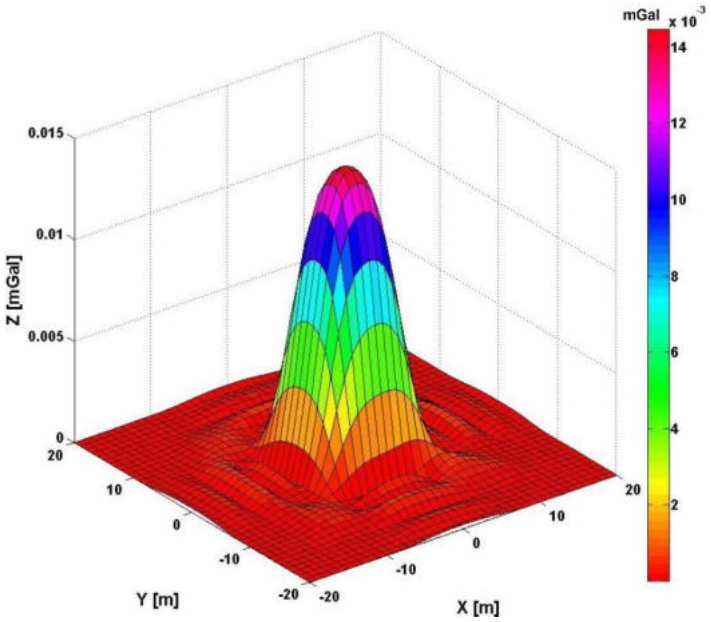

Figure 83. *3D view of the IRLS-FT noise-free regional (low-pass filtered) gravity anomaly data sampled non-equidistantly.*

In addition, it is identical to those produced by both the DFT and IRLS of the regularly sampled noise-free and noisy data *(Figure 73* and *Figure 75* respectively*)*. These findings indicate that the inversion-based FT method employed in our investigation is highly applicable with non-equidistantly sampling procedures, and thus we can confidently recommend it for dealing with real field data measurements sampled either over equidistant or non-equidistant grids.

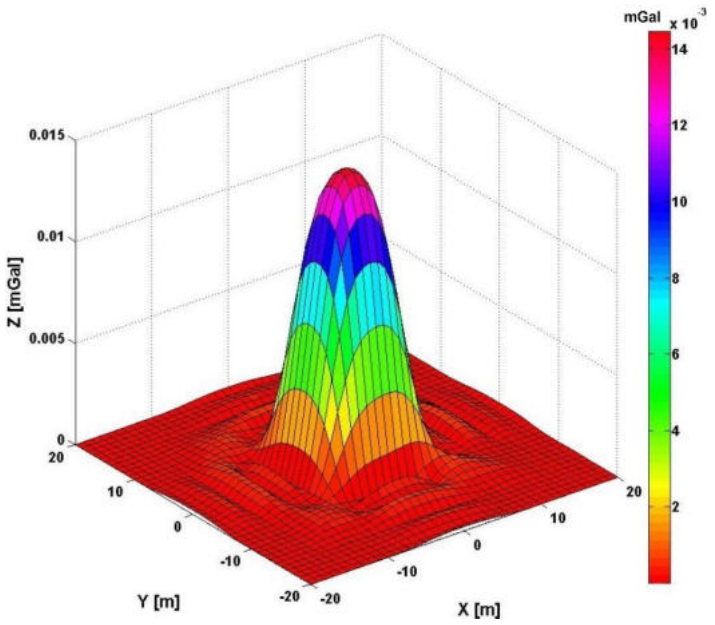

Figure 84. *3D view of the IRLS-FT noisy regional (low-pass filtered) gravity anomaly data sampled non-equidistantly.*

According to the results demonstrated above, I declare thesis statement 3 as follows:

Thesis 3

I developed a new filtering procedure (low-pass filtering) based on the 2D IRLS-FT method and applied it to gravity data. I constructed a model of a right rectangular prism to generate 2D synthetic gravity datasets which are initially created over an equidistant grid before being randomized to provide a random-walk sampling design. In my calculations, the inversion procedures are executed using Hermite functions of order M_x=M_y=25 as an inversion parameter for a compromise between estimation accuracy and stability. According to the results on both space and frequency domains, I proved that the inversion-based filtering procedure has a highly reduced outlier sensitivity in the cases of regular and random-walk sampled datasets.

5.5. TESTING THE INCOMPLETE SAMPLING PROBLEM IN 2D

The economic significance of our inversion in solving the issues made from statistics lacking or misplaced at some stage in the sector survey inspired us to maintain our investigations and supply a comprehensive image of ways the delivered inversion professional is. As discussed in the previous chapter of the 2D magnetic data processing, the incomplete sampling problem is evaluated by the IRLS-based inversion on two-dimensional synthetic gravity datasets in accordance with low-pass filtering. This is also to ensure that our inversion approach is highly effective to treat this kind of data sampling scheme and to improve the various data sets obtained by other geophysical methods. Therefore, equidistantly and non-equidistantly (randomly) sampled incomplete gravity datasets are subjected to the inversion method. To achieve such an aim, the two mentioned sampling configurations are evaluated using both noise-free and noisy datasets.

5.5.1. INCOMPLETE GRAVITY DATASET WITH EQUIDISTANT SAMPLING

5.5.1.1. Noise-free Equidistantly Sampled Gravity Datasets

The same noise-free equidistantly sampled gravity datasets taken over 1681 measurements *(Figure 66)* are subjected to the inversion algorithm at the same missing percentages of the magnetic methodology (15%, 25%, 35%, 45%, 50%, and 60%), revealing the related 2D amplitude-frequency spectra in *Figures 85a, b, c, d, e,* and *f* respectively. In our examples, Hermite functions of order $M_x=M_y=25$ are used. The total number of the measurements is randomly selected through the inversion procedure to be 1411, 1269, 1097, and 918 for 15%, 25%, 35%, and 45% missing data cases respectively, while 845 and 656 are constructed for 50% and 60% cases respectively.

It is seen that the estimated spectra for 15%, 25%, 35%, and 45% are virtually the same as those obtained by both the DFT and IRLS of the equidistantly sampled complete datasets *(Figures 67* and *70* respectively*)*. In addition, the spectrum derived when 50% of the measuring points are randomly cancelled is still fulfilled when compared to that deduced in 60% missing data, which is highly distorted.

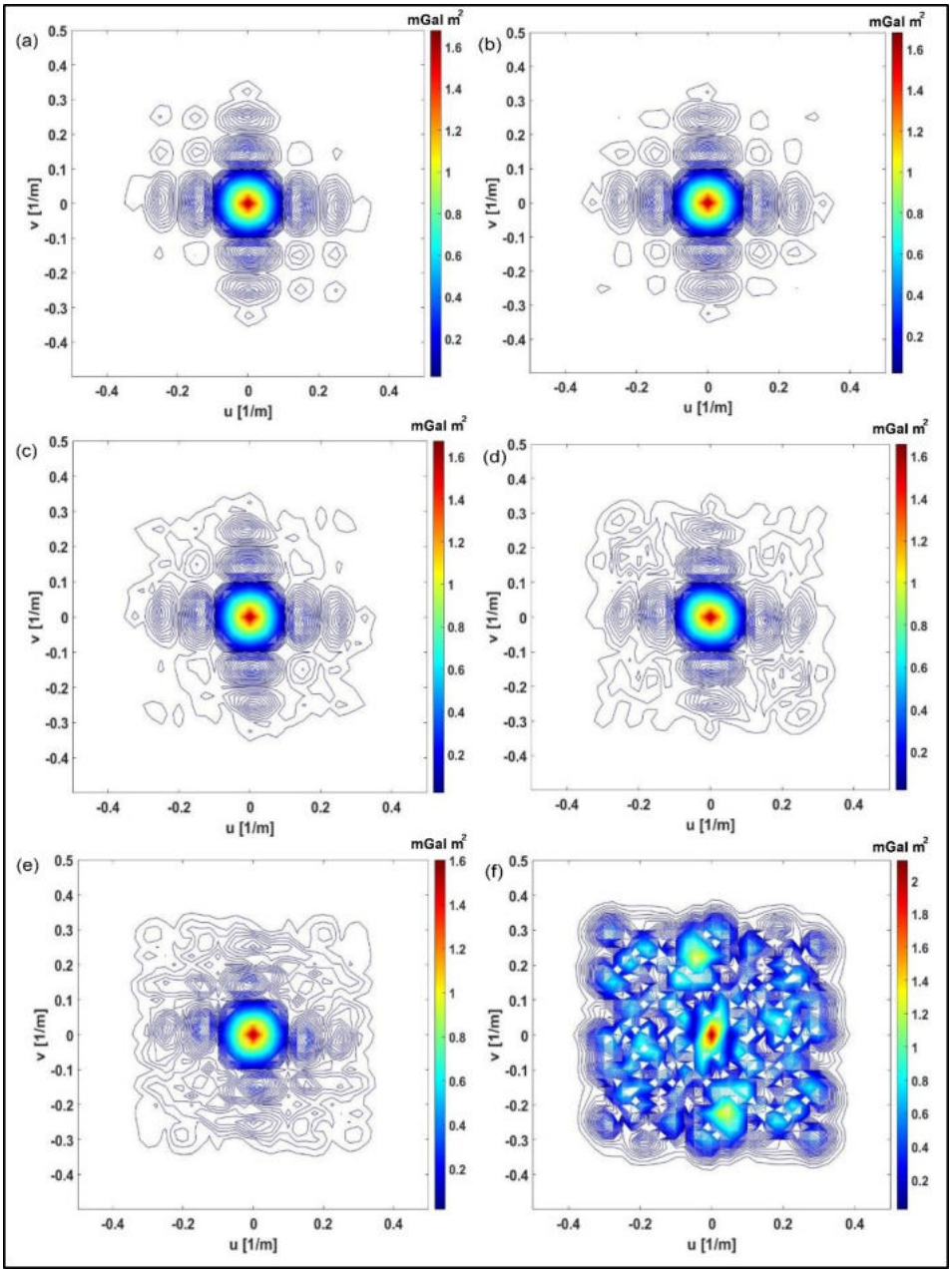

Figure 85. *The 2D IRLS-FT spectra when a) 15% of the equidistantly sampled gravity datasets are missing, b) 25% of the equidistantly sampled gravity datasets are missing, c) 35% of the equidistantly sampled gravity datasets are missing, d) 45% of the equidistantly sampled gravity datasets are missing, e) 50% of the equidistantly sampled gravity datasets are missing, f) 60% of the equidistantly sampled gravity datasets are missing.*

On the other hand, the regional gravity anomaly maps estimated by the inversion-based low-pass filtering of the above-mentioned spectra are illustrated in *Figures 86a, b, c, d, e,* and *f* respectively. Compared to the regional gravity anomaly maps constructed by both the DFT and IRLS of the equidistantly noise-free and noisy sampled complete datasets *(Figures 73* and *75* respectively*)*, there are striking similarities in cases of missing 15%, 25%, 35%, and 45% of measurements, which are numerically supported by low data distance values of 0.0266, 0.0269,

0.0351, and 0.0529, respectively. As in all of our previous examples, these distances increase by increasing the number of excluded measurement data.

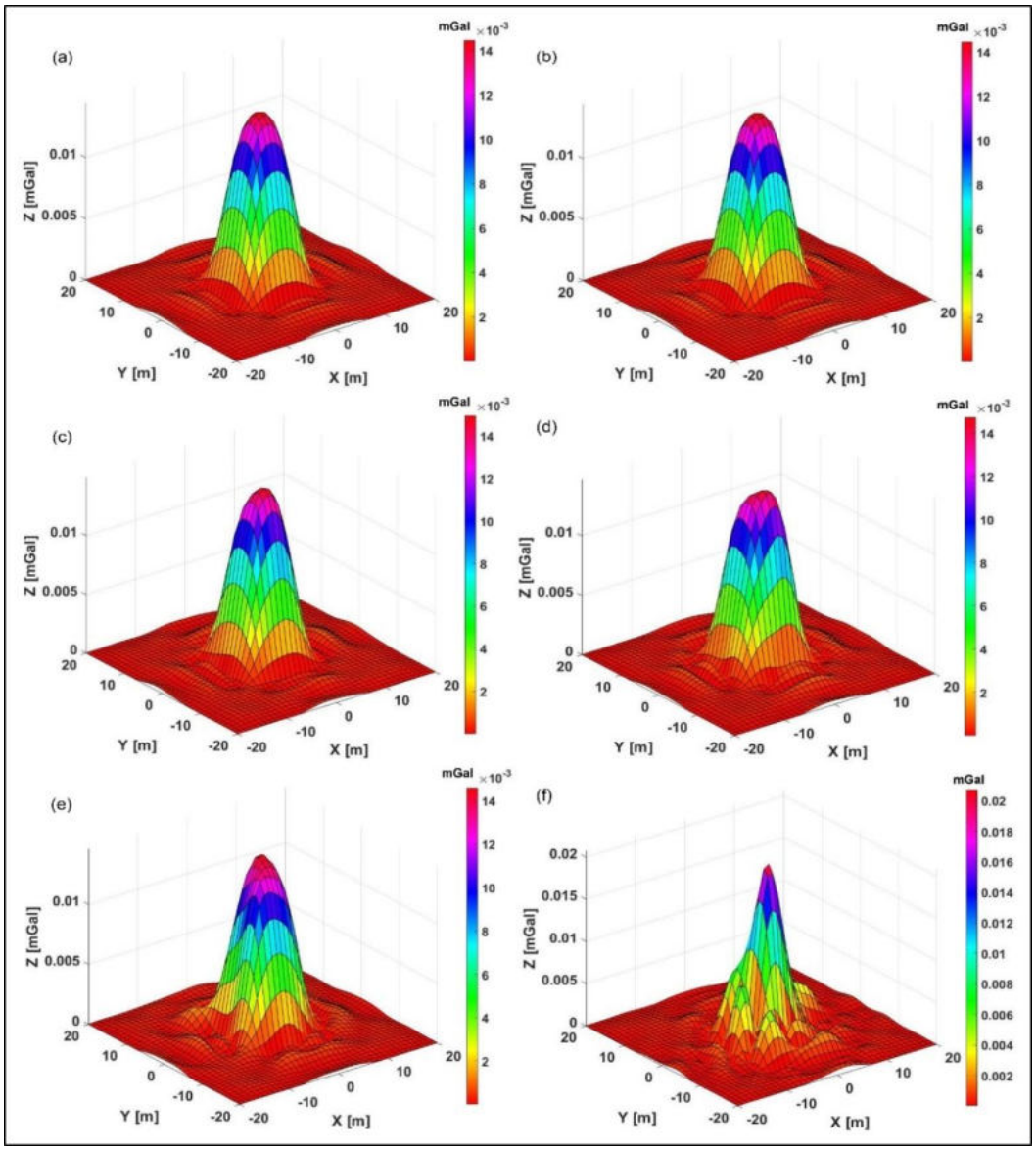

Figure 86. *The regional anomaly maps using IRLS-FT when a) 15% of the equidistantly sampled gravity datasets are missing, b) 25% of the equidistantly sampled gravity datasets are missing, c) 35% of the equidistantly sampled gravity datasets are missing, d) 45% of the equidistantly sampled gravity datasets are missing, e) 50% of the equidistantly sampled gravity datasets are missing, f) 60% of the equidistantly sampled gravity datasets are missing.*

In the 50% missing data case, the mean value of the data distance is 0.0811, whereas a very high data distance value of 0.3831 is given when 60% of the observing data points are randomly neglected. The latter is owing to that as we increase the missing data percentages the number of the measurements becomes closer to the number of the model parameters (M_x M_y=625) and hence, the inverse problem is sounded to be marginally over-determined. As stated in the magnetic methodology, the results show that the IRLS inversion approach is extremely

appropriate for gravitational data processing even when half of the measuring points (50%) are absent as a maximum, where the over-determination ratio has a value of 845/625=1.352.

5.5.1.2. Noise contaminated Equidistantly Sampled Gravity Datasets

To prove the stability and effectiveness of the newly developed inversion method the algorithm is applied to the same noisy datasets sampled equidistantly *(Figure 68)*. The 2D noisy frequency spectra estimated when missing 15%, 25%, 35%, 45%, 50%, and 60% of the datasets are shown in *Figures 87a, b, c, d, e,* and *f* respectively.

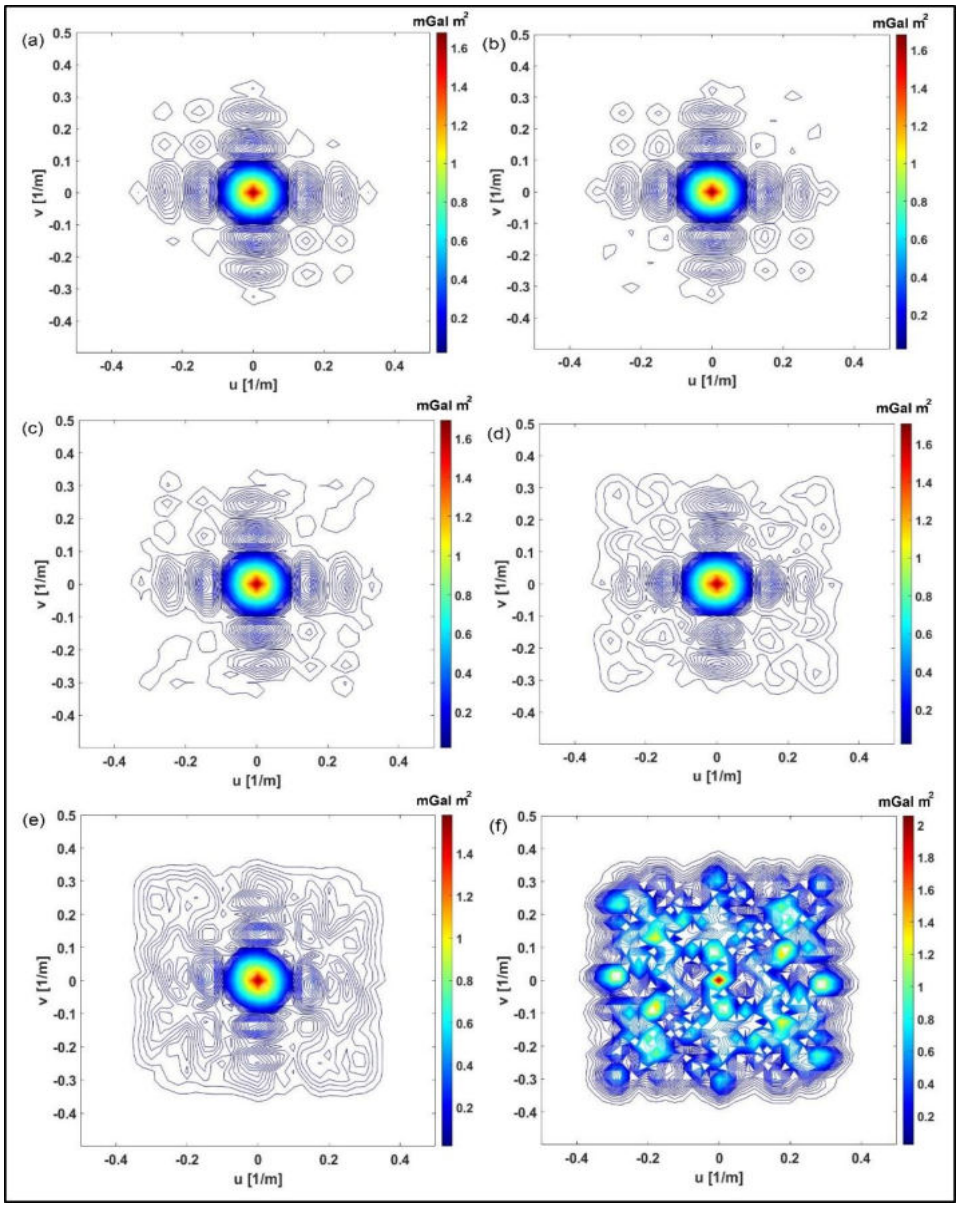

Figure 87. *The 2D IRLS-FT spectra when a) 15% of the equidistantly sampled noisy gravity datasets are missing, b) 25% of the equidistantly sampled noisy gravity datasets are missing, c) 35% of the equidistantly sampled noisy gravity datasets are missing, d) 45% of the equidistantly sampled noisy gravity datasets are missing, e) 50% of the equidistantly sampled noisy gravity datasets are missing, f) 60% of the equidistantly sampled noisy gravity datasets are missing.*

A total of 1422, 1274, 1100, 913, 834, and 651 observations are used for the missing percentages respectively. It is noticed that all the spectra are nearly similar when compared to those found in the equidistantly sampled noise-free datasets (*Figures 85a, b, c, d, e,* and *f* respectively) reflecting the higher inversion adaptability. Furthermore, the 2D inverse Fourier transformations as designated by the regional gravity anomaly maps are presented in *Figures 88a, b, c, d, e,* and *f* respectively.

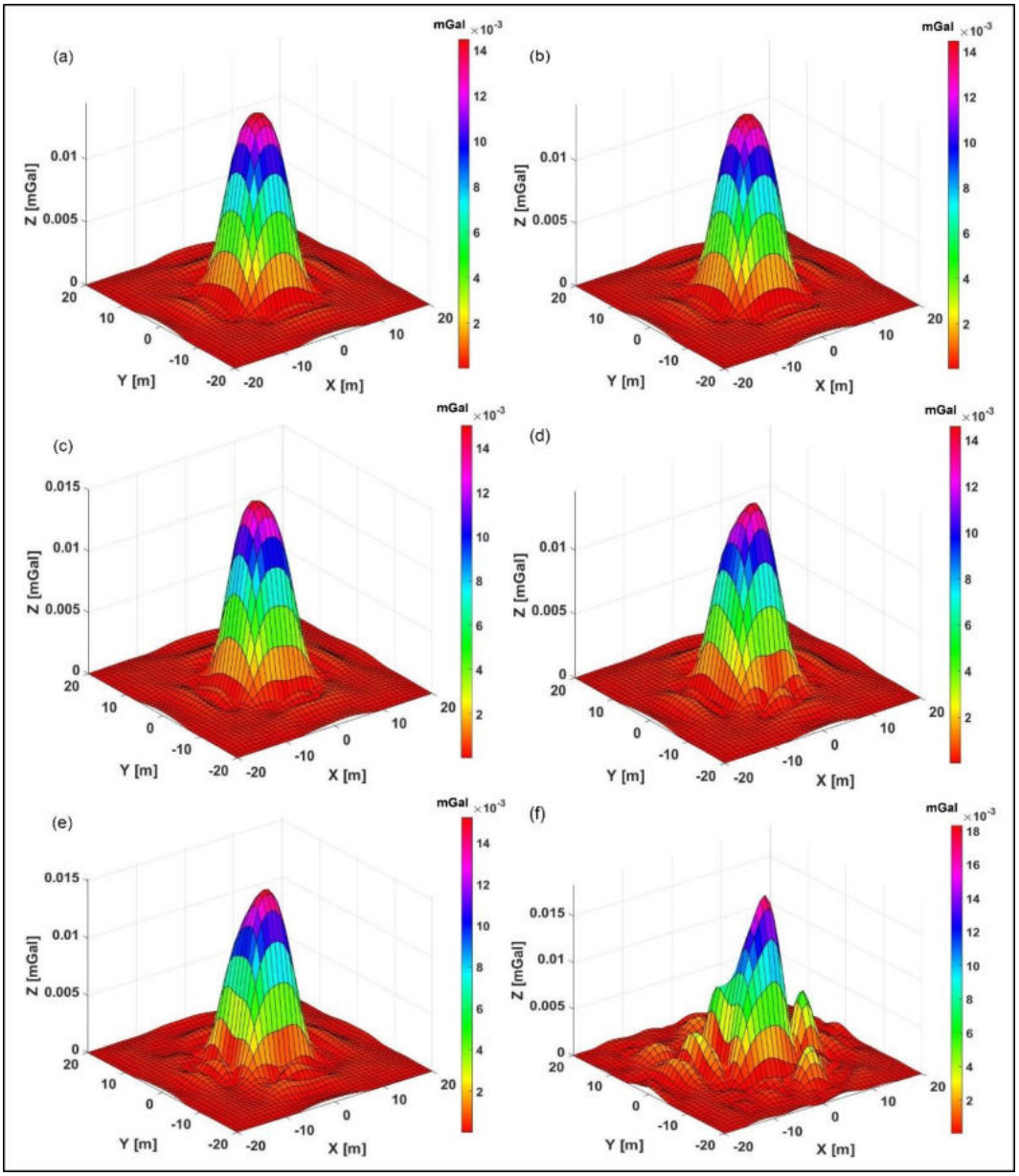

Figure 88. *The regional anomaly maps using IRLS-FT when a) 15% of the equidistantly sampled noisy gravity datasets are missing, b) 25% of the equidistantly sampled noisy gravity datasets are missing, c) 35% of the equidistantly sampled noisy gravity datasets are missing, d) 45% of the equidistantly sampled noisy gravity datasets are missing, e) 50% of the equidistantly sampled noisy gravity datasets are missing, f) 60% of the equidistantly sampled noisy gravity datasets are missing.*

The filtered gravity maps demonstrate low data distance values of 0.0270, 0.0275, 0.0366, and 0.0547 for 15%, 25%, 35%, and 45% missing data cases respectively. When only half of the readings (the percentage at which the inverse problem is completely over-determined 834/625=1.334) are used, the distance is computed as 0.0823, however, for the 60% case, it is calculated as a much higher value of 0.3967 due to the marginal over-determination of the inverse problem. According to these findings, we can conclude that our introduced inversion is robust, stable, and effective for processing gravity datasets-based noise contamination, and therefore reinforces its economically meaningful in the different geophysical explorations.

5.5.2. INCOMPLETE GRAVITY DATASET WITH NON-EQUIDISTANT SAMPLING

5.5.2.1. Noise-free Non-equidistantly Sampled Gravity Datasets

Our analysis has been expanded to the 2D non-equidistantly sampled noise-free gravity datasets (*Figure 78)*. *Figures 89a, b, c, d, e,* and *f* show the 2D frequency spectra estimated at missing percentages of 15%, 25%, 35%, 45%, 50%, and 60% respectively. Measurements of 1429 are used for 15% missing case, 1262 for 25%, and 1083 for 35%, while in 45%, 50%, and 60% missing percentages, a total of 933, 854, and 634 are produced respectively. In all our calculations, we utilize the same order of Hermite Functions (M=25 in both *x* and *y*- directions). Great similarities in both anomaly shapes and amplitudes are presented when compared to those proceeded in the noise-free incomplete gravity data sampled equidistantly *(Figures 85a, b, c, d, e,* and *f* respectively*)*.

For the given missing data percentages, the 2D Fourier transformation of the noise-free spectra are used in Eq. 103 and Eq.104 to produce the regional gravity anomaly maps as seen in *Figures 90a, b, c, d, e,* and *f* respectively. The filtered maps are comparable to those established with the equidistantly sampled noise-free incomplete gravity data *(Figures 86a, b, c, d, e,* and *f* respectively*)*. The numerical estimates of mean distances for 15%, 25%, 35%, 45% missing data cases are 0.0268, 0.0317, 0.0426, and 0.0639 respectively. In addition, well-satisfied results are obtained in 50% missing case which has a distance of 0.0964. In contrast, extreme deformation and distortion are readily visible when 60% of the measurements are randomly ignored, as numerically evidenced by a substantially larger data distance value of 0.4867. In our cases, the marginal overdetermination ratio of the inverse problem persists as we increase the percentages of the missing data by over 50% (the percentage at which the inverse problem is completely over-determined 854/625=1.3664).

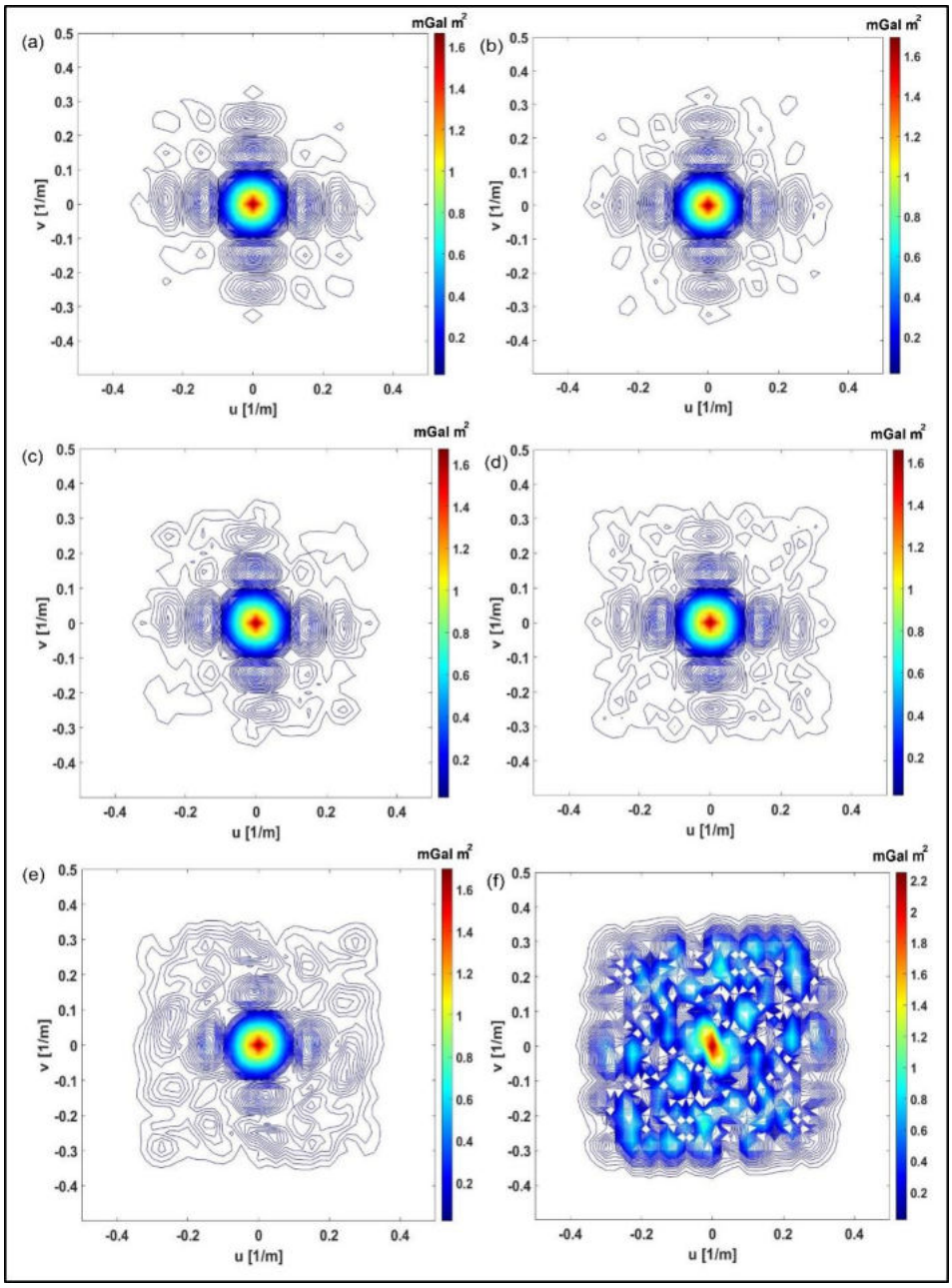

Figure 89. *The 2D IRLS-FT spectra when a) 15% of the non-equidistantly sampled gravity datasets are missing, b) 25% of the non-equidistantly sampled gravity datasets are missing, c) 35% of the non-equidistantly sampled gravity datasets are missing, d) 45% of the non-equidistantly sampled gravity datasets are missing, e) 50% of the non-equidistantly sampled gravity datasets are missing, f) 60% of the non-equidistantly sampled gravity datasets are missing.*

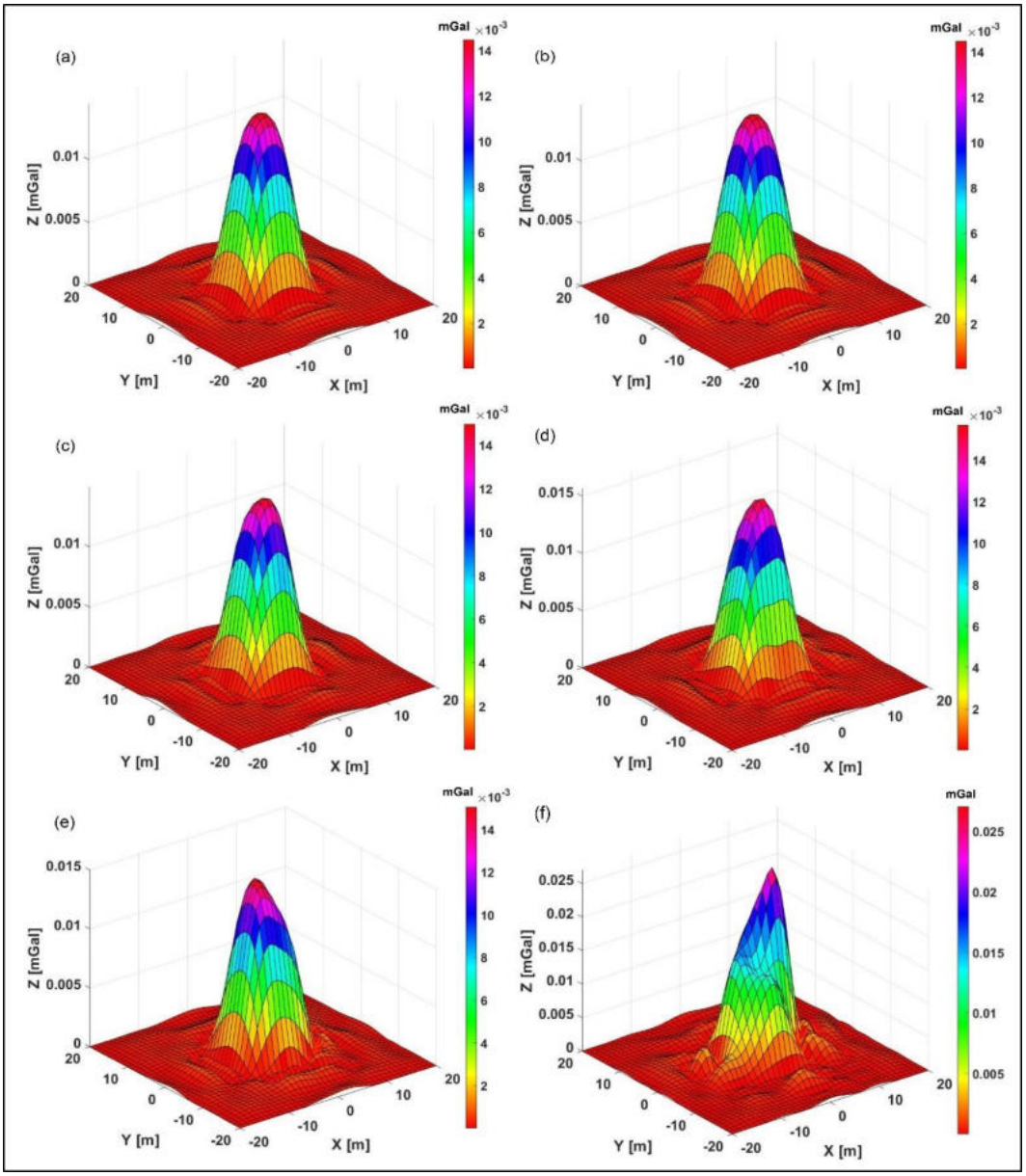

Figure 90. *The regional anomaly maps using IRLS-FT when a) 15% of the non-equidistantly sampled gravity datasets are missing, b) 25% of the non-equidistantly sampled gravity datasets are missing, c) 35% of the non-equidistantly sampled gravity datasets are missing, d) 45% of the non-equidistantly sampled gravity datasets are missing, e) 50% of the non-equidistantly sampled gravity datasets are missing, f) 60% of the non-equidistantly sampled gravity datasets are missing.*

5.5.2.2. Noise contaminated Non-equidistantly Sampled Gravity Datasets

The incomplete sampling problem based on the 2D randomly noise-contaminated gravity datasets shown in *Figure 80* is assessed to prove the inversion efficiency and applicability. *Figures 91a, b, c, d, e,* and *f* illustrate the 2D amplitude-frequency spectra for measurements of 1432, 1277, 1061, 938, 857, and 640 at 15%, 25%, 35%, 45%, 50% and 60% missing data percentages respectively. All the estimated spectra through the inversion procedure are nearly

identical when compared to those found in the equidistantly sampled noise-free situations (*Figures 85a, b, c, d, e,* and *f* respectively).

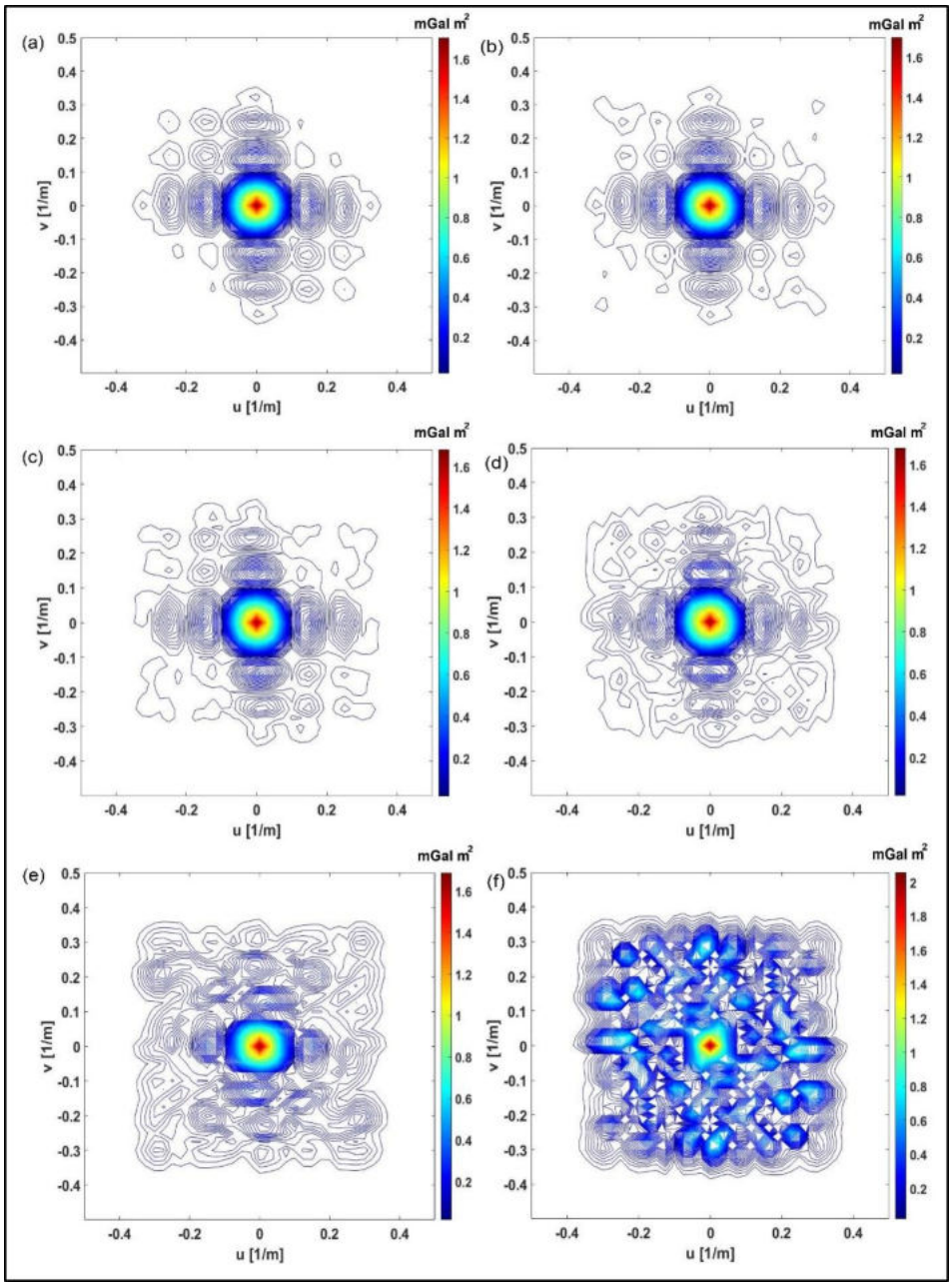

Figure 91. *The 2D IRLS-FT spectra when a) 15% of the non-equidistantly sampled noisy gravity datasets are missing, b) 25% of the non-equidistantly sampled noisy gravity datasets are missing, c) 35% of the non-equidistantly sampled noisy gravity datasets are missing, d) 45% of the non-equidistantly sampled noisy gravity datasets are missing, e) 50% of the non-equidistantly sampled noisy gravity datasets are missing, f) 60% of the non-equidistantly sampled noisy gravity datasets are missing.*

The regional gravity anomaly maps at the missing percentages stated above are shown in *Figures 92a, b, c, d, e,* and *f* respectively. The stability of our IRLS inversion approach is numerically proven by low mean distance values of 0.0274, 0.0328, 0.0439, and 0.0647 when

15%, 25%, 35%, and 45% of the measuring points are randomly cancelled respectively. The value of the mean distance is 0.0976 in the 50% missing case, while the less over-determined ratio of the inverse problem associated with missing 60% of the data leads to a higher data distance value of 0.4881. These findings in both frequency and space domains indicate that half of the measuring data points (50%) are quite enough to judge the inversion efficiency (the over-determination ratio is 857/625=1.3712), opening up a new and economically significant way of handling the incomplete sampling problems encountered during field data acquisition.

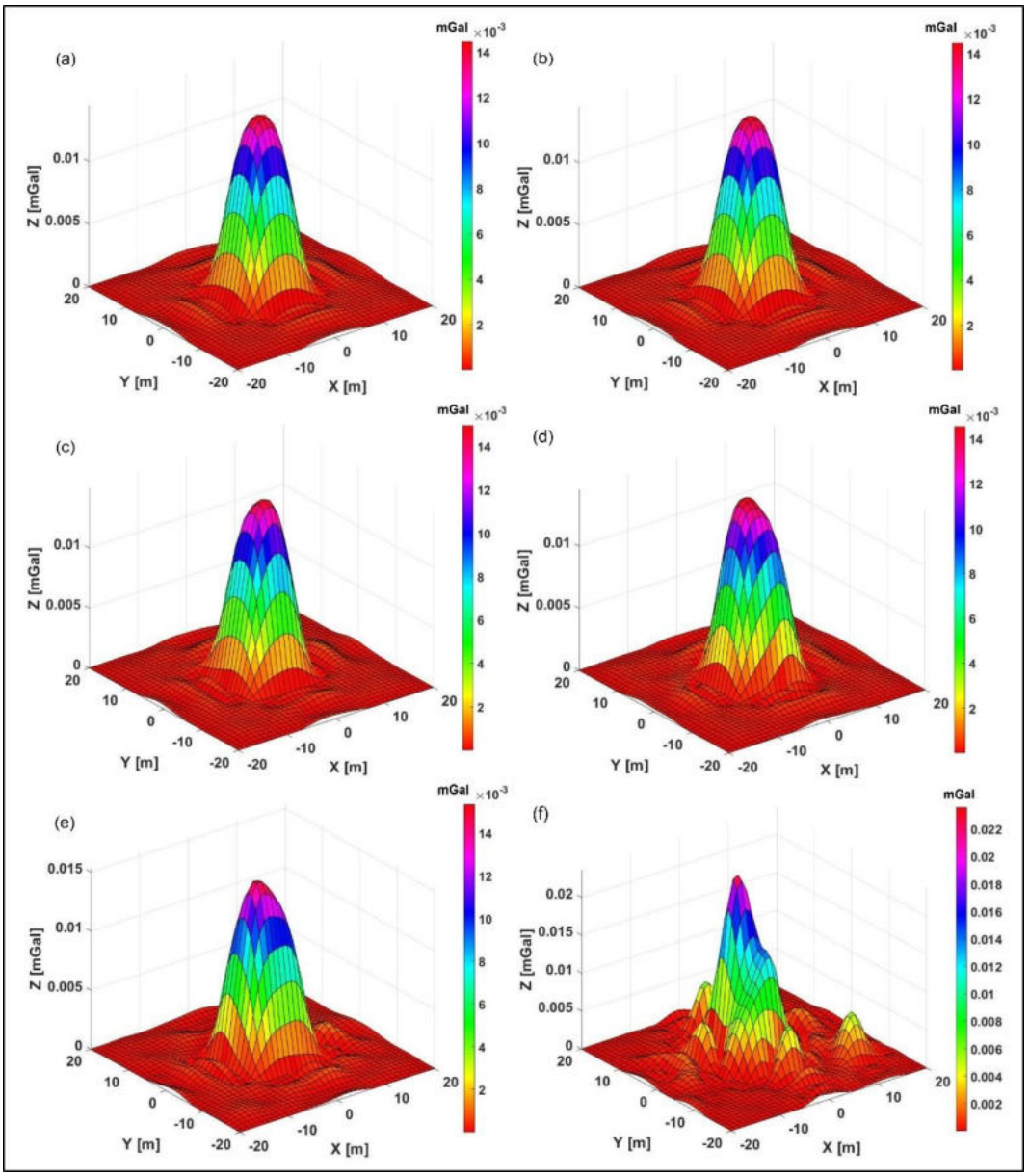

Figure 92. *The regional anomaly maps using IRLS-FT when a) 15% of the non-equidistantly sampled noisy gravity datasets are missing, b) 25% of the non-equidistantly sampled noisy gravity datasets are missing, c) 35% of the non-equidistantly sampled noisy gravity datasets are missing, d) 45% of the non-equidistantly sampled noisy gravity datasets are missing, e) 50% of the non-equidistantly sampled noisy gravity datasets are missing, f) 60% of the non-equidistantly sampled noisy gravity datasets are missing.*

According to the results demonstrated above, I declare thesis statement 4 as follows:

Thesis 4

I gave an extended analysis of the 2D IRLS-FT-based filtering method in terms of the missing gravity data problems in a numerical example containing 41x41 measurement points and 25x25 unknown expansion coefficients.

In my analysis of the missing data filtering problem on noise-free and Cauchy noise-contaminated gravity datasets (containing outliers), I found that the inversion algorithm is stable and highly applicable for handling both the equidistantly and random-walk sampled incomplete gravity datasets even when 50 percent of the data measurements are cancelled or missing. In contrast, distorted anomaly shapes and amplitudes are provided at higher rates of the missing data percentages (when the over-determination rate is below 1.3).

Chapter 6
FIELD EXAMPLE: A CASE STUDY WITHIN SINAI PENINSULA, EGYPT

6.1. STUDY AREA AND GEOLOGICAL SETTING

In the last few decades, Egypt is developing a strategy to entice Egyptians to reside in the Sinai Peninsula by completing economic projects and offering employment possibilities and also to lessen the overwhelming population in the vicinity of the Nile Valley. In the framework of the national Sinai development plan (NSDP), numerous developed communities and settlements, including those involved in mining, tourism, agriculture, and industrial operations have been constructed (ICG, 2007). The Sinai Peninsula is situated in Egypt's northeastern region. It is shaped like a triangle, with its base to the north running along the Mediterranean coast for roughly 210 kilometers, and Ras Mohamed, where it crests to the south. It is bordered to the east by the Gulf of Aqaba and the international boundary, and to the west by the Gulf of Suez and the Suez Canal. Said (1962) has recognized the Sinai Peninsula into three major portions. The first is the southern complex of high mountains such as Sant Katherine (highest peak), Um Shumar, and Al-Thabt. El-Tih and El-Egma plateaus, which make up approximately 40% of Sinai, are placed in the center of the peninsula and slope northward toward Wadi El Arish. On the other hand, a lengthy and parallel row of dunes, some of which are exceeding 100 meters high, isolate the Sinai Peninsula's northern mountains and hills from the Mediterranean Sea coast. The study area lies in the Egma-Tih plateau in the western central part of Sinai Peninsula, Egypt between latitudes 29° 16" to 29° 41" N and longitudes 33° 16" to 34° 00" E as shown in *Figure 93*.

Because of the majority of Sinai's natural resources have not yet been properly utilized to their full potential, the study area, which is inhabited by the Bedouin population still needs to more development and reconstruction. This appears clearly by the land use or land cover map illustrated in *Figure 94*, which is established by Esri's recent Global Land Cover Map (GLCM) based on Sentinel 2 data as well. The land cover map shows that most of the study area is scrubland or bare ground, where a loss in the artificial cover is provided due to human activities. However, there is little consideration given to either crops or built areas attributable to a scarcity of daily use water, where the Bedouin population lives in the valleys and relies mostly on rainfall.

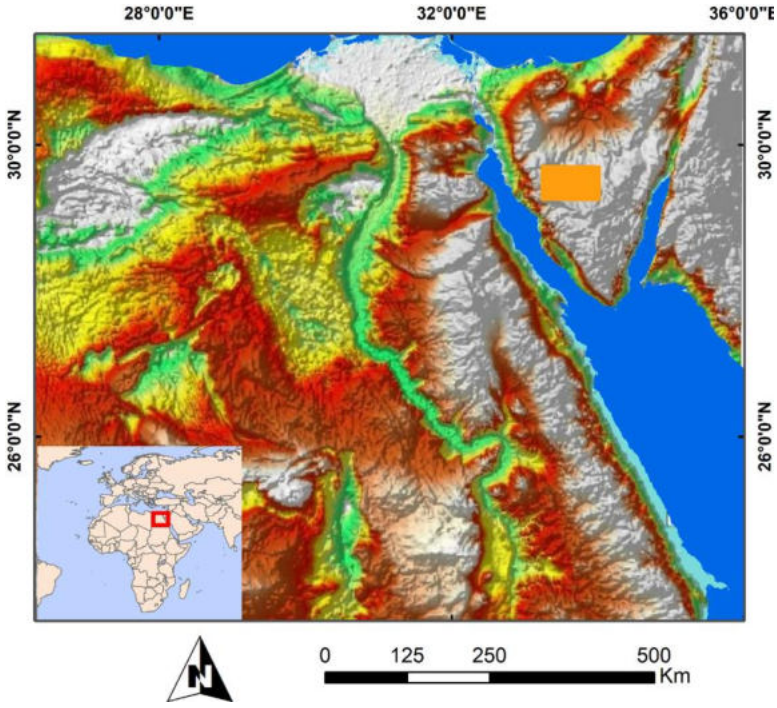

Figure 93. *Location map of the area under investigation.*

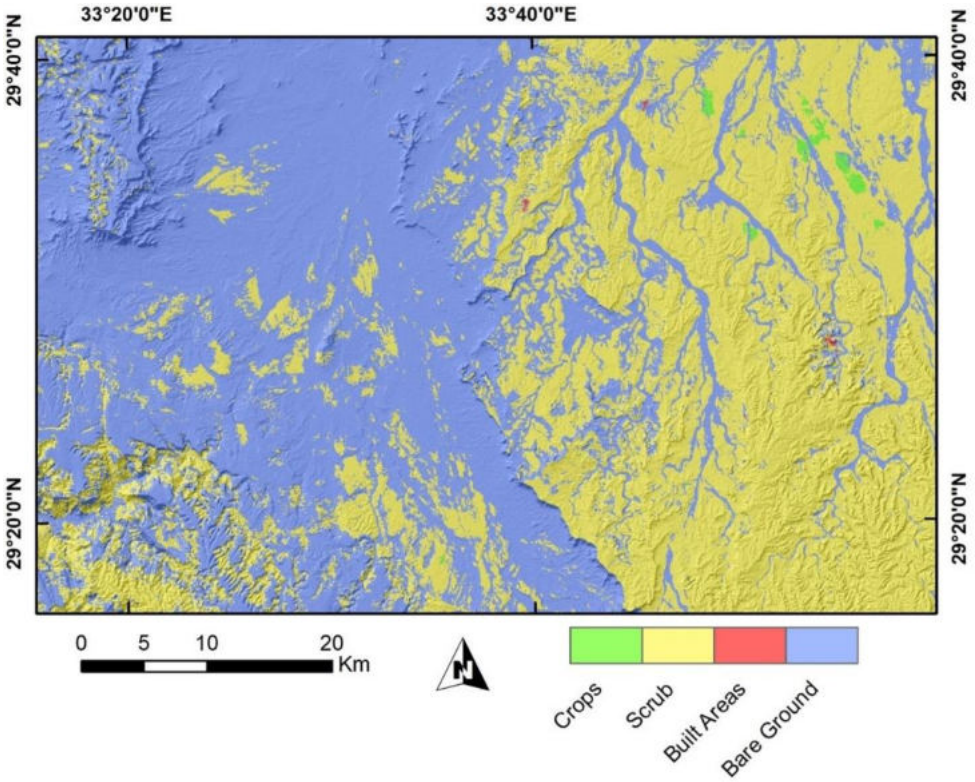

Figure 94. *Land use/land cover map of the study area.*

Furthermore, the surface geological setting of the investigated area is characterized by several rock units. *Figure 95* employs high spectral and spatial resolution sentinel 2 data to construct an accurate geological map of the study area based on the spectral characteristics of each rock unit (Shebl et al., 2021 and 2022). According to the efficiency of the lithological mapping-based machine learning procedures, the lithological map is performed using a support vector machine (SVM). It is indicated that the Wadi deposits and Alluvial Hamadah deposits, which date from the Holocene and Pleistocene ages respectively, constitute the Quaternary sediments. Egma Formation, which is primarily comprised of chalky limestone, serves as a representative of the Lower Eocene deposits, which extend widely and almost cover the majority of the study area.

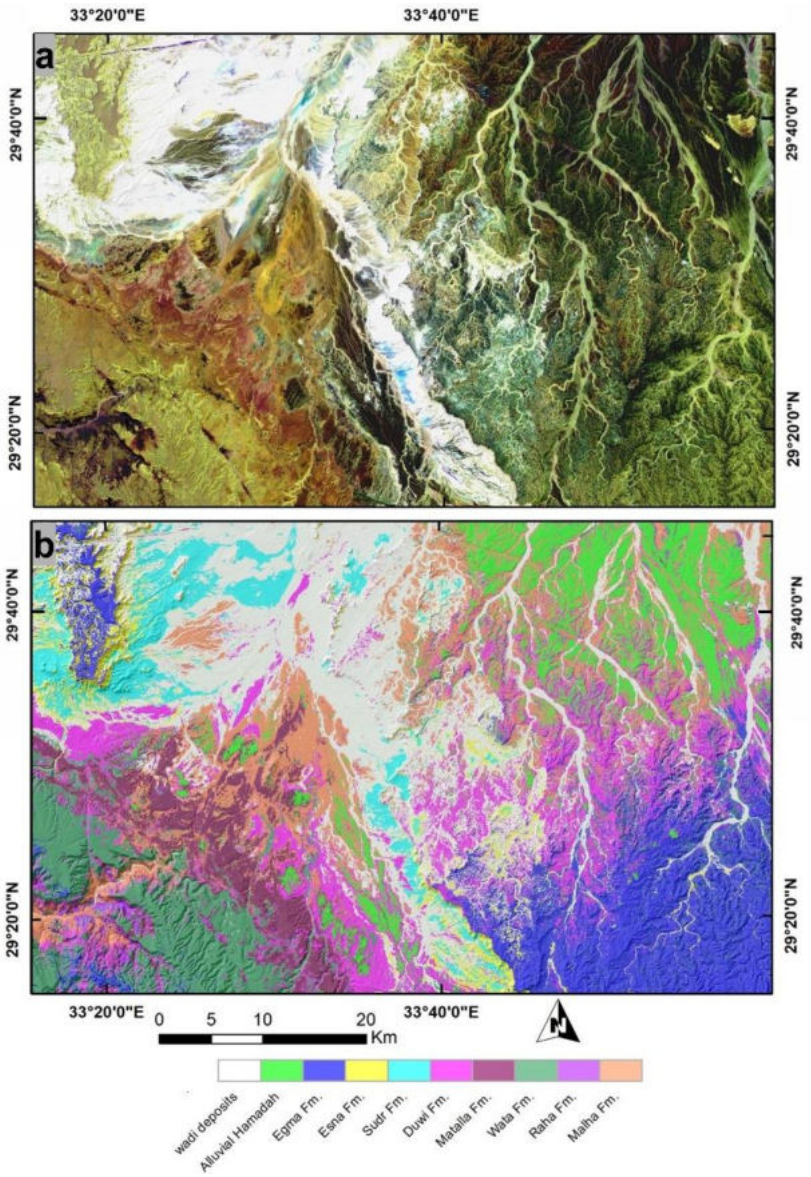

Figure 95. *a) FCC 12/11/2 RGB for lithological differentiation, b) lithological map of the study area created using SVM.*

In addition, shale makes up the majority of the Paleocene deposits represented by the Esna Formation, which are concentrated in the study area's Northwestern portion. The Maastrichtian-aged Sudr Formation is found in the study region from the north to the south and is primarily composed of white to pale grey chalk. Moreover, Duwi Formation is widespread in the study region and is of the Campanian age. It is mostly made up of alternating beds of clastics and carbonates. Matulla Formation, which is of Coniacian and Santonian age and primarily made of limestone, belt-shaped stretched from the western to central south. The study area's southern region is dominated by the Turonian-aged Wata Formation, which is primarily made up of dolomitic limestone. Raha Formation is found in the southwest of the study area and is of the Cenomanian age. It is mostly made up of alternating beds of dolomitic limestone, marl, and claystone. Finally, the Malha Formation (Nubian sandstone), which is formed primarily of kaolinized sandstone intercalated with mudstone, paleosoil, and conglomerate, represents the Lower Cretaceous rocks that are only slightly distributed in the southwest of the study area. More details about the geology of central Sinai have been previously described by many authors (Hume, 1906; Shata, 1956; Said, 1990; Hassanin, 1997; Ali, 2006).

Due to its diversity of both simple and complex structural forms, the Sinai Peninsula is regarded as one of the regions that attain considerable attention from a structural point of view. The surface structural lineaments play a vital role in different purposes, for example as good indicators for groundwater recharge and other geotechnical and environmental aspects. Therefore, we present a lineament density map of the study area utilizing PALSAR DEM (Shebl and Csámer, 2021) as depicted in *Figure 96a*. The surface lineaments are seen to be widely dispersed throughout the study area, with little representation in the southwest. Additionally, the slope map of the area under investigation is demonstrated in *Figure 96b*. It is noticed that the study area is mostly characterized by gentle and moderate slopes, while the central and western portions are predominately steep slopes due to the presence of wadi cliffs and rock outcrops respectively.

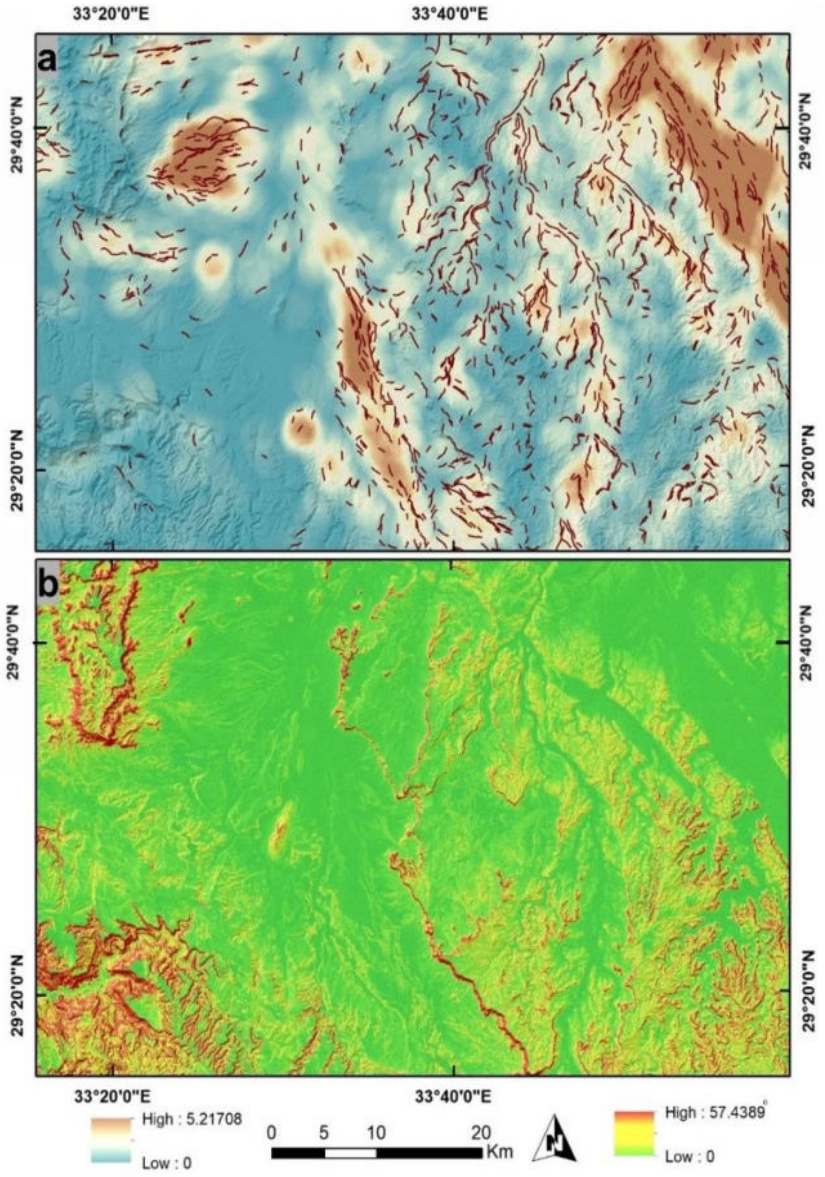

Figure 96. *a) Lineaments extracted from ALOS PALSAR DEM over their density map, b) Slop map in degrees of the study area.*

More importantly, as an extension of the Oasis montaj software program for performing layer-earth models, GM-SYS 3D (Geosoft, 2014) is designed to generate a 3D basement relief magnetic map of the study area as shown in *Figure 97*. In our investigation, the model further assumes that there are two main layers in the subsurface: one for the sediments, with magnetic susceptibility zero, and another for the basement, with magnetic susceptibility referenced roughly by (Araffa et al., 2015) as 0.00075 CGS unit. In addition, the magnetic field strength, field inclination, and field declination input values must also be supplied for the 3D magnetic modeling. The corresponding values are 43134 nT, 43.58°, and 3.31° respectively. The 3D magnetic modeling indicates that the depth of the basement surface varies by about 200 meters,

going from shallow in the western portion of the research region (about 1658 m below sea level) to deeper in the eastern half (about 1857 m below sea level).

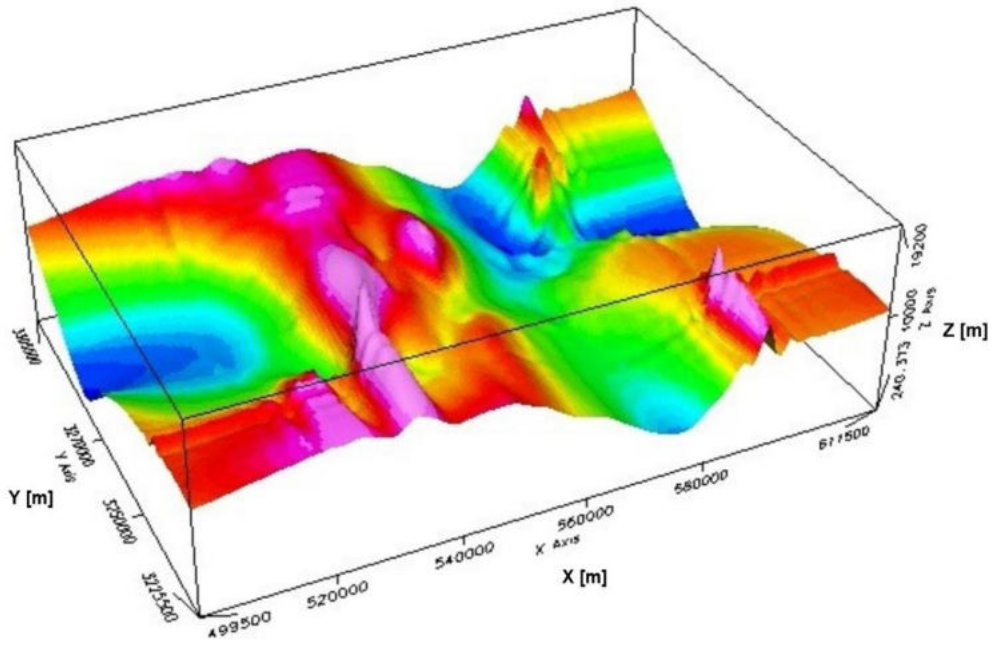

Figure 97. *3D magnetic modeling of the basement relief surface in the study area.*

6.2. 2D IRLS-FT BASED LOW-PASS FILTERING

6.2.1. A FIELD EXAMPLE WITH EQUIDISTANT SAMPLING

More interestingly, the efficiency and stability of our inversion-based 2D Fourier transformation (2D IRLS-FT) can be verified by applying its algorithm to real field data measurements. Consequently, we present a case study that relies on real gravity datasets measured in the investigated area mentioned above. It is well known that the biharmonic spline interpolation gridding approach is capable of resolving the majority of gridding challenges even when the highly noisy measured data sets are distributed over a non-regular sampling grid. In that regard, the gravity field data measurements carried out in the study area were interpolated to be set along with an equidistant sampling grid as shown in *Figure 98*. These gravity observations were taken in the area under exploration using CG-3 Autograv (Automated Gravity Meter) made by Scintrex. The instrument can be employed in both comprehensive local investigations and regional surveys or large-scale geodetic studies since it can combine measurements to range over 7000 mGal without resetting and record with a reading resolution of 0.01 mGal (CG-3 Autograv Manual, 1995). In addition, each gravity station's elevation and location (latitude and longitude) were recorded using GPS for use in the subsequent data

adjustments. *Figure 98* demonstrates that 441 equidistant measuring stations were sampled utilizing intervals of 3347 m and 2381 m in *x* and *y* directions respectively.

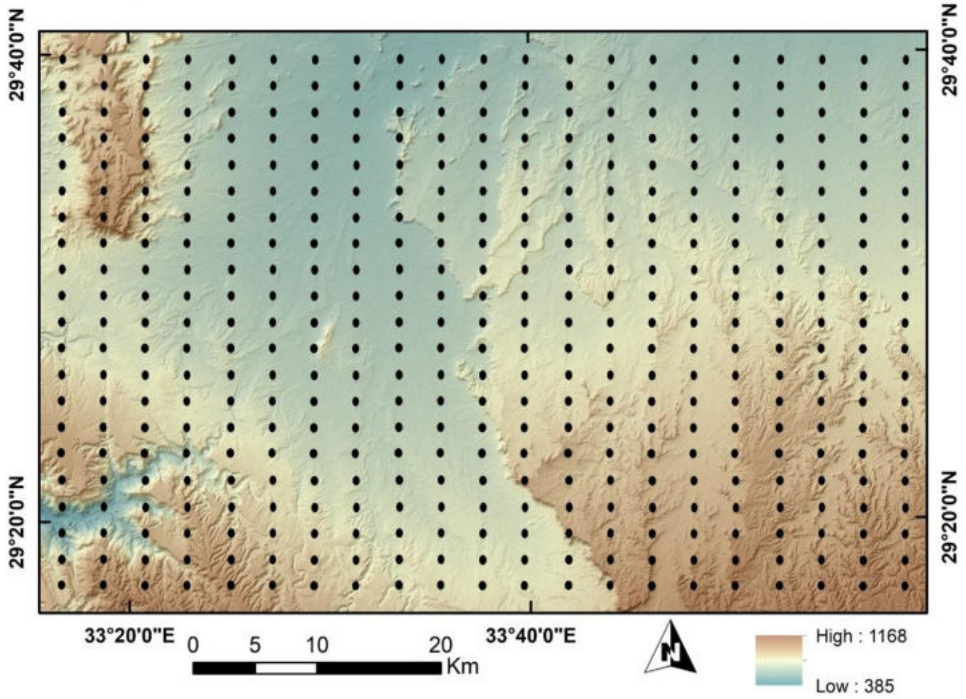

Figure 98. *Location map of the equidistantly measured gravity stations over DEM.*

After gravity data corrections are made to all the readings, the measurements are plotted to obtain the Bouguer gravity anomaly contour map *(Figure 99)*, which is also visualized in a 3D surface map as clearly seen in *Figure 100*.

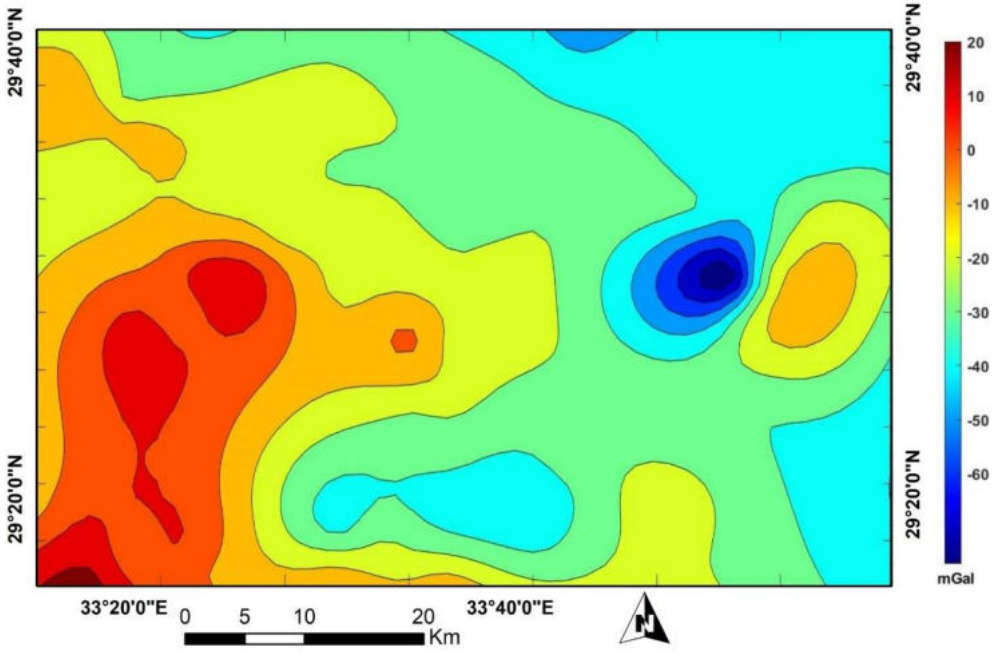

Figure 99. *Bouguer gravity anomaly map of the datasets sampled equidistantly.*

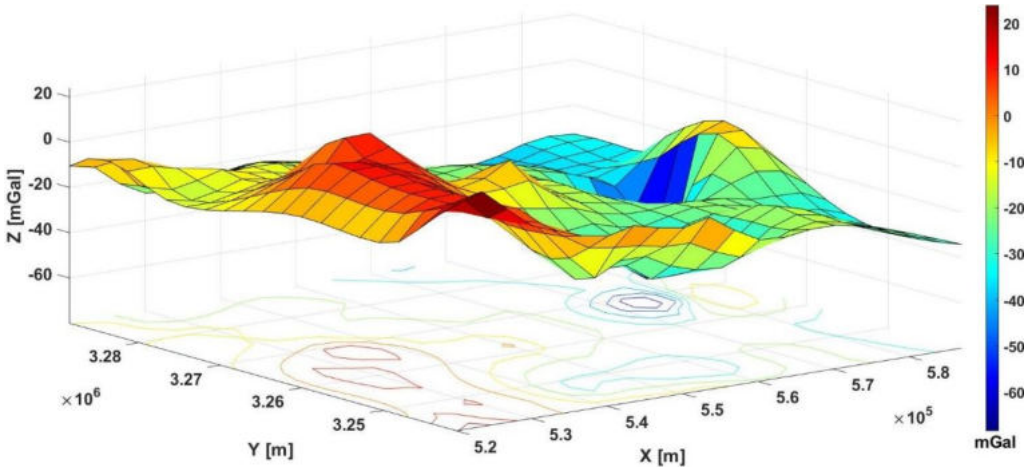

Figure 100. *3D Bouguer gravity anomaly surface of the datasets sampled equidistantly.*

It is observed that the Bouguer gravity anomaly map of the study area indicates the lateral changes in the earth's gravitational field, which is characterized by the presence of relatively high and low positive and negative gravity anomalies of varying sizes and magnitudes. The range of the total intensity gravity values is approximately -60 mGal to 20 mGal. Generally, the Bouguer gravity anomaly map illustrates an increase in the gravity anomaly from east to west which is probably attributed to the density increase associated with the uplifted basement block at the western part of the study area as described earlier in the 3D magnetic modeling of the basement relief surface *(Figure 97).*

The algorithm of the inversion-based FT is then implemented on the above-mentioned equidistantly measured gravity data sets to estimate the 2D Fourier frequency spectrum. Hermite functions of order M_x=M_y=13 were chosen to achieve both accuracy and stability of the inversion procedure. In the framework of the low-pass filtering approach, the spectrum is converted to space domain giving the 2D regional gravity anomaly map and the corresponding 3D regional surface as depicted in *Figures 101* and *102* respectively. It is obvious that the regional gravity anomaly datasets illustrate the gravity impacts brought about by the deep structures that are broad and extend widely in the studied area. Furthermore, the 3D regional gravity anomaly surface of the equidistant measuring datasets becomes more smoother and has virtually equal gravity amplitudes (-60 mGal to 20 mGal) when compared to that of the 3D Bouguer anomaly surface *(Figure 100).* These results allow us to conclude that the 2D IRLS inversion-based Fourier transformation employed in our study is very efficient, reliable, and capable of handling and processing the real field gravitational data sets sampled over equidistant grids.

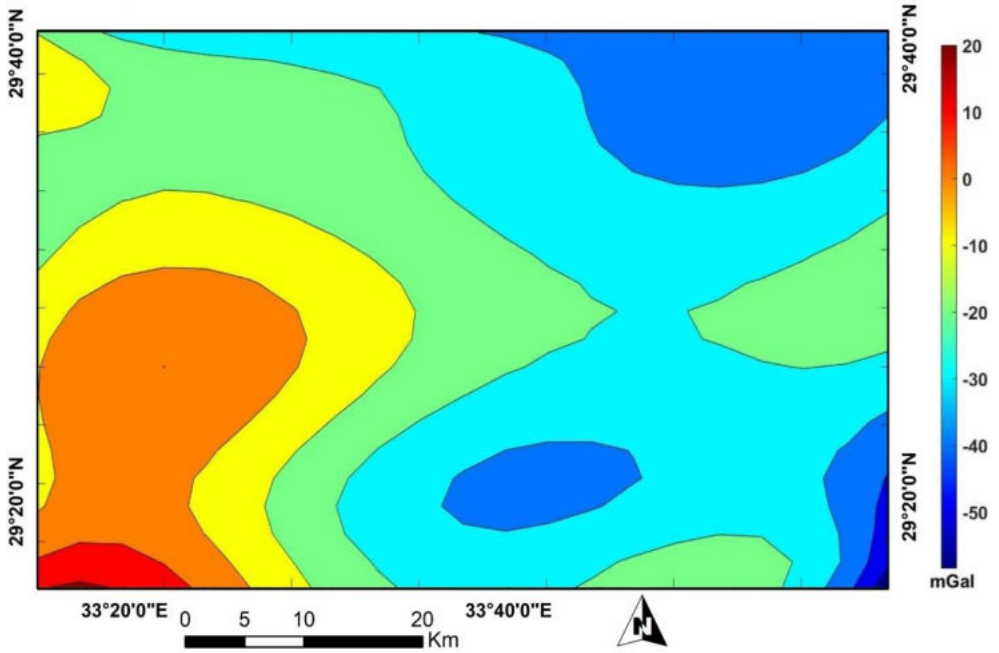

Figure 101. *Regional gravity anomaly map of the datasets sampled equidistantly.*

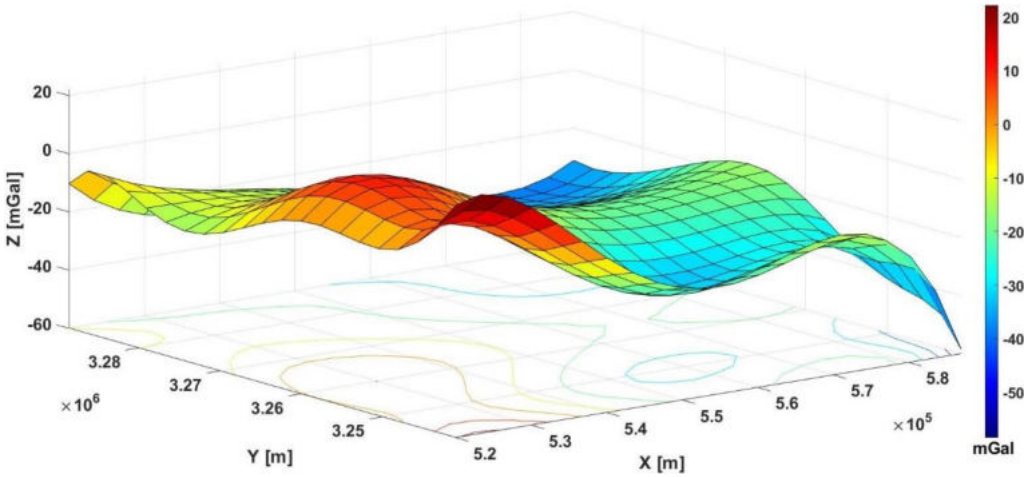

Figure 102. *3D Regional gravity anomaly surface of the datasets sampled equidistantly.*

6.2.2. A FIELD EXAMPLE WITH NON-EQUIDISTANT SAMPLING

Due to the presence of several restrictions that confront many geophysicists during the field survey, some of which are natural and others are artificially created, most of the measured datasets are taken over non-regular grids. Therefore, the stability and efficiency of our inversion-based 2D Fourier transformation algorithm are evaluated on 2D non-equidistantly sampled field gravitational datasets. To achieve our purpose, we randomize the regular gravity readings displayed in *Figure 98* to generate the non-equidistant sampling grid as shown in *Figure 103*. To demonstrate how successful our inversion with irregular sampling processes is, the same number of field data measurements (441 measuring points) is used.

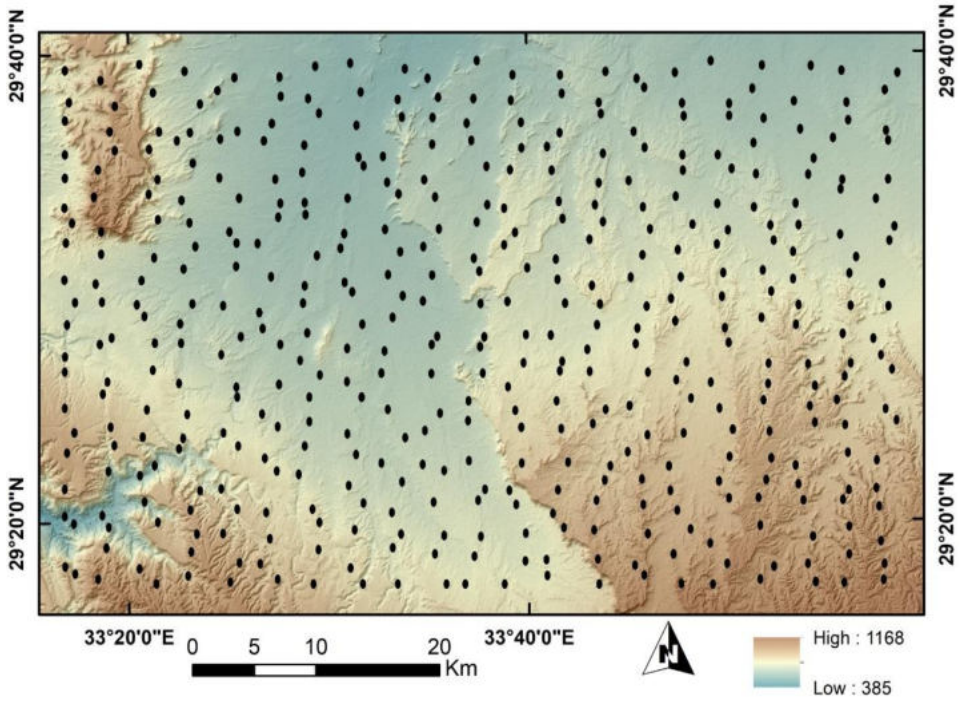

Figure 103. *Location map of the non-equidistantly measured gravity stations over DEM.*

Similarly, the Bouguer gravity anomaly map and the related 3D anomaly surface are demonstrated in *Figures 104* and *105* respectively. Compared to the Bouguer anomaly map of the equidistant measuring datasets *(Figure 99)* as well as the 3D anomaly surface *(Figure 100)*, the extremely like-sharped contour lines do not cover the whole surveyed area, revealing random distributions of rectangular blocks in *x* and *y* directions.

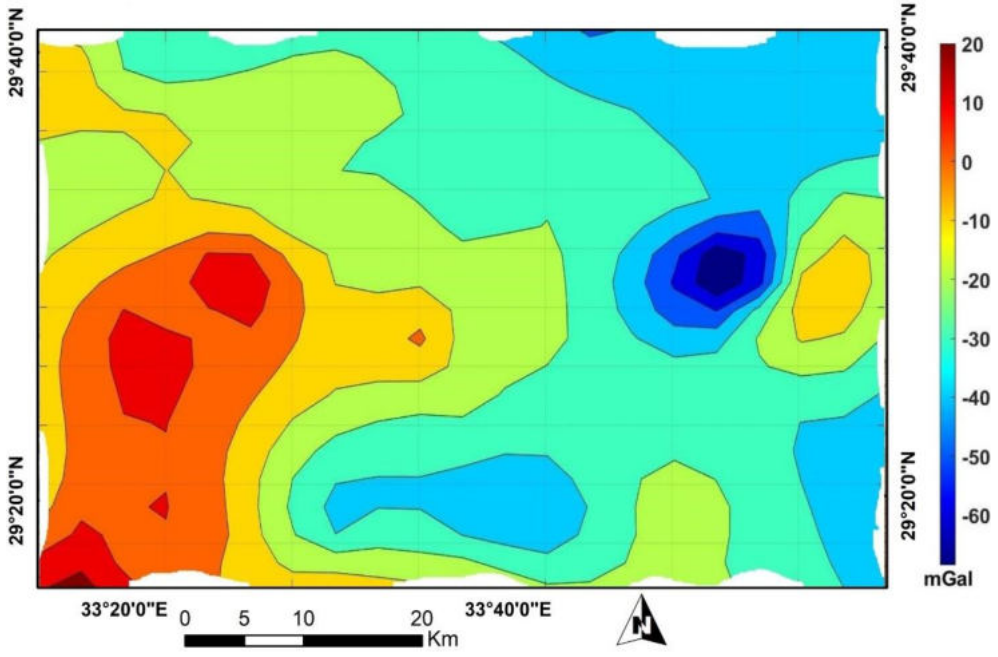

Figure 104. *Bouguer gravity anomaly map of the datasets sampled non-equidistantly.*

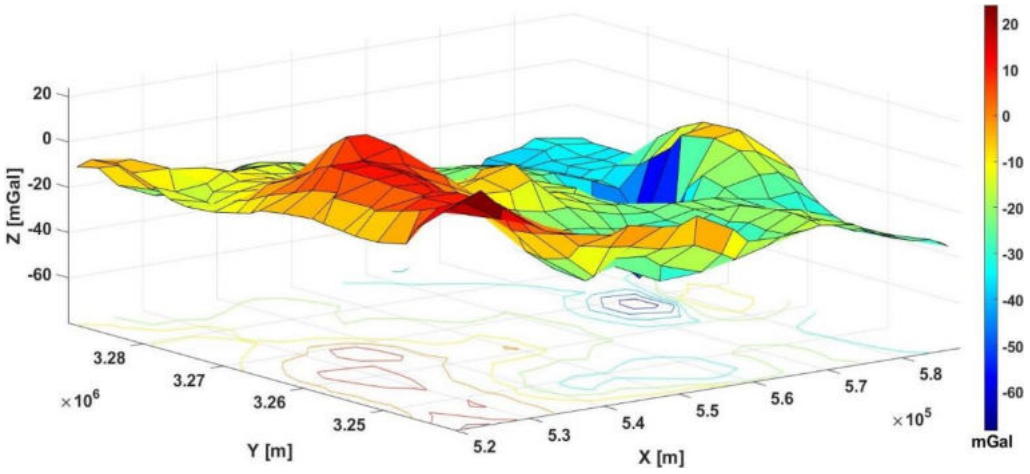

Figure 105. *3D Bouguer gravity anomaly surface of the datasets sampled non-equidistantly.*

The inversion-based low-pass filtering is subsequently executed on the non-equidistant measuring datasets shown above. *Figures 106* and *107* illustrate the filtered regional gravity anomaly map and its 3D anomaly surface view respectively. A close examination indicates that the results are identical in both anomaly shapes and amplitudes when compared to those found in the equidistant sampling case *(Figures 101* and *102* respectively). Most importantly, the 3D regional gravitational anomaly surface *(Figure 107)* is restored to its original form as one measure over an equidistant sampling grid. This reflects the inversion success and applicability when we deal with non-equidistant sampling field datasets, and it is therefore strongly recommended to investigate other geophysical tools' analyses of various datasets.

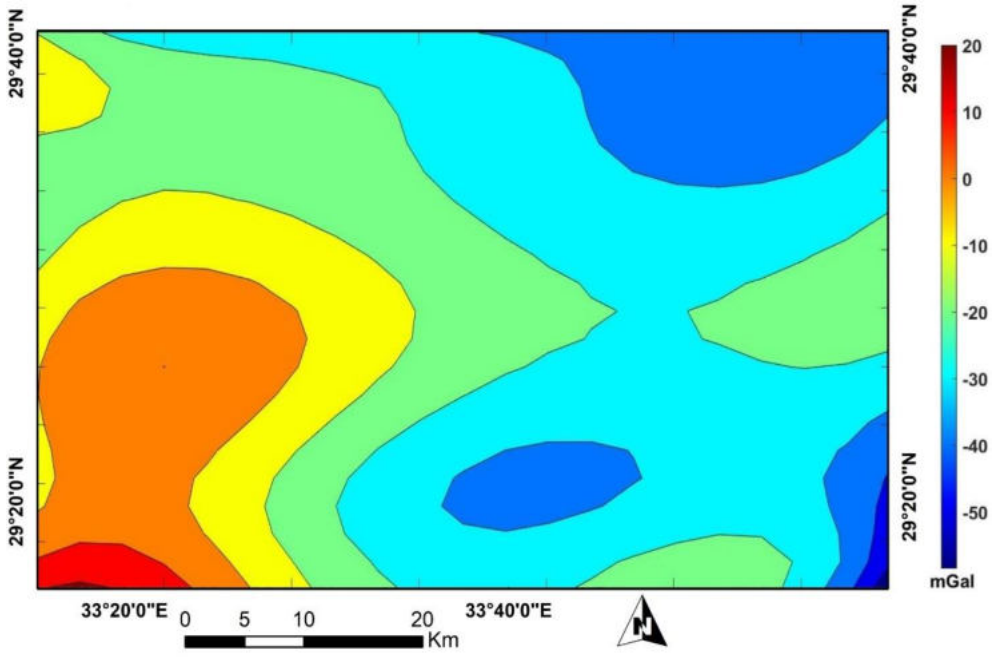

Figure 106. *Regional gravity anomaly map of the datasets sampled non-equidistantly.*

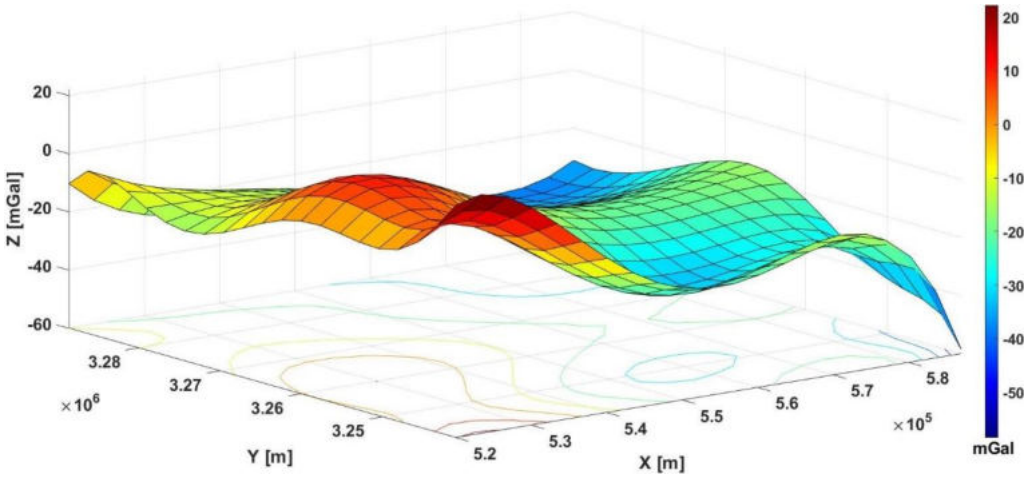

Figure 107. *3D Regional gravity anomaly surface of the datasets sampled non-equidistantly.*

6.3. 2D IRLS-FT BASED INCOMPLETE SAMPLING PROBLEM

Aside from the quality and accuracy of the data gathered during the field application of various geophysical methodologies, the total cost and financial aspects must be taken into consideration as well. Covering the whole area with datasets is sometimes expensive, time-consuming, and also labor-intensive, especially when the measurements are sampled over uniformly spaced grids. Furthermore, the natural or artificially created obstacles that encounter the geophysicists in the studied area are of great issue. Hence, the problem of the data missing or incomplete sampling can arise during the field survey. Consequently, our recently developed inversion-based Fourier transformation (2D IRLS-FT) can be regarded as one of the most crucial and economically advantageous concerns that have a great capability for resolving the difficulties stated. In our investigation, we apply our newly developed inversion algorithms to the above-mentioned equidistantly and non-equidistantly real gravity datasets measured in the western central part of Sinai Peninsula, Egypt to handle the missing data problem.

6.3.1. A FIELD EXAMPLE WITH EQUIDISTANTLY SAMPLED INCOMPLETE DATASETS

The same real gravity data taken equidistantly over 441 measuring points *(Figure 98)* are subjected to the IRLS-based inversion. In an attempt to validate the inversion's stability and the point at which it stops operating, we provide four examples in the context of low-pass filtering at different missing data percentages (25%, 35%, 50%, and 60%). In our analyses, Hermite functions of order M_x=M_y=13 are selected. In the first two cases when 25% and 35% of the data are randomly canceled through the inversion process, the total number of the measurements is reduced to be 327 and 281 respectively. The 3D regional gravity anomaly surfaces derived at

these missing data percentages are shown in *Figures 108a* and *b*. When compared to gravitational forms and amplitudes calculated by the inversion approach in both equidistant and non-equidistant sampling complete datasets *(Figures 102* and *107* respectively), it is found that these regional anomaly surfaces have the same properties. In the 50% missing data case, a total of 226 measurements are used to construct the 3D regional gravity anomaly surface as depicted in *Figure 108c*. The outcome is remarkably similar and satisfactory when compared to the equidistantly and non-equidistantly sampled regional anomaly surfaces based on the complete datasets. On the other hand, *Figure 108d* shows how the 3D regional gravity anomaly surface gets severely deforested and distorted in both anomaly shapes and amplitudes when the missing data percentage is increased up to 60%, using only 181 measuring points. This attributes to the same problem of the marginal over-determination ratio that exposed us to synthetic gravity and magnetic processing, where the number of measured datasets is nearly close to the model parameters (M_x M_y=169). In our investigation, measuring only half the study area's total data capacity (50%) is quite enough for the inversion introduced to accomplish our purpose of data processing as estimated by the completely over-determination value of 226/169=1.337.

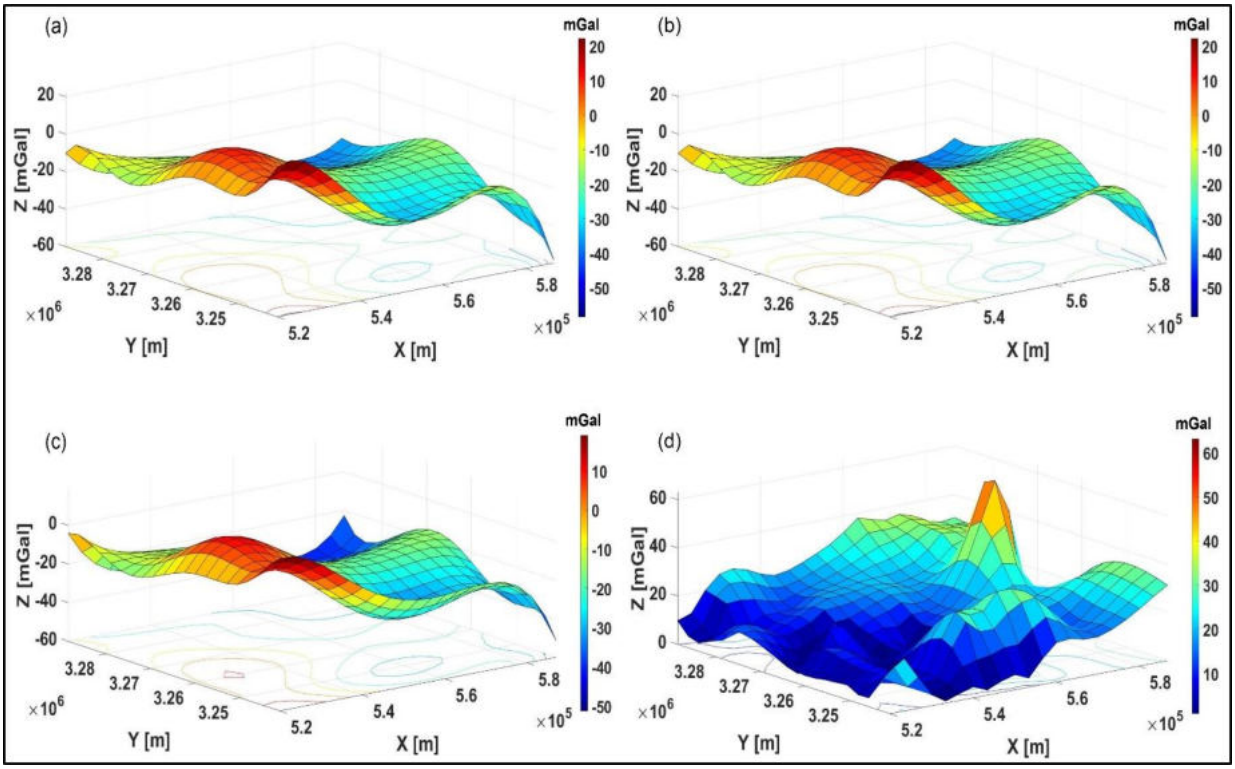

Figure 108. *The 3D regional anomaly surfaces using IRLS-FT when a) 25% of the equidistantly measured gravity field datasets are missing, b) 35% of the equidistantly measured gravity field datasets are missing, c) 50% of the equidistantly measured gravity field datasets are missing, d) 60% of the equidistantly measured gravity field datasets are missing.*

6.3.2. A FIELD EXAMPLE WITH NON-EQUIDISTANTLY SAMPLED INCOMPLETE DATASETS

To demonstrate the inversion effectiveness with the non-equidistantly sampled real gravity incomplete datasets, the algorithm is implemented on the non-regular sampling gravity datasets displayed in *Figure 103*. For comparison, the same missing data percentages are used when 25%, 35%, 50%, and 60% of the measurements are randomly cancelled. A total of 331, 288, 223, and 179 are randomly selected at each proportion respectively. The corresponding 3D regional gravity anomaly surfaces are shown in *Figures 109a, b, c,* and *d* using Hermite functions of order $M_x=M_y=13$.

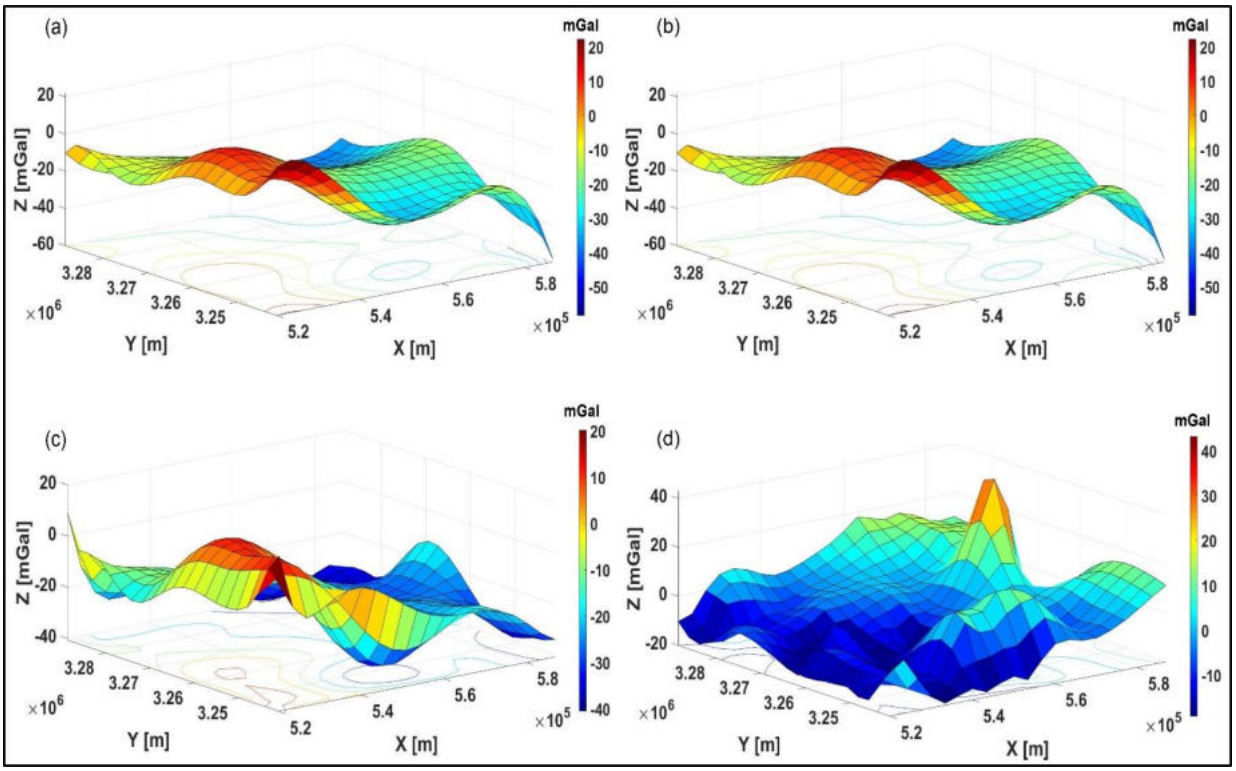

Figure 109. *The 3D regional anomaly surfaces using IRLS-FT when a) 25% of the non-equidistantly measured gravity field datasets are missing, b) 35% of the non-equidistantly measured gravity field datasets are missing, c) 50% of the non-equidistantly measured gravity field datasets are missing, d) 60% of the non-equidistantly measured gravity field datasets are missing.*

It is obvious that the regional surfaces obtained in 25% and 35% of missing cases are identical to those of the equidistantly and non-equidistantly sampled complete gravity data *(Figures 102* and *107* respectively). Moreover, there are slight variations in anomaly shapes and amplitudes while dealing with the 3D regional gravity anomaly surface of the 50% missing case but in general, it still provides us with satisfactory results. The inversion stops

operating when 60% of the measured data are randomly missing, in contrast to all the missing data percentages indicated previously due to the marginal over-determination where the breaking value of the completely over-determined inverse problem is approximately obtained at 50% missing case as 223/169=1.319. This is evident in the substantially deformed 3D regional surface. These results show that even when we measure only half of the datasets (50%) during the field survey, the inversion is resilient, efficient, and appropriate for processing the real incomplete datasets measured either equidistantly or non-equidistantly. Far beyond this proportion, the inverse problem is sounded to be slightly over-determined where the number of the observing data points used is nearly close to the number of the model parameters (M_x M_y=169). Thus, as we previously stated, our newly developed inversion-based Fourier transformation opens up new and economically viable possibilities in data processing, as well as in the planning of field observations assembled using various geophysical tools.

According to the results demonstrated above, I declare thesis statement 5 as follows:

Thesis 5

I assessed the efficiency and accuracy of the 2D IRLS-FT inversion method in outlier sensitivity and missing data problem using a field example of real gravity measurements sampled both equidistantly and non-equidistantly in the western central part of Sinai Peninsula, Egypt. A total of 441 measuring points were carried out over a grid of 21x21, whereas I set the unknown expansion coefficients as 13x13 to match the inversion procedure's accuracy and stability.

In the framework of the low-pass filtering, I found that the inversion method has a higher capacity for noise rejection and gave me similar regional gravity anomaly maps reflecting the inversion's efficiency and robustness for processing both the equidistantly and non-equidistantly measured field datasets. In the case of a missing data problem, I found that measuring only half of the study area's total data capacity (50%) is sufficient to present the inversion stability and accuracy, whether I treat equidistantly or non-equidistantly sampled field data measurements. The results are distorted at the missing percentages over 50%, where the over-determination rate decreases below 1.3.

Chapter 7
SUMMARY

The Fourier transformation is widely regarded as one of the most commonly utilized tools in geophysical data processing. In signal processing, several Fourier transformation algorithms methods are provided to enhance the quality of the geophysical datasets being interpreted, and hence, a comprehensive picture of the subsurface geology can be effectively drawn. It is well known that geophysical datasets collected in the field might occasionally encounter a slew of problems. This is owing to the presence of obstacles, either naturally or purposefully manufactured, in the surface area being investigated. In particular, advances in geophysical equipment as well as various surface topographical characteristics, like caverns, hills, and mountains, have a significant impact on data qualification. The constraints that necessitate the survey to be carried out on non-regular or random grids of measuring data points. In that case, it is quite likely that a fraction of noise outliers unrelated to subsurface geological anomalies will be included in the geophysical information gathered. Additionally, the random surveying of the areas under investigation causes gaps in missing datasets, and therefore, a significant amount of information is susceptible to lose. Due to the inefficiency and limitations of the traditional Fourier transformation (DFT) approach in processing the outlier noisy data, as well as the non-regular and incomplete sampling designs, this thesis introduces an algorithm of the inversion-based 1D and 2D Fourier transformation (IRLS-FT) to cope with these challenges. Dobróka et al (2015) used a one-dimensional (1D) inversion technique, which was further extended and adapted to two-dimensional (2D) inversion (Dobróka et al, 2017). Their research findings revealed sufficient improvements in both the space and frequency domains.

To give a complete analysis of the applicability of the IRLS-FT method, this thesis' objectives can be broken down into three distinct categories. The first is reducing the outlier sensitivity by applying the inversion-based Fourier transformation to the synthetic 1D wavelet and 2D magnetic and gravity datasets, as well as real gravity field measurements. Moreover, the second is to analyze the inversion approach in processing the non-regularly sampled complete datasets in 1D and 2D. For the two mentioned goals, the proposed IRLS-FT inversion method is based mainly on the iteratively reweighted least-squares Fourier transformation. The series expansion method is employed to discretize the Fourier frequency spectrum, with the expansion coefficients approximated as the solution to the over-determined inverse problem. The Hermite functions are constructed as basis functions, taking advantage of the fact that they

are eigenfunctions of the Fourier transformation, permitting quick and precise calculation of the elements of the Jacobian matrix. Besides, the most frequent value (MFV) method is used to handle the problem of the scale parameters by iteratively calculating the Cauchy-Steiner weights through an internal iteration loop with the least amount of data loss. The thesis' final goal, on the other hand, is to overcome the incomplete sampling problem at various degrees of missing data. This procedure is employed by reconstructing the inversion algorithms stated above and testing them on regularly and non-regularly sampled incomplete synthetic 1D wavelet and 2D magnetic and gravity datasets, as well as incomplete gravity field data.

First, the inversion-based Fourier transformation (IRLS-FT) method is tested on synthetic datasets sampled in one dimension. For noise rejection capacity, a comparison between our inversion approach and the conventional Fourier transformation (DFT) method is presented (Dobroka et al, 2012; Szegedi and Dobróka, 2014; Nuamah et al, 2021). In this case, the time-domain 1D equidistant synthetic Morlet wavelet is constructed over 401 measurement points. To evaluate the noise sensitivity, the aforementioned noise-free wavelet is subsequently contaminated with both Gaussian and random noise of Cauchy distribution, emulating the real filed data measurements. The real and imaginary components of the Fourier frequency spectrum are then computed using the conventional (DFT) and inversion-based IRLS-FT methods, which convert the noisy datasets from the time domain to the frequency domain. The results demonstrated that the inversion-based Fourier transformation (IRLS) approach is resilient and has a substantial ability to reduce outlier spread when compared to those found by the traditional DFT method, as numerically evidenced by the data and model distance values. However, the same data is randomized to show how effective the inversion is in the non-regular sampling case (Nuamah and Dobróka, 2019). Two different strategies are constructed, the datasets are first randomly selected and then generated in a framework of the random walk measurement positioning. The visualized and numerical findings in both sampling procedures provide extreme effectiveness, robustness, and applicability of the inversion method for processing 1D complete datasets gathered in non-regular configurations. Finally, we identify an interesting research point on the incomplete sampling problem that we'd like to tackle. Hence, a newly developed algorithm of the inversion-based 1D Fourier transformation method is applied to the 1D incomplete datasets sampled both equidistantly and non-equidistantly. In both cases, the measuring data points are randomly missed or cancelled achieving different missing data percentages (for example, 15%, 30%, and 50%). According to the outcomes, it can be demonstrated that the new inversion approach is highly appropriate for processing both

regularly and non-regularly collected incomplete datasets, even when half of the observations (50%) are missing. Further, the inversion is followed by the solution of a more sophisticated data geometry as defined by the block incomplete sampling design. To do so, different block intervals are specified, with their related data contents being erased or cancelled at random. It is shown that the inverted wavelet by our inversion method restored its original shape as if we were operating with complete datasets reflecting its higher applicability, and it may thus be suggested for geophysical field datasets gathered in incomplete block designs.

Due to the efficacy of the IRLS inversion approach in processing 1D datasets, the inversion algorithm is developed to be applied to equidistantly and non-equidistantly sampled 2D synthetic magnetic data. Our work is dominated by a practical application in the field of reduction to the Earth's magnetic pole (RTP) (Dobróka et al, 2017; Abdelaziz and Dobróka, 2020; Nuamah et al, 2021). The pole reduction (RTP) enables us to locate the anomaly directly above the centre of the causative body, the problem which arose from the dipolar nature of the Earth's geomagnetic field. The 2D synthetic magnetic data is generated above an FT-shaped magnetic anomaly (1681 measurements) using appropriate values of intensity, inclination, and declination. For measuring the outlier sensitivity, the datasets created along with an equidistant grid are contaminated with random noise of Cauchy distribution using a scale parameter of 0.03. The Cauchy noise-contaminated datasets are then subjected to both the traditional 2D DFT and the 2D IRLS inversion algorithms yielding in 2D Fourier frequency spectra. To achieve both inversion accuracy and stability, Hermite functions of order (M_x=M_y=25) are chosen as basis functions in all our calculations of the magnetic data processing. The results of the spectra as well as the inverted RTP anomaly maps demonstrate a superior noise reduction capability of the inversion used, which can be effectively employed as a viable alternative to the conventional DFT method. After randomizing the observed data points, the 2D inversion is deployed to non-regular complete datasets in a manner analogous to that used in the 1D sampling scenario (Nuamah and Dobróka, 2019). Random selection and random walk measurement positionings are also used in the evaluation. The results demonstrate the inversion algorithm's efficiency and stability in handling the non-regular sampling concerns. In addition, the newly developed 2D inversion algorithm is employed for treating the incomplete sampling problem of the 2D magnetic datasets. Hence, regularly, non-equidistantly, and randomly sampled magnetic datasets are all subjected to the inversion algorithm using different missing percentages. To present a comprehensive impression of inversion efficiency and acuity, missing 15%, 25%, 35%, 45%, 50%, and 60% of the data measurements are selected in all mentioned

sampling configurations. At the same missing data percentages stated above, the 2D noisy magnetic datasets carried out over both regular and random sampling grids are then evaluated. According to the visualized and numerical results based on both noise-free and noisy magnetic datasets, the newly developed inversion-based 2D Fourier transformation is highly applicable and gives adequate improvements up to 50% of missing data as a maximum, far beyond this point, the inversion is ineffective. That is because our inverse problem is sounded to be marginally over-determined where the observing data points used are nearly close to the number of the model parameters (M_x M_y).

The applicability of our inversion-based Fourier transformation method is extended to the 2D synthetic gravitational datasets for evaluating its capacity in noise reduction. Unlike the magnetic data processing described above, we present a practical geophysical deployment based on low-pass filtering, in which deep gravitational sources with low wavenumber (regional) are enhanced, while those of higher wavenumber related to the local shallow sources (residuals) are ignored or rejected. This filtering procedure results in regional gravity anomaly maps. To accomplish these goals, a model of a right rectangular prism using formulas given by (Nagy, 1966; Blakely 1996) is constructed to generate 2D synthetic gravity datasets (1681 measurements). The model is initially created over an equidistant grid before being randomized to generate a non-equidistant sampling design. Furthermore, the synthetic datasets are then contaminated with random noise of Cauchy distribution to simulate real data. For signal processing, the conventional Fourier transformation (2D DFT) approach is only performed on the gravitational datasets sampled equidistantly, whereas the 2D IRLS-FT inversion methodology is applied to data dispersed over both equidistant and non-equidistant grids. In both, the inversion procedures are executed using Hermite functions of order M_x=M_y=25 as an inversion parameter for a compromise between accuracy and stability. In the equidistant sampling case, the inversion algorithm gives a much better resolution and has a higher noise rejection capability when compared to the results found by the traditional DFT method, especially when the Cauchy noise is dominating. This appears clearly in the noisy regional gravity anomaly maps estimated by IRLS-FT which is the same as that obtained by the DFT method with the noise-free data. The improvements are further verified by mathematical computations of data and model distance values in the space and frequency domains respectively. Similarly, the 2D inversion-based Fourier transformation is applied to the non-equidistantly sampled noisy and noise-free datasets. According to the findings represented by the data filtering enhancements as well as the numerical calculations in space and frequency

domains, the inversion seems to reset the data to its original shapes and amplitudes as one would measure over an equidistant grid. It reflects our inversion approach's efficiency, stability, and adaptability in processing gravitational data sampled non-equidistantly, making it highly recommended to treat diverse data taken via other geophysical tools. Finally, the incomplete sampling problem is assessed using our newly developed IRLS-based inversion on noise-free and noisy 2D gravity datasets sampled at the same missing data percentages used for magnetic data processing (15%, 25%, 35%, 45%, 50%, and 60%). In both equidistant and non-equidistant sampling cases, stability and effectiveness of the inversion used are evidenced for all missing data below 50 percent (the percentage at which the inverse problem is completely over-determined) which are distorted after that point due to the same problem of marginal over-determination ratio.

Finally, the assessment of the above-mentioned inversion-based Fourier transformation method (2D IRLS-FT) is completed with a field example of real gravity measurements acquired using CG-3 Autograv (Automated Gravity Meter) manufactured by Scintrex (CG-3 Autograv Manual, 1995). Hence, a case study within the western central part of Sinai Peninsula, Egypt is involved. The study area is of great importance where the Egyptian government is developing a strategy in the last few decades to encourage tourism, agriculture, and industrial enterprises as a part of the national Sinai development plan (NSDP). The geological setting of the study area is characterized by the presence of several rock units belonging to Lower and Upper Cretaceous, Paleocene, Lower Eocene, and Quaternary deposits. In addition, to acquire a comprehensive image of the area under investigation, other maps including those showing land use/cover, lineaments density, and slope are supplied. A 3D magnetic modeling is constructed showing the basement relief surface of the study area, which varies by roughly 200 meters from west to east. For gravity data processing, the inversion-based low-pass filtering is implemented on the field data (441 measurements) taken over an equidistant grid which is subsequently randomized to construct a non-equidistant sampling design. In both cases, we provide the Bouguer gravity anomaly contour maps and related 3D visualization. Our inversion then uses Hermite functions of order $M_x=M_y=13$ to estimate the regional gravity anomaly maps and surfaces. The results showed the same properties reflecting the inversion's efficacy and robustness for handling both the equidistantly and non-equidistantly measured field datasets. Furthermore, the two configurations of field data indicated above are examined for the incomplete or missing sampling problem utilizing our newly developed inversion method. Hence, missing 25%, 35%, 50%, and 60% of the data measurements are selected in both

sampling geometries to compare the corresponding results and also to accomplish the inversion stability by identifying the missing percentage at which the inversion is insufficiently effective or ceases to operate. The 3D low-pass filtered regional gravity anomaly surfaces showed that regardless of whether we treat equidistantly or non-equidistantly sampled field data measurements, measuring only half of the study area's total data capacity (50%) is sufficient to present our newly developed inversion stability and accuracy. Dealing with our real datasets, it is established that as we increase the missing data percentages after 50%, the obtained results are distorted where the number of the field data measuring points is nearly close to the number of the model parameters (M_x M_y), and hence the inverse problem is slightly overdetermined. According to the findings of 1D and 2D synthetic datasets in space and frequency domains, as well as the real gravitational field data measurements, the newly developed inversion-based Fourier transformation approach is highly applicable, robust, and noise resistant, opening up a new and economically powerful way of handling the incomplete sampling problems encountered during wide and diverse field data acquisition.

Acknowledgment

First and foremost, I would like to express my sincerest gratitude and deepest thanks to ALLAH, who helped and guided me in bringing this thesis to light. Praise is to God.

I would like to express my sincere gratitude to my supervisor Professor. Dr. Mihály Dobróka, professor emeritus at the Department of Geophysics, University of Miskolc, Hungary for his continuous support, motivation, valuable scientific discussions, and providing all assistance and facilities as this research is being done. I could not have imagined having a better supervisor for my Ph.D. research either on the professional or human levels. Without his advice, immense experience, and moral support, the success of this thesis cloud have not been possible.

My sincere gratitude also goes to my co-supervisor Dr. Endre Turai, a private professor at the Department of Geophysics, University of Miskolc, Hungary for his kind supervision, continuous help, encouragement, and imparting his knowledge and experience during the progress of my academic research.

I would like to express my warmest thanks to all the faculty members of the Department of Geophysics, University of Miskolc, Hungary, especially Professor. Dr. Norbert Péter Szabó for the kind cooperation and appreciation. My sincere gratitude also extends to the Egyptian cultural affairs and missions sector for their continuous support. I would like to thank my PhD colleagues for the lovely moments and moral support I received during my academic stay. Special appreciation to Dr. Armand Abordan for his cooperation and personal assistance in my daily life. I wish to express my thanks to the reviewers for letting my defense be an enjoyable moment and for their brilliant comments and suggestions.

I would like to express my heartfelt thanks to my lovely wife (Dr. Abrar Samir Awad) and cute son (Ziad). Their tremendous understanding and encouragement in the past few years were the motivating force behind the achievement of my doctoral dissertation.

Last but not the least, I would like to dedicate all this work to the soul of my mother, my father, and my lovely baby (Rakan). Words cannot actually express how I feel about your loss. I will never forget my parents' continuous encouragement and moral support in their life to push me forward. I hope paradise will be your dwell.

Mahmoud Ibrahim Abdelaziz

REFERENCES

Abdelaziz, M. I. (2019). Application of Magnetic and Geoelectrical Resistivity Sounding Methods in Groundwater Exploration in a Part of Central Sinai, Egypt. MultiScience – XXXIII. microCAD International Multidisciplinary Scientific Conference, Hungary, 9p. DIO:10.26649/music.2019.079. 23-24

Abdelaziz, M. I., Dobróka, M. (2020). Testing the Noise Rejection Capability of the Inversion Based Fourier Transformation Algorithm Applied to 2D Synthetic Geomagnetic Datasets. Journal of Geoscience and Engineering, 8 (13), pp. 72-80.

Aki, K., Richards, P. G. (1980). Quantitative Seismology, Theory and Methods. Vol. I, 557p and vol II, 373p Freeman and Co, San Francisco, Geological Magazine, 118(2), 208-208. DOI: https://doi.org/10.1017/S0016756800034439

Al Hinai, N. (2020). Biomedical Signal Processing and Artificial Intelligence in Healthcare. Elsevier Inc, ISBN 978-0-12-818946-7, Chapter 1.4.2 Frequency domain filters, pp. 1-28.

Ali, O. F. (2006). Hydrogeological studies on Lower Cretaceous rocks (Nubia sandstone aquifer) in Central Sinai, Egypt. Ph.D. Thesis, Fac. Sci., Azhar University, Egypt, 282 p.

Amundsen, L. (1991). Comparison of the least-squares criterion and the Cauchy criterion in frequency-wavenumber inversion. Geophysics 56, pp. 2027-2035.

Araffa, S. A. S. (2013). Delineation of groundwater aquifer and subsurface structures on North Cairo, Egypt, using integrated interpretation of magnetic, gravity, geoelectrical and geochemical data. Geophysical Journal International, 192 (1), pp. 94–112. DOI: https://doi.org/10.1093/gji/ggs008

Araffa, S. A. S., Sabet, H. S., Gaweish, W. R. (2015). Integrated geophysical interpretation for delineating the structural elements and groundwater aquifers at central part of Sinai Peninsula , Egypt. Journal of African Earth Sciences, 105 (2015), pp. 93 – 106. DOI. https://doi.org/10.1016/j.jafrearsci.2015.02.011

Araffa, S. A. S., Sabet, H. S., Mahmoud, M. H. (2019). Tectonics mapping using geophysical data around Nuweiba area, Gulf of Aqaba, Sinai, Egypt. NRIAG Journal of Astronomy and Geophysics, 8 (1), pp. 73-83. DOI: https://doi.org/10.1080/20909977.2019.1615802

Ateya, I. L., Takemoto, S. (2002). Gravity inversion modeling across a 2-D dike-like structure-A case study. Earth, Planets and Space, 54, pp. 791-796. DOI: https://doi.org/10.1186/BF03351732

Backus, G. E. (1970). Inference from inadequate and inaccurate data, I. Proc. Natl. Acad. Sci. USA, 65, pp.1 –7. DOI: 10.1073/pnas.65.1.1

Backus, G. E., Gilbert, F. (1967). Numerical applications of a formalism for geophysical inverse problems. Geophys. J. R. Astr. Soc, 13,pp. 247 –276.

Baranov, V., Naudy, H. (1964). Numerical calculation of the formula of reduction to the magnetic pole. Geophysics., 29, pp. 67–79. DOI: https://doi.org/10.1190/1.1439334

Bevan, B. W. (2010). Geophysical Data Processing, an Introduction. Technical report, 62 p. DOI: 10.13140/RG.2.2.35060.96647

Blakely, R. J. (1996). Potential Theory in Gravity and Magnetic Applications, Cambridge, Cambridge University Press. DOI: https://doi.org/10.1017/CBO9780511549816

Budden, N. A. (1999). Mars field geology, biology and paleontology workshop: Summary and recommendations: Houston, Lunar and Planetary Institute, Contribution 968, 80 p.

Cady, J. W. (1980). Calculation of gravity and magnetic anomalies of finite-length right polygonal prisms. Geophysics, 45 (10), pp. 1507-1512. DOI: https://doi.org/10.1190/1.1441045

CG-3 Autograv Manual. (1995). SCINTREX, Automated gravity meter, Operator manual, version 5, 198p .

Chai, X., Gu, H., Li, F., Duan, H., Hu, X., Lin, K. (2020). Deep learning for irregularly and regularly missing data reconstruction. Sci Rep 10, 3302. DOI: https://doi.org/10.1038/s41598-020-59801-x

Claerbout, J. F. (1985). Fundamentals of Geophysical Data Processing. United States, 174 p.

Claerbout, J. F., Muir, F. (1973). Robust modeling with erratic data. Geophysics, 38, pp. 826 – 844.

Dimri, V. P., Srivastava, R. P. (2007). Design of Gravity Survey Network and Its Application: a Case Study. Innovation in EM, Grav and Mag Methods:a new Perspective for Exploration Capri, Italy, April 15 – 18, 3p. DOI: https://doi.org/10.3997/2214-4609-pdb.166.C_PP_01

Dobrin , M . B ., 1976. Introduction to geophysical prospecting . McGraw-Hill Book Co., New York , USA , 630 p.

Dobrin, M. B. (1981). Introduction to geophysical prospecting , third edition , McGRAW-HILL INTERNATIONAL BOOK COMPANY, 630 p.

Dobroka, M., Volgyesi, L. (2008). Inversion reconstruction of gravity potential based on gravity gradients. Math Geosci, 40(3), pp. 299–311. DOI: https:// doi. org/ 10. 1007/ s11004- 007-9139-z

Dobróka, M., Gyulai, A., Ormos, T., Csókás. J., Dresen, L. (1991). Joint inversion of seismic and geoelectric data recorded in an under-ground coal mine. Geophys Prospect, 39, pp. 643–665. DOI: https://doi.org/10.1111/j.1365-2478.1991.tb00334.x

Dobroka, M., Pracser, E., Kavanda, R., Turai, E. (2013). Quick imaging of MT data using an approximate inversion algorithm. Acta Geod Geophys J, 48(1), pp. 17–25. DOI: https:// doi. org/ 10. 1007/ s40328- 012- 0011-3

Dobróka, M., Szabo, N. P. (2010). Sorfejteses inverzio II. Melyfurasi geofizikai adatok feldolgozasa intervallum inverzios eljarassal. Magyar Geofizika, 51(1), pp. 25–42.

Dobróka, M., Szabo, N. P., Toth, J., Vass, P. (2016). Interval inversion approach for an improved interpretation of well logs. Geophysics, 81(2), pp. D155–D167. DOI: https://doi.org/10.1190/geo2015-0422.1

Dobróka, M., Szegedi, H., Vass P. (2017). Inversion-based Fourier transform as a new tool for noise rejection, INTECH 2017. DOI: https://doi.org/10.5772/66338

Dobróka, M., Szegedi, H., Vass, P., Turai, E. (2012): Fourier transformation as inverse problem - an improved algorithm. Acta Geodaetica et Geophysica Hungarica, Vol. 47(2), pp. 185- 196. DOI: 10.1556/AGeod.47.2012.2.7

Dobróka, M., Szegedi, H., Somogyi Molnár, J. Szűcs, P. (2015). On the Reduced Noise Sensitivity of a New Fourier Transformation Algorithm. Mathematical Geosciences., 47 (6), pp. 679-697. DOI: https://doi.org/10.1007/s11004-014-9570-x

Dobroka, M., Volgyesi, L. (2010). Sorfejteses inverzio IV. A nehezsegi erőter potencialfuggvenyenek inverzios előallitasa. Magyar Geofizika, 51(3), pp. 143−149.

Ellis, G. (2012). Control System Design Guide. Elsevier Inc, ISBN 978-0-12-385920-4, Chapter 9.2.1.5 Butterworth Low-Pass Filters, pp. 165-183.

Essa, K. S. (2014). New fast least-squares algorithm for estimating the best-fitting parameters due to simple geometric-structures from gravity anomalies. Journal of Advanced Research, 5 (1), pp. 57-65. DOI: https://doi.org/10.1016/j.jare.2012.11.006

Fancsik, T., Turai, E., Szabó, N. P., Molnár, J. S., Dobróka, T. E., Dobróka, M. (2021). Evaluation of induced polarization measurements using a new inversion method. Acta Geod Geophys J, 56, pp. 623–643. DOI: https://doi.org/10.1007/s40328-021-00355-3

Fedi, M. (2007). DEXP: A fast method to determine the depth and the structural index of potential fields sources. Geophysics, 72 (1), pp. I1-I11. DOI: https://doi.org/10.1190/1.2399452

Ferenczy, L., Kormos, L., Szucs, P. (1990). A new statistical method in well log interpretation, paper O. In: 13th European formation evaluation symposium transactions: Soc. Prof. Well Log Analysts, Budapest Chapter, 17p.

Fuchs, K., Müller, G. (1971). Computation of synthetic seismograms with relectivity method and comparison with observations. Geophys. J. Roy. Astr. Soc . 23, pp. 417 −433.

Gauss, C. F. (1809). Theory of motion of the heavenly bodies moving about the sun in conic sections. Hamburg, Perthes et Besser.

Geosoft Inc. (2014). GM-SYS 3D Modeling, Montaj extension developed by Geosoft. In, Geosoft oasis montaj manual, pp . 1 – 4.

Gill, M. K., Asefa, T., Kaheil, Y., McKee, M. (2007). Effect of missing data on performance of learning algorithms forhydrologic predictions: Implications to an imputation technique. Water resources research, 43, W07416. DOI:10.1029/2006WR005298

Gyulai, A., Ormos, T., Dobroka, M. (2010). Sorfejteses inverzio V. Gyors 2D geoelektromos inverzios modszer. Magyar Geofizika, 51(4), pp. 185−195.

Gyulai, A., Szűcs, P., Turai, E., Baracza, M. K., Fejes, Z. (2017). Geoelectric characterization of thermal water aquifers using 2.5D inversion of VES measurements. Surv Geophys, 38(2), pp. 503 526. DOI: https:// doi. org/ 10. 1007/ s10712- 016- 9393-z

Hassanin, A. M. (1997). Geological and geomorphological impacts on the water resources in Central Sinai, Egypt. Ph. D. Thesis, Geol. Dept., Fac. of Sci., Ain-Shams Univ, ARE, 373 p.

Hermite, C. (1864). Sur un nouveau développement en série de fonctions [On a new development in function series]. C. R. Acad. Sci. Paris. 58: 93–100. Collected in Œuvres II, pp. 293–303. DOI: https://doi.org/10.1017/CBO9780511702761.022

Hinze, W. J., Von Frese, R. R. B., Saad, A. H. (2013). Gravity and Magnetic Exploration: Principles, Practices and Applications. New York, Cambridge University Press, 512 p.

Hochsten, M. P., Hunt, T, M. (1970). Seismic, Gravity and Magnetic Studies, Broadlands Geothermal Field, New Zealand. Geotherraics, 2 (1), pp. 333-346. DOI: https://doi.org/10.1016/0375-6505(70)90032-5

Hoffman, S. J., Kaplan, D., editors. (1997). The reference mission of the NASA Mars Exploration Study Team: NASA Special Publication 6107.

Hubbard, S. S., Linde, N. (2011). Treatise on Water Science. Earth Systems and Environmental Sciences, Vol 2, pp. 401-434. DOI: https://doi.org/10.1016/B978-0-444-53199-5.00043-9

Hume, W. F. (1906). The Topography and Geology of the Peninsula of Sinai. Geological Survey, Cairo.

ICG (International Crisis Group). (2007). Egypt's Sinai question. Middle East / North Africa Report no, 61, ICG, Brussels, 32p.

Jackson, D. D. (1979). The use of a priori data to resolve non-uniqueness in linear inversion. Geophys. J. R. Astr. Soc, 57, pp. 137 –157. DOI: https://doi.org/10.1111/j.1365-246X.1979.tb03777.x

Karcol, R., Pašteka, R. (2019). On the Two Different Formulas for the 3D Rectangular Prism Effect in Gravimetry. Pure and applied geophysics. 176, pp.257–263. DOI: https://doi.org/10.1007/s00024-018-1966-y

Kaur, P., Garg, D. K. (2020). Construction of incomplete Sudoku square and partially balanced incomplete block designs. Communications in Statistics-Theory and Methods, 49 (6), pp. 1462-1474. DOI: https://doi.org/10.1080/03610926.2018.1563177

Kazufumi, I., Bangti, J. (2014). Inverse Problems : Tikhonov Theory and Algorithms (Series on Applied Mathematics). World Scientific Publishing Company. DOI: https://doi.org/10.1142/9120

Kearey, P., Brooks, M., Hill, I. (2002). An Introduction to Geophysical Exploration. 3rd Edition, Blackwell Science Ltd, 288 p.

Kondrashov, D., Ghil, M. (2006). Spatio-temporal filling of missing points in geophysical data sets. Nonlin. Processes Geophys, 13, pp. 151–159. DOI: https://doi.org/10.5194/npg-13-151-2006

Kunaratnam, K. (1981). Simplified expressions for the magnetic anomalies due to vertical rectangular prisms. Geophysical Prospecting, 29, pp. 883-890. DOI:https://doi.org/10.1111/j.1365-2478.1981.tb01032.x

Lafehr, T. R., Nabighian, M. N. (2012). Fundamentals of Gravity Exploration. Tulsa, OK: Society of Exploration Geophysicists, 211 p.

Laplace, P. S. (1812). Théorie analytique des probabilités [Analytic Probability Theory]., Ibidem, t. 7, No. 1–2, Paris, 1886. DOI: https://doi.org/10.1016/B978-044450871-3/50105-4

Legendre, A. M. (1805). New Methods for the Determination of the Orbits of Comets, Paris. F. Didot., 13, 80 p.

Luan Thanh Pham., Duong Anh Nguyen., Eldosouky., A. M., Abdelrahman, K., Tich Van Vu., Al-Otaibi, N., Ibrahim, E., Kharbish, S. (2021). Subsurface structural mapping from high-resolution gravity data using advanced processing methods. Journal of King Saud University – Science, 33 (5) 1014882, pp. 1-8. DOI: https://doi.org/10.1016/j.jksus.2021.101488

Luo Yao., Dian-Jun XUE., Ming Wang. (2010). Reduction to the pole at the geomagnetic equator. Chinese Journal of Geophysics, 53 (6), pp.1082-1089. DOI: https://doi.org/10.1002/cjg2.1578

Ma, J. (2013). Three-dimensional irregular seismic data reconstruction via low-rank matrix completion," Geophysics, 78 (5), pp. V181-V192. DOI: https://doi.org/10.1190/geo2012-0465.1

Meju, M. A. (1994). Geophysical Data Analysis: Understanding inverse problem theory and practice. SEG Course Notes Series, Vol. 6, Society of Exploration Geophysicists, Tulsa, 296p.

Mekhemer, H. M., Sultan, S. A., Abd Alla, M. A., Brimich, L., Santos, F. M. (2007). Integrated geophysical interpretation for groundwater potentiality at Wadi Ghubba, Central Sinai, Egypt. Contributions to Geophysics and Geodesy, 37 (4), pp. 387-402.

Menke, W. (1984). Geophysical Data Analysis: Discrete Inverse Theory. Academic Press, New York.

Menke, W. (1989). Geophysical Data Analysis; Discrete Inverse Theory. Revised Edition: Academic Press, Inc., Harcourt Brace Jovanovich Publishers, San Diego.

Menke, W. (2018). Geophysical Data Analysis; Discrete Inverse Theory. Fourth Edition, Academic Press, 322p. DOI: https://doi.org/10.1016/C2016-0-05203-8

Meyer, C., Treiman, A. H., and Kostiuk, T. (1995). Planetary surface instruments workshop: Houston, Lunar and Planetary Institute, LPI Technical Report 95-05.

Milsom, J. (2003). Field Geophysics. 3rd Edition; John Wiley & Sons Ltd, University College London, England, 232 p.

Montsion, R. M., Perrouty, S., Lindsay, M. D., Jessell, M. W., Frieman, B. M. (2021). Mapping structural complexity using geophysics: A new geostatistical approach applied to greenstone belts of the southern Superior Province, Canada. Tectonophysics, 812 (5), 19 p. DOI: https://doi.org/10.1016/j.tecto.2021.228889

Murty, B. V. S., Raghavan, V. K. (2002). The gravity method in groundwater exploration in crystalline rocks: a study in the peninsular granitic region of Hyderabad, India. Hydrogeol. J, 10, pp. 307 – 321. DOI: https://doi.org/10.1007/s10040-001-0184-2

Nagy, D. (1966). The gravitational attraction of a right rectangular prism. Geophysics, 31 (2), pp. 362 – 371. DOI: https://doi.org/10.1190/1.1439779

Nettleton, L. L (1976). Gravity and Magnetics in Oil Prospecting. New York: McGraw-Hill Book Co.

Newton, Sir Isaac. (1687). The mathematical principles of natural philosophy. London.

Nuamah, D. O. B., Dobróka, M. (2019). Inversion-based Fourier transformation used in processing non-equidistantly measured magnetic data. Acta Geodaetica et Geophysica, 54, pp. 411–424. https://doi.org/10.1007/s40328-019-00266-4

Nuamah, D. O. B., Dobróka, M., Vass, P., Baracza, M. K. (2021). Legendre polynomial-based robust Fourier transformation and its use in reduction to the pole of magnetic data. Acta Geodaetica et Geophysica,56, pp. 645–666.DOI: https://doi.org/10.1007/s40328-021-00357-1

Oldenburg, D. W., Yaoguo Li. (2005). Inversion for Applied Geophysics: A Tutorial. Near Surface Geophysics, Society of Exploration Geophysicists, 85 p. DOI: https://doi.org/10.1190/1.9781560801719.ch5

Parker, R. L. (1977). Understanding inverse theory. Ann. Rev, Earth Planet. Sci, 5, pp. 35 –64. https://doi.org/10.1146/annurev.ea.05.050177.000343

Plouff, D. (1976). Gravity and magnetic fields of polygonal prisms and application to magnetic terrain corrections. Geophysics, 41(4), pp. 727–741. DOI: https://doi.org/10.1190/1.1440645

Rabeh, T., Miranda, M. (2008). A tectonic model of the Sinai Peninsula based on magnetic data, J of Geophysics and Engineering, 5 (4), pp. 469–479. DOI: https://doi.org/10.1088/1742-2132/5/4/010

Rasmussen, R., Pedersen, L. B. (1979). End corrections in potential field modeling. Geophysical Prospecting, 27 (4), pp. 749-760. DOI: https://doi.org/10.1111/j.1365-2478.1979.tb00994.x

Reynolds, J. M. (1997). An Introduction to Applied and Environmental Geophysics, PDF, West Sussex, PO19 IUD, England , 796 P.

Roy, A. (1966). The method of continuation in mining geophysical interpretation. Geoexploration, 4 (2), pp. 65-83. DOI: https://doi.org/10.1016/0016-7142(66)90012-3

Sacchi, M. D., Ulrych, T. J., Walker, C. J. (1998). Interpolation and extrapolation using a high-resolution discrete Fourier transform. IEEE Transactions on Signal Processing, 46 (1), pp. 31–38. DOI: 10.1109/78.651165.

Said, R. (1962). The geology of Egypt. Elseveir Publishing Company, Amsterdam and New York.

Said, R. (1990). The geology of Egypt. Egyptian General Petroleum Corporation, Conoco Hurghada Inc. and Repsal Exploration, S. A, 734 p.

Salat, P., Tarcsai, Gy., Cserepes, L., Vermes, M., Drahos, D. (1982). Information-statistical methods of geophysical interpretation (in Hungarian), Tankonyvkiadó.

Sangpara, P; Borkowski, J. J; Srisuradetchai, P. (2021). Using Balanced Incomplete Block Designs to Generate New Sampling Designs. Thailand Statistician, 19 (1), pp. 125-139.

Scales, J. A., Gersztenkorn, A., Treitel, S. (1988). Fast Lp solution of large, sparse, linear systems: Application to seismic travel time tomography. Journal of Computational Physics 75, pp. 314- 333.

See, K., Song, S. Y., Stufken , J. (1997). On a class of partially balanced incomplete block designs with applications in survey sampling. Communications in Statistics-Theory and Methods, 26 (1), pp. 1-13. DOI: https://doi.org/10.1080/03610929708831898

Sen, M. K., Stoffa, P. L. (2013). Global optimization methods in geophysical inversion. 2nd eddition, Cambridge University Press, New York. DOI: https://doi.org/10.1017/CBO9780511997570

Sharma, P. V (1997). Environmental and engineering geophysics , Cambridge University Press , 475 p.

Shata, A. (1956). Structural Development of the Sinai Peninsula, Egypt. Bull. Inst. Desert, Egypt, 6 (2), pp. 117-157.

Shebl, A., Abdellatif, M., Hissen, M., Abdelaziz, M. I., Csámer, Á. (2021). Lithological mapping enhancement by integrating Sentinel 2 and gamma-ray data utilizing support vector machine: A case study from Egypt. Int. J. Appl. Earth Obs. Geoinf, 105, 102619, 16 p. DOI: https://doi.org/10.1016/j.jag.2021.102619

Shebl, A., Csámer, Á. (2021). Reappraisal of DEMs, Radar and optical datasets in lineaments extraction with emphasis on the spatial context. Remote Sens. Appl. Soc. Environ. 24, 100617. https://doi.org/10.1016/J.RSASE.2021.100617

Shebl, A., Kusky, T., Csámer, Á. (2022). Advanced land imager superiority in lithological classification utilizing machine learning algorithms. Arab. J. Geosci. 2022, 15, 923, pp. 1–13. https://doi.org/10.1007/S12517-022-09948-W

Shuey, R. T., Pasquale, A. S. (1973). End corrections in magnetic profile interpretation. Geophysics, 38 (3), pp. 507-512. DOI: https://doi.org/10.1190/1.1440356

Singh, S. C., West, G. F., Bregman, N. D., Chapman , C. H. (1989). Full waveform inversion of relection data. J. Geophys. Res, 94, pp. 1777 −1794.

Steiner, F. (1988). Most frequent value procedures (a short monograph). Geophysical Transactions,34, pp.139-260.

Steiner, F. (1997). Optimum Methods in Statistics. Budapest: Academic Press.

Steiner, F., Hajagos, B. (1994). Practical definition of robustness. Geophys Trans, 38, pp. 193–210.

Studt, F. E. (1964). Geophysical prospecting in New Zealand's hydrothermal fields. U.N. Con/. New Sources Energy, Rome, 2, 380 p.

Szabó, N. P. (2015). Hydraulic conductivity explored by factor analysis of borehole geophysical data. Hydrogeol J, 23(5), pp. 869–882. DOI: https ://doi.org/10.1007/s1004 0-015-1235-4

Szabó, N. P. (2011). Shale volume estimation based on the factor analysis of well-logging data. Acta Geophys, 59 (5), pp. 935–953. DOI: https ://doi.org/10.2478/s1160 0-011-0034-0

Szegedi, H., Dobróka, M. (2014). On the use of Steiner's weights in inversion-based Fourier transformation: robustification of a previously published algorithm. Acta Geophysica, 49 (1), pp. 95-104. DOI: 10.1007/s40328-014-0041-0

Szucs, P., Civan, F. (1996). Multi-layer well log interpretation using the simulated annealing method. J Pet Sci Eng, 14, pp. 209–220.

Szucs, P., Civan, F.,Virag, M. (2006). Applicability of the most frequent value method in groundwater modeling. Hydrogeol J, 14, pp. 31–43. DOI: https://doi.org/10.1007/s10040-004-0426-1

Tarantola, A. (1987). Inverse Problem Theory, Methods of Data Fitting and Model Parameter Estimation. Elsevier Publishing Company, New York .

Thompson, M. T. (2013). Intuitive Analog Circuit Design. Elsevier Inc, ISBN 978-0-12-405866-8, Chapter 14 Analog Low-Pass Filters, pp. 531-583.

Trad, D. (2009). Five-dimensional interpolation: Recovering from acquisition constraints. Geophysics 74 (6), pp. V123-V132. DOI: https://doi.org/10.1190/1.3245216

Turai, E., Dobróka, M. (2011). Data processing method developments using TAU transformation of time-domain IP data I. Theoretical basis. Acta Geodaetica et Geophysica Hungarica, 46 (3), pp. 283-290. DOI: https://doi.org/10.1556/AGeod.46.2011.3.1

Turai, E., Dobroka, M., Herczeg, A. (2010). Sorfejteses inverzio III. Gerjesztett polarizacios adatok inverzios feldolgozasa. Magyar Geofizika, 51(2), pp. 88−98.

Yagle, A. E., Levy, B. C. (1985). A layer-stripping solution of the inverse problem for a one-dimensional elastic medium. Geophysics J, 50, pp. 425 −433.

Zhang, J. (2017). Most frequent value statistics and distribution of 7Li abundance observations, Monthly Notices of the Royal Astronomical Society, 468 (4), pp. 5014–5019, DOI: https://doi.org/10.1093/mnras/stx627