DEVELOPMENT OF UNIFIED INTEGRATED MODEL TO CHARACTERIZE GEOPHYSICAL DATA USING IMAGE PROCESSING TECHNIQUE

By

ISHOLA KEHINDE SAHEED

April 2015

DEVELOPMENT OF UNIFIED INTEGRATED MODEL TO CHARACTERIZE GEOPHYSICAL DATA USING IMAGE PROCESSING TECHNIQUE

By

ISHOLA KEHINDE SAHEED

Thesis submitted in fulfillment of the Requirement for the degree of Doctor of Philosophy

MARCH, 2015

DEDICATION

This work is dedicated in memories of my parents Pa Abdulganiu and Bintu Fatimah Ishola Kabala Alaiyeogun and my lovely eldest brother, Zulkifli Oladipo Ishola for their commitments in their life time at ensuring a good legacy for me –Education.

ACKNOWLEDGEMENTS

In the name of Allah, Most Gracious, Most Merciful

First and foremost I give all praises and adoration to Allah (SWT) for granting me good health, patience, and wisdom to complete this study. I am highly indebted to my supervisors Professor Dr. Mohd Nawawi Bin Mohd Nordin and Associate Professor Dr. Khiruddin Bin Abdullah for their guidance, time devotion, and supports throughout my study. I have been privileged to carry out this work under their tutelages.

Also, I am grateful to the members of our research group Messrs Sobri, Naim, Khaizal, and Klein. All your contributions to this work are highly appreciated.

I am particularly grateful to the Technical Staff of the Geophysics Unit, School of Physics, Universiti Sains Malaysia (USM), for their commitment and perseverance during the data acquisitions Messrs Mydin Jamal, Yaakob Othman, Shahil Ahmad Khosani and Azmi Abdulah (AA) and not to forget Mr. Low Weng Leng (Rtd.). Also, to members of the academic staff of the geophysics unit, Dr. Hassan Baiuomy, Dr. E. Amin Khalil and Dr. Nordiana Mohd Muztaza. All my postgraduate colleagues in geophysics unit, School of Physics, USM. I thank you for your supports during field data acquisitions and in using some of the equipments. Also, I express my deepest appreciations to Dr. Dale Rucker of hydroGEOPHYSICS. Inc., Tucson, USA for providing part of the field data used in this research.

My appreciations go to my family for their understanding, encouragement, and for being my unfailing pillar of support. My dearest wife, Mrs Shakirah Ishola, my lovely children Abdullah Ishola, Aisha Ishola, and Maryam Ishola. Also, Fatimah Ishola and Abdulganiu Ishola and their mother Mrs Monisola Ishola. I thank you all. Also, my siblings, Saheef Taiwo Ishola (twin brother), Mrs Abiodun Adeniran and her family as well as Feyiropo Ishola. My in-laws Ahmed, Alhaji and Mrs Williams. I appreciate all your supports, understanding, and contributions. My colleagues at my place of work Drs. L. Adeoti, M. Olopade, and I. Olasupo and others too numerous to mention. May Allah reward you all in manifold for your contributions. Very special thanks to my friends too many to mention. I am also grateful to Messrs Bashir Ishaku and Rami Ibrahim for his immense assistance. I am thankful to University of Lagos, Akoka, Lagos, Nigeria for granting me study leave in pursuit of PhD degree and the Federal Government of Nigeria through TETFUNDS scholarship for funding my studies. I also express my profound appreciations to Profs. S.B. Olobaniyi, (Head, Department of Geosciences) A.B. Ayolabi and M.S. Ilori, Dean of the Faculty of Science, University of Lagos, Akoka for their supports towards my program.

To my examiners, Prof. Abdulrahman Samsudin, Prof. Fauziah Ahmad, and Asso. Prof. Kamar Shah, I am thankful for their invaluable advice which has improved my thesis. Finally, my gratifications go to the people of Malaysian for their hospitalities throughout my study.

TABLE OF CONTENTS

LIST OF APPENDICES

LIST OF TABLES

LIST OF FIGURES

Page

LIST OF SYMBOLS

F	Function
$\boldsymbol{m}$	model space
$\boldsymbol{d}$	data space
σ	electrical conductivity
∇	Differential operator
V	Voltage
I	Current
δ	Dirac delta function
α	regularization parameter
w_d	Data weighing matrix
m_{ref}	reference model
Wm	Model weighting matrix
$\vec{E}$	Electric field intensity
q	Charge density
ε_o	Permittivity of free space
U	Electric potential
$\vec{j}$	Current density
ρ	Resistivity
K	Geometric factor
r	radius
ρ_a	Apparent resistivity
a	Electrode spacing
n	dipole length factor/propagation factor
m_a	Apparent chargeability
V_s	Secondary Voltage
V_p	Primary Voltage
η	Chargeability
M	Current dipole
J_e	Primary current due to external source
J_n	actual current density
ϕ_e	Potential field due to primary current
λ^1	Lamé constant
μ	shear modulus
K	Bulk modulus

v	Poisson's ratio
φ	potential displacement for p-wave
Φ	potential displacement for s-wave
dx	Geophone interval
V_f	Phase velocity
Δt_f	time interval between geophones

LIST OF ABBREVIATIONS

AC	Alternating current
AE	Absolute error
Avg	Average

CI	Classification image
Dpd	Dipole-dipole
DOI	Depth of investigation
DC	Direct current
DIP	Dispersive property
DI	Dispersion image
EC	Electrode configuration
ERI	Electrical resistivity imaging
FD	Finite difference
Grd	Gradient array
GIS	Geographical information system
HIS	Intensity-hue-saturation
HVL	High velocity layer
IP	Induced polarization
IMP	Image processing technique
ISODATA	Iterative Self-Organizing Data Analysis Technique
IRM	Inversion resistivity method
IVL	Intermediate velocity layer
JH	Jenderam Hilir
KK	Kuala Kangsar
LVL	Low velocity layer
Max	Maximum
Min	Minimum
Med	Median
MAE	Mean absolute error
MAPE	Mean absolute percentage error
MASW	Multichannel analysis of surface waves
NNINT	Natural neighbor interpolation
OA	Overall accuracy
PA	Producer's accuracy
Pdp	Pole-dipole
RGB	Red Green Blue
RMS	Root mean square
SC	Supervised classification
Sch	Schlumberger
TK	Teluk Kumbar
UM	Unified Model
USC	Unsupervised classification
UA	User's accuracy
Vp	Compressional wave velocity
Vs	Shear wave velocity
Wsc	Wenner-Schlumberger

LIST OF PUBLICATIONS

Ishola K. S., Nawawi, M.N.M., Abdullah; K., Adiat, K. A.N and Abdulraham, A. (2013). Evaluation of two-dimensional electrical resistivity models using image processing method. Short paper on 13th SAGA Biennial Conference and Exhibition in South Africa. Conference proceedings.

Ishola K. S., Nawawi, M.N.M., Abdullah; K., Sabri, A.I.A and Adiat, K. A.N. (2014a). Assessment of the reliability of reproducing two-dimensional resistivity models using an image processing technique. Springer Plus 3: 214, 1-12, doi: 10.1186/2193-1801-3-214

Ishola K. S., Nawawi, M.N.M., Abdullah, K. (2014b). Application of k-means algorithm to two-dimensional electrical resistivity imaging. Second International conference on Engineering Geophyics. EAGE. Al Ain, United Arab Emirates, 24-27, November 2013. Conference proceedings.

Ishola K. S., Nawawi, M.N.M., Abdullah, K. (2014c). Application of data mining for geophysical images Classification and interpretation. IEEE conference on geoscience and remote sensing. 19-20 November, 2014. IWRS, Selangor.

Ishola K. S., Nawawi, M.N.M., Abdullah, K. (2014d). Combining multiple electrode arrays for 2-D electrical resistivity imaging using unsupervised Classification Technique. Pure and Applied Geophysics. Doi: 10.1007/s00024-014-1007-4.

Ishola K. S., Nawawi, M.N.M., Abdullah, K and A. Abdulrahman. (2015). 11th International Geosciences Conference of the Saudi Society of Geosciences, Riyadh, Saudi Arabian. Accepted for oral presentation. 11-15th May, 2015.

PEMBANGUNAN MODEL BERSEPADU BERSATU BAGI MENCIRI DATA GEOFIZIK MENGGUNAKAN TEKNIK PEMPROSESAN IMEJ

ABSTRAK

Kajian geofizik menggunakan pengimejan kerintangan elektrik telah dijalankan di mana susunatur elektod yang berbeza atau pelbagai kaedah geofizik telah digunakan bagi mendapatkan maklumat lengkap mengenai model lapisan Bumi. Tujuan penyelidikan ini adalah untuk menggabungkan kaedah-kaedah tersebut kepada satu imej bersepadu yang akan meningkatkan keseluruhan kualiti dan kebolehpercayaan imej-imej geofizik bagi pencirian subpermukaan. Bagi tujuan ini, kaedah pengelasan tanpa pengawasan dengan algoritma berkelompok telah digunakan. Bagi mencapai tujuan tersebut, pengimejan kerintangan berdasarkan set-set data daripada susunatur elektrod berlainan telah digunakan bagi kedua-dua data sintetik dan lapangan sebagai contoh. Bagi kes sintetik, parameter-parameter statistik asas (iaitu minimum, maksimum, median dan purata) telah diperkenalkan bagi menggabungkan imej-imej yang berlainan kepada satu imej tunggal. Juga, imej-imej pasca songsangan 2-D telah digabungkan kepada satu imej tunggal menggunakan kaedah pengelasan K-min dengan beberapa parameter awal yang ditentukan sebelum melaksanakan prosedur-prosedur pengelompokan dan pengelasan. Kesemua imej-imej songsang geofizik telah dipra-proses, dimanipulasi dan analisis imej dijalankan menggunakan perisian pemprosesan imej, PCI Geomatica. Prestasi bagi imej-imej geofizik dijalankan dalam contoh-contoh sintetik menggunakan ralat mutlak min,

ralat peratusan mutlak min dan pembangunan jadual matrik ralat manakala bagi imej-imej bersatu data lapangan log-log litologi lubang gerudi telah digunakan.

Keputusan menunjukkan bahawa imej yang paling baik mewakili model-model sebenar yang sama dengan susunatur elektrod individu diperolehi daripada capaian maksimum. Justru, ketepatan keseluruhan dan pekali-pekali kappa menunjukkan persetujuan yang baik di antara imej-imej model bersatu dan sebenar. Tambahan pula, imej-imej bersatu yang diperolehi daripada penggabungan pasca-pengelasan bagi beberapa kelompok menunjukkan terdapat dua ke empat kumpulan yang mewakili ciri-ciri bagi model-model dalam contoh sintetik manakala dalam contoh lapangan, model-model bersatu mengandungi antara tiga ke enam kumpulan. Setiap kumpulan ini dicirikan dengan parameter-parameter geofizik diukur (iaitu kerintangan elektrik, kebolehcasan atau halaju seismik pembiasan) bersama dengan unit-unit litologi tersimpul. Keseluruhannya, geologi tempatan subpermukaan bagi kawasan-kawasan kajian dicirikan sebagai tanah liat, pasir berliat, lempung, pasir dan pasir berbatu. Kes-kes yang dipertimbangkan di dalam kajian ini dapat dilihat sebagai penggunaan yang berjaya bagaimana teknik pemprosesan imej ini dapat menjadi alat tambahan yang menyakinkan dimana penggabungan bagi data geofizik yang berlainan adalah kunci untuk pencirian subpermukaan yang lengkap.

DEVELOPMENT OF UNIFIED MODEL FROM THE INTEGRATED GEOPHYSICAL DATA USING IMAGE PROCESSING TECHNIQUE

ABSTRACT

Conducting geophysical surveys using electrical resistivity imaging where different electrode array configurations are used or multiple geophysical techniques are employed in order to obtain comprehensive information about the layered Earth model. The purpose of this research was to combine these techniques into an integrated unified image that increases the overall quality and reliability of the geophysical images for subsurface characterization. To this end, unsupervised classification technique via clustering algorithm was employed. To meet this need, resistivity imaging based on data sets from different standard electrode arrays was conducted for both synthetic and field data examples. For the synthetic case, basic statistical parameters (i.e. minimum, maximum, median, and average) were introduced to merge the different images to a single unified image. Also, the 2-D post inversion images were combined to a single unified image using k-means clustering technique with some initial parameters defined before implementing the clustering and classification procedures. All the inverse geophysical images were pre-processed, manipulated, and image analysis carried out using an image processing package, PCI Geomatica. The performance of the geophysical images

was carried out in the synthetic examples using mean absolute error, mean absolute percentage error and construction of an error matrix table while for the field data unified images available borehole lithologic logs were used.

Results show that the best images representing the true models comparable to those from individual electrode configurations were obtained from maximum approach. Additionally, the overall accuracy and kappa coefficients show good agreement between the unified and true models' images. Furthermore, the unified images obtained by the post-classification merging of some clusters show that there are two to four groups representing features of the models in the synthetic examples while in the field examples, the unified models contain between three to six groups. Each of these groups is characterized by the measured geophysical parameter(s) (i.e., electrical resistivity, chargeability or seismic refraction velocity) together with inferred lithological units. Overall, the subsurface local geology of the study areas is characterized by clay, clayey sand, silt, sand and gravelly sand. The cases considered in this study can be viewed as a successful application of how image processing technique can be a promising additional tool where combination of different geophysical data is the key to a comprehensive subsurface characterization.

CHAPTER 1

INTRODUCTION

1.1 Background

Geophysical imaging is used to picture Earth's subsurface. The use of near surface geophysical methods allows subsurface features to be located, mapped and characterized in responses to changes in physical, electrical or chemical properties in the subsurface. The location and orientation of anomalies are essential for modeling of the subsurface geology. In applied geophysics, modeling has become an essential tool for comparison of the resolution of different electrode configurations (Martorana et al., 2009). Geophysical images obtained are used for mapping the extent of occurrence of some natural resources in the subsurface and to interpret them on the basis of their physical properties (e.g., electrical conductivity, density, velocity, and electric permittivity).

As earth scientists our main goal is to obtain adequate and reliable characterization of the subsurface. For instance, an aquifer unit is mapped for optimum water production in hydrology and hydro-geophysical studies or the estimation of soil's strengths/parameters for geotechnical purposes. The bottom line in any of these fields of active research is to quantify and reduce uncertainties, unambiguities or mis-interpretations of results for subsurface exploration, characterization and management.

In electrical resistivity surveys, high resolution and reliable and better imaging depends on the choice of electrode configuration (EC). The electrode configuration used should provide adequate information about the earth's model (Dahlin and Zhou 2004). The selection of the most appropriate EC has continued to draw rapt attention among researchers in view of their merits and limitations (Olayinka and Yaramanci 1999). The type of arrays chosen and the model parameters of the investigated structures would influence substantially the results of the survey.

Several studies have been carried out regarding the performance and efficacy of various ECs. Some commonly used ECs in resistivity studies are the Wenner (Wen), Schlumberger (Sch), Wenner-Schlumberger (Wsc), dipole-dipole (Dpd), pole-dipole (Pdp), and pole-pole (Pop). Others are gradient (Grd) and square arrays (Squ) (e.g., Reynolds 1997; Sharma 1997). It is generally recognized that Wenner and Schlumberger arrays are less sensitive to noise and have high vertical resolution whereas Dpd array has lower signal-to-noise ratio but better lateral resolution (Barker 1979; Dahlin and Zhou 2004). Also, Roy and Apparao (1971) and Barker (1989) studied the depth of detection of different array types. The resolution and accuracy of the inverted data sets were investigated (Sasaki 1994; Beard and Tripp 1995; Candansayar and Basokur 2011; Dahlin and Zhou 2004) while Ward (1990) reviewed the performances of four EC on some geologic structures.

There has been increasing quest for more holistic approach to characterizing the subsurface through the application of quantitative methods to improved understanding of the structures beneath the earth surface. To this end, several

geophysical techniques have been used and are still being employed for probing beneath the earth's surface. Some of the pioneering works in this regard include Corwin and Hoover, 1979; Sill, 1983; Butler, 1984; Fitterman and Stewart, 1986; deGroot-Hedlin and Constable, 1990; Greaves et al., 1996 and Xia et al., 1999.

Most of the instrumentations used in geophysical surveys respond only to a physical parameter e.g., in electrical resistivity method to electrical conductivity, induced polarization to chargeability and seismic methods to velocity and ground penetrating radar to dielectric permittivity. It is therefore not adequate to say that surveys carried out at a particular location with one geophysical technique could offer greater insights into the subsurface conditions since each method will yield information about a relatively independent aspect of the subsurface (Weymouth, 1986). For instance, in a survey carried out using electromagnetic method, the interpreted result using this technique might suggest the presence of foundations of an historical structure. On the other hand, if the survey is carried out with magnetic method, the inferred interpretation could indicate that the delineated anomaly is of hearth. This rather puzzle nature (ambiguity) of the subsurface could be resolved when more than one methods is employed (Clay, 2001; Gaffney and Gater, 2003).

The use of an individual technique in a survey can only be appropriate in the detection or discrimination of a target provided there exist a good contrast in the measured physical properties between the target (structure) and the surrounding material (background). This contrast could be in terms of the subsurface characteristics to which the technique responds (e.g., electrical conductivity,

magnetic susceptibility etc). Each technique or group of techniques has/have some degree of success depending on the goal or aim of the study.

Against this background, when a geophysical investigation is conducted over a region containing unknown structures and only one method is employed in this survey, there is a high likelihood that the method might miss part of the structure or mis-interpret the structure completely (Dutta et al., 2013). Therefore, it follows that since no single geophysical method can yields an optimal information about the subsurface, combining a suite of geophysical methods for a proper characterization becomes imperative (Sauvin et al., 2013). Thus, an optimum strategy is to combine the best of each method to produce results much more informative than would be possible from deployment of individual methods alone.

In order to produce a robust, reliable and adequate characterization of the subsurface local geology of an area, the different information that are obtained from different methods or sensors at the same time or different time should be integrated. Some of the pioneering works that combine different geophysical techniques include: electrical resistivity (ER) and induced polarization (IP) surveys (Oldenburg, 1999); ground penetration radar (GPR), seismic and electrical methods (Garambois et al., 2002); electrical methods, seismic refractions, and GPR (Demanet et al., 2001). Others are Gallardo and Meju, 2003, 2004 in joint inversion of two-dimensional (2-D) resistivity and seismic travel time; Van-Dam, 2012 used a suite of geophysical techniques for characterizations of landforms. Also, concerted efforts have been made to combining geophysical data using statistical techniques particularly for hydrological investigations Ezzedine et al., 1999; Hubbard and Rubin, 2000; Chen et

al., 2001; Bosch, 2004; Tronicke and Holliger, 2005; Linde et al., 2006; Fregoso and Gallardo, 2009; Dubreuil-Boischair et al., 2011).

The use of a surface geophysical method alone for subsurface measurements is sometimes subjective due to uncertainties that are inherent in data acquisition, processing, and interpretation. It is therefore necessary to express the information contents of the subsurface structures through quantitative integrative approach. Combining data of different kinds can sometimes bring in information that can help reduce uncertainties in measurements. For example, complementing electrical imaging with induced polarization methods would allow the interpreter to distinguish between e.g. sand formations with seawater intrusion and clay formations or help to delineate landfills (Weller et al., 2000; Dahlin et al., 2002; Binley and Kemna, 2005; Marescot et al., 2008).

Also, measurements of both seismic refraction compression P-wave velocity and electrical resistivity is important as the former provide not only greater depth of investigation (DOI) especially mapping of bedrocks but also help resolve the poor resolution of ERI at greater depths. Furthermore, studies have shown that shear wave velocities (Vs) in addition to compressional wave (Vp), can help to resolve ambiguities in lithological identification and discrimination (Baucer, 2003; Jongmans et al., 2009). The individual inversion models of seismic refraction compressional (P) and shear (S) waves velocities measurements do not provide consistent subsurface models but a combined model resolve better the velocity reversal for instance associated with S-waves than individual model alone (Song et al., 2007).

Subsurface modeling and characterization due to limited accessibility and lack of direct observation of the heterogeneity of the investigated area is inherently a difficult task. Nevertheless, an integration of diverse sources of information into a single model could provide better understanding of the subsurface conditions (Deutsch, 2003; Caers, 2005). However, the techniques of integrating different geophysical data sets still remain a challenge. One of the constraints of using an integrative approach is linked to the different scale and attributes or the measured parameters of the geophysical techniques. Another reason that could be advanced for the difficulty in integrated approach is the lack of knowledge of the relationship between the different physical properties, for instance, between velocity and resistivity measurements (Carcione et al., 2007). However, the establishment of a precise relationship between different subsurface geo-materials still remains a problem (Kotyrba and Schmidt, 2013).

Against the backdrop of the limitations in some of the available techniques for combining geophysical data, this study attempt to explore the applicability and relevance of an unsupervised classification technique (USC) in the context of clustering for the integration of different geophysical data sets leading to the development of unified classification models (UCM) for the characterization of the subsurface geological units in the study areas.

The purpose of clustering is to partition a set of data (objects/pixels of the image) into a predefined number of groups or clusters such that the intra-cluster variability is minimized, at the same time maximize the inter-cluster variability between the objects. Clustering has been increasingly used in many areas of active

research in solving problems related to medicine, psychology, biology, sociology, pattern recognition and image processing. In geosciences, it has proven to be a vital tool in establishing certain patterns by organizing the measured physical parameters/attributes of the data sets into clusters. USC using clustering algorithms have proven to be vital tools for not only automatically extracting information required for structural exploration but also useful for the integration of multi-dimensional datasets especially for remote sensed datasets. In this research, we employ the clustering technique using *k*-means algorithm to combine and zone the resulting geophysical images into clusters that bear relationships with physical parameters/attributes of the models.

1.2 Problem Statements

Geophysical techniques are widely used in near surface characterization. Sometimes different geophysical datasets are acquired and separately interpreted in order to obtain relevant information beneath the earth's surface. As a result, the synergies between the different techniques in terms of acquisition and inversion are often ignored. The limitations of each geophysical method could be overcome by the simultaneous integration of individual data sets with the assurance that different information obtained from the datasets could provide a comprehensive understanding and characterization of the subsurface structures or targets.

To this end, only a holistic data analysis based on integrating multiple geophysical techniques can provide reliable, all-embracing information about the subsurface conditions or geology. Some techniques of data integration have been introduced and implemented in the field of earth sciences. Gallardo and Meju, 2003,

2004; Beaujean et al., 2010; Haber and Oldenberg, 1997; Tryggvason and Linde, 2006 used joint inversion techniques to combine various geophysical images data sets on the basis of prior knowledge of structural relationships or existing empirical relationships between the physical attributes of the different data sets.

However, the results of the simultaneously joint inversions of the different geophysical data sets were limited due to the need for establishment of an empirical or mathematical relationships and common structural links between the measured parameters of the models jointly combined (Haber and Oldenburg, 1997). Paasche et al., 2006; Fraser and Dickson, 2007; Paasche and Eberlie, 2009; Dietrich and Tronicke, 2009 and Linder et al., 2010 in their respective studies have successfully implemented data integration through image classification techniques. They focused mainly on the use of Fuzzy clustering technique where degree of membership of the data sets is highly essential. Using this technique involves assigning membership to each data point leading to long period of implementation.

Against the backdrop of some of the challenges militating against reliable integration and interpretations of geophysical data, this study attempts to apply an unsupervised classification technique using the *k*-means clustering for the integration of geophysical data. The importance of this clustering technique is that rather than relying on the existence of empirical relationships (e.g. a well-known Archie's law which allows the direct conversion of electrical resistivity into porosity of granular sediments are usually scaled using a number of sparse observations where the target parameter porosity and the constrain parameter, electrical resistivity distribution are commonly known) or presence of common structural features at the boundary of the

measured parameters (Haber and Oldenburg, 1997 and Gallardo, 2007), the clustering technique naturally partition the data sets into groups or clusters without any recourse to empirical relationship between the attributes parameters of the models.

1.3 Objectives of this research

The primary objective of the study is to develop unified models for the characterization of near-surface targets by integrating multiple geophysical data using clustering technique.

The specific objectives that summarize the importance of this research are to:

(i) Investigate the suitability of different electrode arrays for imaging of the near subsurface geological structures using synthetic models
(ii) Determine the applicability of image classification technique in processing and analyzing multiple geophysical data sets
(iii) Develop an approach for combining the post inversion models for subsurface structural characterization and unified models from integration of multiple geophysical techniques
(iv) Verify through accuracy assessment the performance of the developed models using different validation parameters.

1.4 Scope and limitations of study

The focus in this study is on the imaging capabilities of three conventional electrode arrays using both 2-D numerical synthetic resistivity models to simulate

various geological conditions ranging from hydrogeological, archaeological to environmental applications and real field geophysical data collected at different study locations. The image processing technique used throughout this research is the PCI Geomatica package which is widely used in remote sensing environment for images processing. However, in this study its applicability is extended to manipulating, processing, and analyzing subsurface geophysical images. Among the functionalities available in this package that were explored includes interpolation, re-sampling and clipping for images pre-processing and EASI (Engineering Analysis and Scientific Interface) Modeler and image classification of data merging and post classification analysis.

Although, the geophysical techniques used for the subsurface investigations are restricted to electrical resistivity imaging, induced polarization, and seismic refraction methods (i.e., active methods), however, the principles can be extended to other passive geophysical techniques. These measurements and analysis are carried out through separate inversion and integrative approach for comprehensive understanding of the subsurface structures. This study is only limited to use of resistive target(s) embedded in conductive background for the synthetic resistivity models only and unsupervised classification technique using only *k*-means classifier algorithm.

1.5 Motivation and significances of study

Subsurface characterization is a crucial step in understanding the depositional era or period and properties of the geological units that made up a particular study area and therefore extremely vital in near surface exploration and applications to

environmental, engineering and related problems. The study has provided a scientifically based technique through which multi-parameters data sets obtained from geophysical measurements can be integrated for proper characterization and understanding of the subsurface geology.

Traditionally, skilled interpreters delineate 2-D geophysical images by visual inspection and assessment, overlaying of two or more images to obtain common anomalies or target of interest. However, most often, the results of this task are usually subjective, ambiguous, misleading, and biased. For these reasons, considerable efforts have been devoted to a more efficient, robust, and holistic approach to combining geophysical data through clustering techniques (Piro et al., 2000; Anderson-Mayes, 2002; Tronicke et al., 2004; Enderle and Weih, 2005; Barainne et al., 2006; Kvamme, 2006; Mauriello and Patella, 2008; Altdorff and Dietrich, 2009; Ernenwein, 2009; Passche and Eberle, 2009; Ogden et al., 2009; Song et al., 2010; Abedi et al., 2012; Apostolopoulos and Orfanos, 2013; Kotyrba and Schmidt, 2013 and Di Giuseppe et al., 2014).

This research investigates the application of cluster technique to geophysical data sets from different methods to provide robust geological information of the subsurface structures. Unlike some of the integration techniques available, it does not require any a priori knowledge of the structural similarity between the models as well as mathematical interrelationships between the different datasets. Image classification techniques have made crucial contribution to a broad range of areas but this potential has not been fully utilized in the field of earth sciences in particular for the characterization of the near-surface targets. Consequently, this study attempts to

apply unsupervised classification for the integration of different geophysical data with a view to developing, a reliable, robust and comprehensive model for the subsurface conditions in the study areas.

Thus, other significances of this research are establishing a methodology for combining post inversion models of different electrode array configurations for near-surface electrical resistivity imaging, and application of clustering technique to integrate multi-parameters images to diverse geologic problems leading to the establishment of cluster-litho attributes relationships. The novel approach allows image classification to be performed on different types of data and different applications with minimal human effort required.

Since this research was designed for near surface characterization by incorporating multiple geophysical techniques, the proposed techniques for merging different types of geophysical data sets show the potentials of being applied in many areas of interest. These include, environmental monitoring (i.e., groundwater studies, contaminants modeling and remediation), locating stratigraphic features in the subsurface (e.g., geological formation), mining/minerals exploration. Also, the models developed through this integration technique could satisfy the ever growing need within the private sector and scientific community for a vital time-and cost-effective approach for integrative analysis of two or more geophysical data.

1.6 Novelties

The novelties in this study are:

(i) The study will be the first to be embarked upon to quantitatively interpret subsurface local geology of the study areas by developing unified classification models using clustering technique. Application of the unsupervised classification technique results in successful analysis of the layered earth model for which inferences are not possible/might be difficult through methods used in past studies.

(ii) An evaluation of results demonstrates that unsupervised classification approach discovers meaningful groups/clusters which are linked to the various geological units in the study areas and the establishment of clusters-litho attributes relationships.

1.7 Thesis structure

The layout of this thesis is structured as follows:

In chapter 1, background to study is introduced. The aim and objectives of the research to be achieved are highlighted. Also, the research questions to addressed, scope and limitations as well as contributions are presented.

Chapter 2: This presents a detail literature review of different geophysical techniques that are used for probing the subsurface. Also, concerted effort was devoted to discussing the basic concepts and theoretical background of the main geophysical methods used in this study. Furthermore, the concepts of image processing and its area of applications were enumerated. In connection to this, clustering technique including commonly used algorithms for partitioning is

explained. Finally, a review of previous research on integration techniques employed in geophysical data is presented.

Chapter 3: It discusses the first phase of the methodologies used in this study. First the imaging capabilities of three electrode configurations namely the dipole-dipole, the Wenner-Schlumberger, and Pole-dipole were presented using synthetic numerical models. Then, the reliability of an image processing package for the reconstruction of 2-D inverse resistivity models was mentioned. Manual ways of merging models' resistivity values as well as clustering procedures using k-means clustering algorithm are explained. The steps used to evaluate and validate the accuracy of the both manual and classification approaches for both individual and combined models' images are presented.

In addition, the geophysical techniques employed during the study are presented and explained. A total of four sites have been surveyed by means of geophysical techniques. The first two test sites, discusses the collection of electrical resistivity imaging data using different electrode arrays followed by the application of the image processing technique. The next two test sites emphasized on the application of the image processing technique on rather multi-parameters data acquired using the electrical resistivity imaging, induced polarization, and seismic refraction (i.e., both P-and S-waves) methods. Then, the integration of the post inversion models data sets using the clustering algorithm and the generation of unified classification models are examined.

Chapter 4: The results and discussion of the results are provided herein. The first part of this chapter, presents the results and discussion of the application of

image processing technique on the synthetic resistivity models. Also, the results of each site surveyed are presented and explained. This is followed by results and discussion of using different electrode arrays for the electrical resistivity surveys are presented followed by the application of clustering technique on multi-parameters measurements taken from electrical resistivity imaging, induced polarization, and seismic refraction and seismic multichannel analysis of surface waves are thoroughly described and interpreted with borehole available at the sites. Also, validations of the results for the models are presented.

Chapter 5: Final conclusions are drawn in this chapter. The conclusions are related to: the numerical simulations of different electrode configurations using synthetic models results, findings on the capabilities of the geophysical techniques in relation to the study areas and integrated approach leading to the generation of classification models. Also, the contributions of this study to knowledge are highlighted. Finally, a number of topics/issues for further research following the line of investigations of this thesis are proposed.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

One major task of this work is the application of image processing techniques through clustering, an unsupervised image classification on geophysical datasets to assist in subsurface conditions characterization. Two applications area are focused on: 1) clustering/classification of different electrode array configurations electrical resistivity data sets for synthetic model and 2) the identification of clusters/groups from the integration of different geophysical techniques datasets. Due to the focus in k-means algorithm, the motivation for the development of unified model for different geophysical techniques is reviewed. As the aim is towards data integration, relevant background works are covered. In this way, it is hoped that the knowledge gaps which this research aim to fill are identified.

2.2 Classification of basic geophysical methods and parameters

Geophysical investigations are based on measurements of some physical parameters of the subsurface structures. The variations in these parameters which depend on soil types and interstitial fluid properties form the basis for the utilization of different geophysical techniques. Kearey et al. (2009) provide a detailed description of available geophysical methods. A classification of the available methods in accordance with their underlying physics and the measured physical property of the earth are provided is presented as Table 2.1. According to the table, many different physical properties of the subsurface can be measured. These

measured parameters give different information about the subsurface conditions. The physical properties of interest for the purpose of this research are the electrical conductivity, chargeability, and velocity. On this account, only electrical resistivity imaging, induced polarization, and seismic refraction methods are reviewed.

Table 2.1 Classification of geophysical methods via measured data and physical parameters

Method	Measured data	Physical parameter
Gravity	Gravitational field of the Earth in space and time	Density
Geo-electric	Earth resistance	Electrical resistivity
Electromagnetic	Response to electromagnetic waves	Electrical resistivity
Magnetic	Geo-magnetic field in space and time	Magnetic susceptibility
Seismic	Travel time for refracted or reflected wave	Density and elastic moduli
Induced polarization	Voltage decay	Electrical chargeability
Radar	Travel time of reflected radar	Dielectric constant
Self potential	Electric potential	Electric resistivity
Nuclear magnetic resonance	Relaxation electromagnetic field	Fluid content and relaxation constants

2.3 Forward Problem

In Earth sciences, measurements of the earth model are taken in order that a set of model parameters could be drawn about the underlying volume of the earth. It is important to note that direct measurements of these parameters are difficult to carry out. Thus, an experiment must be set up in which knowledge of the physics is assumed. In order to determine whether or not an adequate model has been selected, there is a need to choose model to reproduce the observed data. Forward problem is the ability to simulate measurements of an arbitrary model.

In mathematical context, forward problem is equivalent to applying a function, F, to a vector ***m*** in model space to arrive at a vector ***d*** in data space as shown in Figure 2.1. In practical sense, forward problem is the act of taking measurements. These measurements then provide information about the earth model wherein the physics involved can be thought of as the functional.

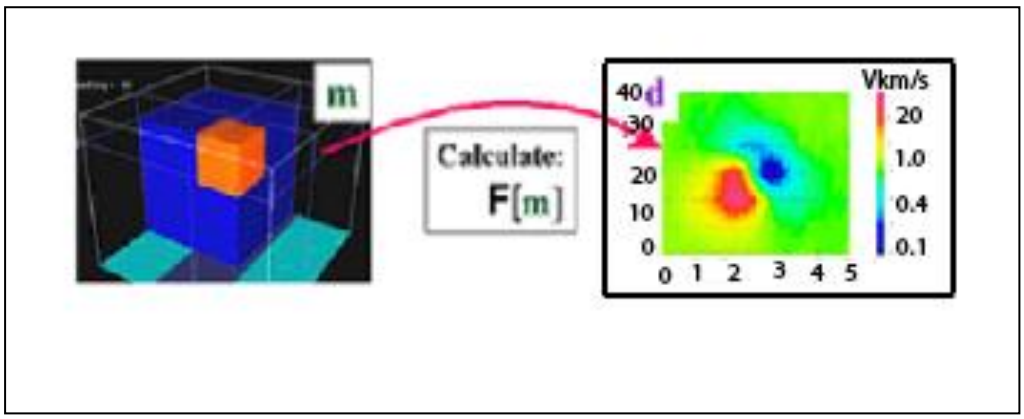

Figure 2.1 illustrating forward modeling (modified after Park et al., 2014)

2.4 Inverse Problem

The inverse modeling is the process of trying to recover parameters from experimental measurements i.e. is an attempt to solve an inverse problem. In this case, we consider a vector in data space, ***d***, and we would like to map it back to a vector in model space, ***m,*** as illustrated in Figure 2.2. An inverse problem is fundamentally non-unique since there exists an infinite number of models which can fit the data (Tarantola, 1987; Oldenburg and Li, 2005).

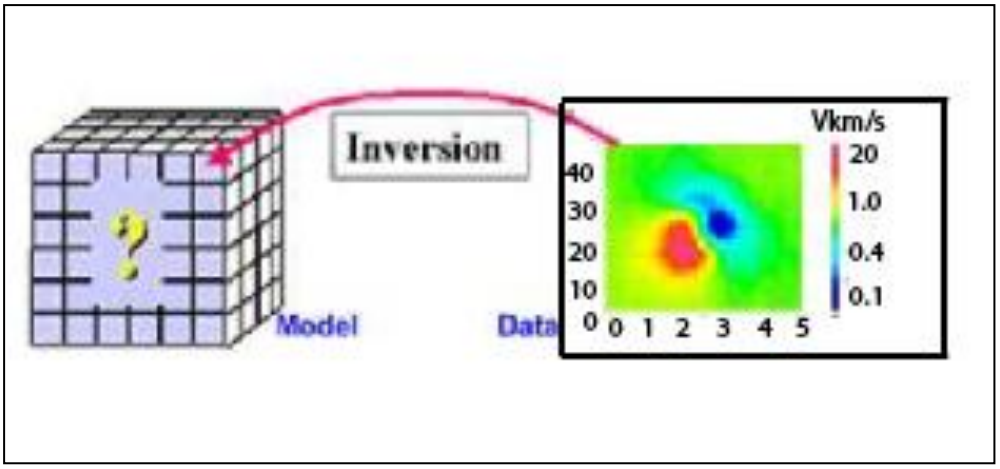

Figure 2.2 Illustrating inverse modeling (modified after Park et al., 2014)

To formulate an inverse problem, the electrical properties considered is discretized into a set of parameters defined by a model vector, ***m***. For a 1-D problem, ***m*** normally contains the parameter (e.g. conductivities, or velocities) and thickness of a multilayer model while for arbitrary 2-D and 3-D distributions its elements generally correspond to the measured parameters of individual elements or cells of the finite difference mesh used in forward modeling that is:

$$m_j = \log \sigma_j \ (j = 1,...,M) \tag{2.1}$$

The logarithm accounts for the large possible range in earth conductivity (Loke, 2005). The inverse problem tries to find a model ***m*** which using the forward mapping according to:

$$\nabla.(\sigma \nabla V) = -I\delta(r) \tag{2.2}$$

to reproduces data ***d*** to the specified level of uncertainty. However, due to inherent non-uniqueness of the resistivity inverse problem, together with the presence of data errors, can effectively lead to an extremely ill-posed numerical problem, additional constraints must be imposed on the inversion. This is normally accomplished by solving the inverse problem as a regularized optimization problem (Tikhonov and Arsenin, 1977), where an objective function to be minimized is of the form:

$$\varphi(m) = \varphi_d(m) + \alpha\varphi_m(m) \tag{2.3}$$

$$\varphi_d(m) = \left\| w_d \left[d - f(m) \right] \right\|^2 \tag{2.4}$$

The first term on the right hand side of equation (2.3) is a measure of the data misfit, with f denoting the forward operator, w_d represents the data weighing matrix associated with the individual (uncorrelated data errors). In addition, $\varphi(m)$ contains a stabilizing model objective function usually expressed as:

$$\varphi(m) = \left\| Wm\left[m - m_f \right] \right\|^2 \tag{2.5}$$

It is used to incorporate certain model constraints relative to a reference model $\mathbf{m_{ref}}$ by suitable choice of a model weighting matrix, Wm. The regularization parameter α in equation (2.3) controls the tradeoff between the data misfit and the model objective function.

2.5 Electrical Resistivity Imaging (ERI)

Electrical resistivity imaging is based on the measurements of electrical resistivity and a rapid means of generating spatial models of physical properties of the subsurface (Daily et al., 2004; Chambers et al., 2006). It requires precision of electrical measurements to be made repeatedly with the recording systems. After collecting sufficient measurements, a solution to the electrical properties of the subsurface can be obtained through inversion scheme. The application of ERI has been on the increase especially in near-surface investigations involving hydrogeological, geotechnical, and other shallow explorations in the last two decades. The increasing interest in this method started with the invention of the multi-electrode systems (Griffiths and Turnbull, 1985). ERI data acquired can be interpreted using appropriate inversion techniques with a view to obtaining resistivity distributions as close as possible to the true earth model. An arrangement of electrodes for resistivity imaging and sequence of measurements for 2-D survey (i.e., where it is assumed that resistivity vary both vertically and laterally along the survey line) is shown in Figure 2.3.

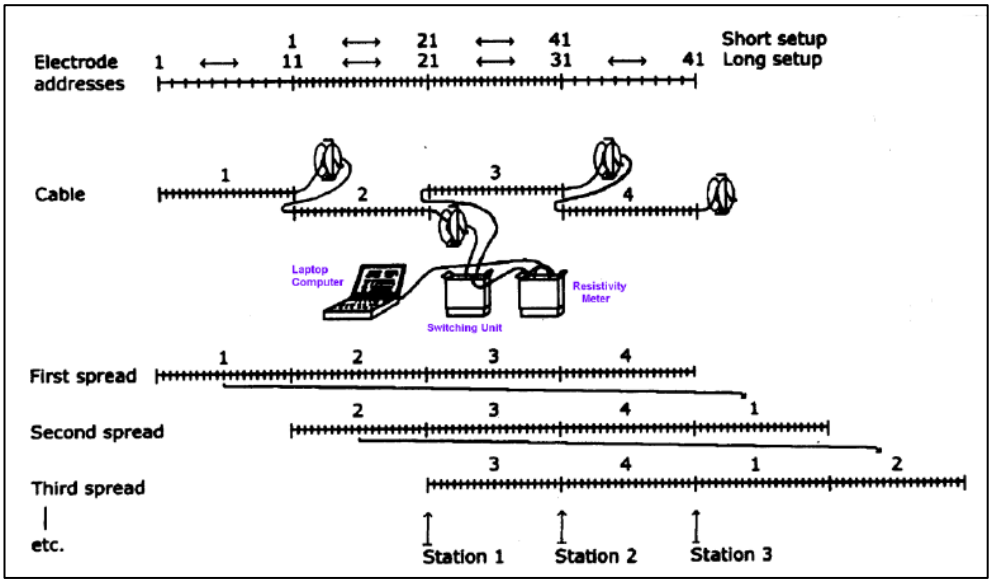

Figure 2.3 Principle of measurement using resistivity meter and automated switching unit for ERI/IP measurements (Loke, 2001).

2.5.1 Basic concepts of ERI

A fundamental property of any volume of a material is it resistance measured in Ohm. The resistance is defined as the material's opposition to the flow of electrical current (Reynolds, 1997). Resistivity (Ohm-m) is related to this property and is expressed as a resistance through a distance, which makes it independent of material geometry. Resistivity is considered as function of rock porosity, volumetric fraction of saturated pores, and the resistivity of the pore water (Archie, 1942). In most cases, it is the pore fluids of the rock that account for the overall resistivity signature rather than the host rock (Lowrie, 1997). In resistivity measurements, the basic procedure is to establish a subsurface distribution of resistivity by injecting current into the ground between two current electrodes. The resulting potential differences are measured between pairs of potential electrodes in a line or grid

(Ramirez et al., 1993). The interpretations of the measured parameters yield information about the electrical conductivity beneath the subsurface.

The input current and the knowledge of the potential distribution are required in determining the resistivity of the subsurface. For two current electrodes A and B as shown in Figure 2.4, the potential at arbitrary point M is given by:

$$U_M = \frac{I\rho}{2\pi}\left[\frac{1}{r_1} - \frac{1}{r_2}\right] \tag{2.6}$$

where r_1 is the distance between M and A and r_2 is the distance between M and B. The potential between M and N is obtained by subtracting the potential at point N from that at point M and is denoted by ΔU. It is given by:

$$\Delta U = \frac{I\rho}{2\pi}\left[\frac{1}{r_1} - \frac{1}{r_2} - \frac{1}{r_3} + \frac{1}{r_4}\right] = \frac{I\rho}{K} \tag{2.7}$$

where r_3 is the distance between N and A and r_4 the distance between N and B. Since *K* only contains distances between electrodes. Then, on re-arranging we have:

$$\rho = K\frac{\Delta U}{I} \tag{2.8}$$

For an inhomogeneous earth, equation 2.8 will produce values that vary according to the geometrical arrangement of electrodes on the surface. This is called the resistivity is called apparent resistivity ρ_a. In Figure 2.5 a range of resistivity values for many earth materials available is displayed.

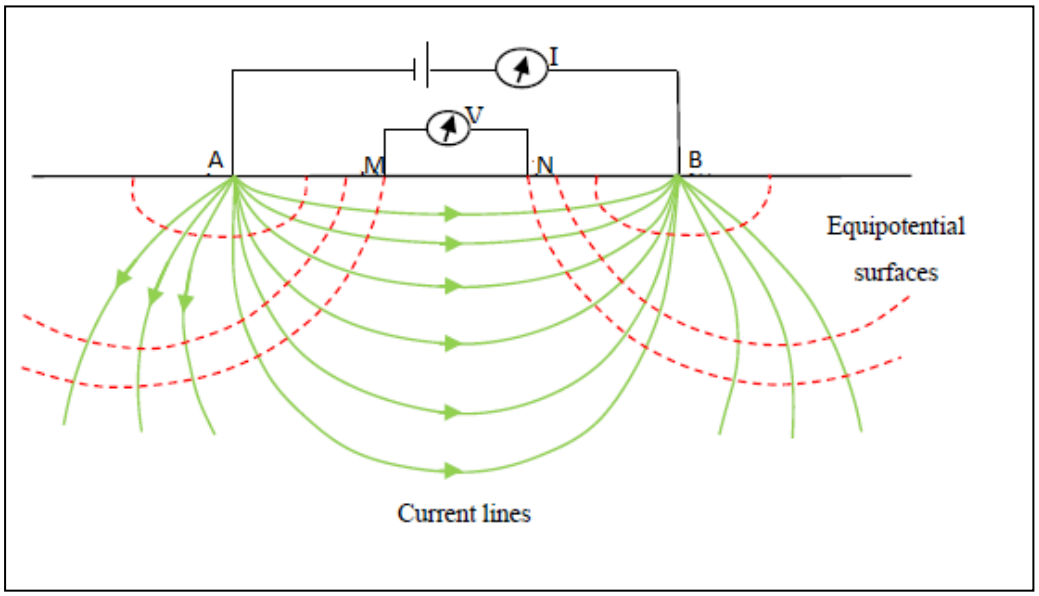

Figure 2.4 Distribution of current and potential lines for two current electrodes at the surface of a homogenous half-space (modified after Van Nostrand and Cook, 1966)