ESTIMATION OF NEAR SURFACE SOILS' POROSITY USING RESISTIVITY IMAGING DATA

NAJMIAH BINTI ROSLI

UNIVERSITI SAINS MALAYSIA

2020

ESTIMATION OF NEAR SURFACE SOILS' POROSITY USING RESISTIVITY IMAGING DATA

by

NAJMIAH BINTI ROSLI

Thesis submitted in fulfilment of the requirements
for the degree of
Doctor of Philosophy

July 2020

ACKNOWLEDGEMENT

All thanks and praises are due to Allah (S.W.T.). This research work is not a solo endeavour as many people have contributed to it in many ways. My main supervisor, Dr. Nur Azwin Ismail, who has given me endless knowledge and helped refine this thesis to be at its best. With her intellectual guidance, she is the best mentor one could ask for.

The highest gratitude goes to my family; Professor Dr. Rosli Saad, Hartini Abd Hamid, Dr. Aishah, Nehlah, Farid Najmi, Athirah and Ahmad Za'im for their constant moral and intellectual support. Even though busy, they took the time to read my thesis and journal papers without asking for anything in return, especially my father and eldest sister. I find it very difficult to thank them enough.

A very special appreciation to my best moral support; Muhammad Nazrin A Rahman. He, too, gave his time to assist me in any ways possible, from idea development, data acquisition until thesis writing. He listened tirelessly to my ideas, while contributing his into the fold. Without his enormous support, this study would be much harder to pursue.

Not forgetting, my co-supervisor Professor Dr. Khiruddin Abdullah, Mr. Yaakob Othman, Mr. Shahil Ahmad Khosaini and the rest of Geophysics staff who have provided much assistance throughout the study. For financial aids, my profound appreciation to my supervisor; Dr. Nur Azwin Ismail, Professor Dr. Rosli Saad and Professor Dato' Dr. Mokhtar Saidin (Director of Centre for Global and Archaeological Research, USM). Thanks also to my colleagues; Mustafa Muhammad Adejo, Rais Yusoh, Mark Jinmin, Kiu Yap Chong, Nabilah Husna and Iffah Zalikha, all

postgraduate students of Geophysics, USM and all those who have contributed to the success of this thesis in whatever form.

This work is dedicated to Allah, the Most Beneficent, the Most Merciful, by Whose grace I become what I am today despite my weaknesses.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

LIST OF SYMBOLS

I	Ammeter
F_a	Apparent resistivity formation factor
ρ_b	Bulk density
ρ_o	Bulk resistivity of saturated soil
$\phi_{calculated}$	Calculated porosity from SPyCRID
m	Cementation exponent
cm	Centimetre
GC	Clayey gravels, clayey sandy gravels
SC	Clayey sands
(SW)	Coarse sand, fine sand
R^2	Coefficient of determination
°	Degree
σ_{eff}	Effective conductivity
e	Exponent
BQv	Effects of surface conductivity
a	Electrodes spacing
χ	Fraction volume
(GM-GP)	Gravel
>	Greater than
≥	Greater than or equal to
”	Inch
CL	Inorganic clays, silty clays, sandy clays of low plasticity
CH	Inorganic clays of high plasticity
MH	Inorganic silts of high plasticity
ML	Inorganic silts, silty or clayey fine sands, with slight plasticity

F_i	Intrinsic resistivity formation factor
<	Less than
≤	Less than or equal to
$\phi_{measured}$	Measured porosity from samples
μ	Mean
m	Metre
µm	Micrometre
µmol/L	Micromole per litre
mA	Milliampere
mm	Millimetre
mL	Millilitre
-	Minus
Z	Missing values
n	Number of data points used from data set
Ωm	Ohm-metre
OL	Organic silts and organic silty clays of low plasticity
%	Percentage
n	Percentage of data points used
+	Plus
±	Plus, or minus
GP	Poorly graded gravel, sandy gravel, with little or no fines
SP	Poorly graded sands, gravelly sands, with little or no fines
ϕ	Porosity
ρ	Resistivity
σ	Sigma in standard deviation
GM	Silty gravels, silty sandy gravels
(CL-OL)	Silty or sandy clay

SM	Silty sands
V_{total}	Total volume
F	Resistivity formation factor
ρ_w	Resistivity of pore-fluid
V_{void}	Void volume
V	Voltmeter
(ML)	Uniform inorganic silt
GW	Well graded gravel, sandy gravel, with little or no fines
SW	Well graded sands, gravelly sands, with little or no fines

LIST OF ABBREVIATIONS

ASTM	American Society for Testing and Materials
CAT	Calibrated automated thrombogram
z-score	Data standardisation via *z*-score
E	East
X	Lower limit of data set
N	North
NE	North-east
NW	North-west
1-D	One-dimensional
PSD	Particle size distribution
PSU	Practical Salinity Unit
Res2Dinv	Resistivity two-dimensional inversion
RMS	Root-mean-square
SPyCRID	Sample-free Porosity Calculation from Resistivity Imaging Data
SAS4000	Signal averaging system 4000
SCS	Soil core sampling
S	South
SE	South-east
SW	South-west
SD	Standard deviation
TG	Thrombin generation
TDS	Total dissolved solids
2-D	Two-dimensional
2-DRI	Two-dimensional Resistivity Imaging

USCS	Unified Soil Classification System
USM	Universiti Sains Malaysia
Y	Upper limit of data set
VES	Vertical electrical sounding
W	West

LIST OF APPENDICES

ANGGARAN KEPOROSAN TANAH DEKAT PERMUKAAN MENGGUNAKAN PENGIMEJAN DATA KEBERINTANGAN

ABSTRAK

Pengimejan keberintangan dua dimensi (2-DRI) merupakan kaedah yang diguna pakai dengan meluas dalam kajian tanah termasuklah dalam anggaran keporosan kerana kepekaannya yang tinggi terhadap perubahan keberintangan elektrik. Keporosan mempunyai pengaruh yang ketara terhadap sifat tanah dan lazimnya diperoleh melalui pensampelan fizikal yang mahal dan menelan masa, maka persamaan Archie adalah kebiasaannya diguna pakai untuk menganggar keporosan sesuatu bahan. Walau bagaimanapun, kebanyakan kajian masih menjalankan pengukuran makmal terhadap sampel-sampel tanah untuk memperoleh nilai-nilai pemboleh ubah Archie seperti faktor penyimenan dan keberintangan bendalir liang sebelum menganggar keporosan untuk kawasan sasaran. Ini menunjukkan bahawa masih tiada kaedah untuk menganggar keporosan dengan tepat tanpa pensampelan fizikal. Kajian ini datang dengan sebuah pendekatan novel (SPyCRID) untuk menganggar keporosan tanah dengan efektif menggunakan data 2-DRI tanpa pensampelan. Dengan memfokuskan hanya kepada tanah peroi, kajian ini menunjukkan penghasilan SPyCRID yang mana tentukurannya telah dijalankan menggunakan dua model untuk mewakili peritus butiran halus yang berbeza dengan keadaan keporosan air tawar dan payau. Pemboleh ubah Archie, iaitu keberintangan bendalir liang dan keberintangan pukal tanah tepu telah diekstrak daripada model songsangan 2-DRI. Dengan menetapkan nilai faktor penyimenan, semua pemboleh ubah Archie sekarang telah dipenuhi dan menjadi input dalam SPyCRID untuk menganggar keporosan tanah setiap model sebelum lelaran data. Dengan mengambil

kira bahawa SPyCRID menjana >20 set data dalam lelaran, kekangan data telah diwujudkan untuk membantu dalam memilih set-set data dengan nilai Archie terbaik yang mewakili tanah tersebut. Kekangan data tersebut adalah berdasarkan kecerunan regresi Waxman-Smits, bilangan titik data yang digunakan, pekali penentuan (R^2) dan sisihan piawai pemboleh ubah. Ketepatan SPyCRID telah disahkan daripada nilai-nilai purata keporosan yang diukur daripada sampel-sampel *in situ* yang tidak terganggu. Ketika mewujudkan jalan kerja dan kekangan data SPyCRID, kaedah tersebut telah diuji ke atas lima tapak ujian dengan tetapan geologi yang berbeza di sekitar kawasan Penang dan Kedah (Malaysia). Keporosan yang dikira oleh SPyCRID di setiap tapak ujian telah disahkan dengan membandingkan dengan nilai-nilai keporosan sampel tanah yang diukur untuk menilai ketepatan dan kebolehpercayaan SPyCRID dalam menganggar keporosan untuk tanah peroi. Hasilan keporosan SPyCRID mempunyai ralat kurang daripada 9.1 % untuk kesemua tapak ujian dan mampu untuk mencapai ralat yang boleh diabaikan iaitu serendah 1.1 %, oleh itu, menunjukkan prestasi cemerlang dengan kadar kejayaan yang tinggi. Ini juga menandakan bahawa SPyCRID adalah efisien dalam tanah peroi walaupun dengan tetapan geologi yang berbeza. Dengan >91.9 % keyakinan kepada SPyCRID, keporosan tanah peroi boleh dianggar secara langsung daripada pengimejan keberintangan 2-dimensi tanpa pensampelan fizikal.

ESTIMATION OF NEAR SURFACE SOILS' POROSITY USING RESISTIVITY IMAGING DATA

ABSTRACT

Two-dimensional resistivity imaging (2-DRI) is a widely employed method in ground studies, which includes porosity estimations due to its high sensitivity to slight electrical resistivity variations. Porosity has significant influence on other ground properties and is conventionally is obtained through physical samplings, which are costly and time consuming; thus, Archie's equation is commonly employed to estimate a material's porosity. However, most studies still conduct laboratory measurements on soil samples to obtain the values for Archie's variables such as cementation exponent and pore-fluid resistivity before calculating porosity for the targeted area. This demonstrates that no method is yet available to accurately estimate porosity without physical samplings. This study comes up with a novel approach (SPyCRID) to effectively estimate porosity of soils using 2-DRI data that is sample-free. Focusing only on unconsolidated soils, this study demonstrates the development of SPyCRID, where its calibrations were conducted using two models to represent different fine grains' percentages with fresh and brackish pore-fluid conditions. Archie's variables; pore-fluid resistivity and bulk resistivity of saturated soil, were extracted from 2-DRI inversion model. With fixed cementation exponent value, all of Archie's variables are now satisfied and became input in SPyCRID to estimate each model's soil porosity prior to data iterations. Considering that SPyCRID generates >20 data sets in the iterations, data constraints were established to assist in selecting data sets with Archie's values that best represents the soil. The data constraints are based on Waxman-Smits' regression gradient, the number of data points used, coefficient of

determination (R^2) and standard deviations of the variables. SPyCRID's accuracy was validated from the averaged porosity values measured from undisturbed in-situ soil samples. Upon establishment of SPyCRID's workflow and data constraints, the method was tested onto five test sites with varying geological settings around Penang and Kedah regions (Malaysia). The estimated porosity by SPyCRID at each test site was validated by comparing to the measured porosity values of soil samples to evaluate SPyCRID's accuracy and reliability in calculating porosity of unconsolidated soils. The resultant SPyCRID's porosity outputs are less than 9.1 % error for all test sites while is able to achieve down to a negligible error of 1.1 %; therefore, showcases an excellent performance with high success rate. This also signifies that SPyCRID is efficient even under varying geological settings of the unconsolidated soils. With >91.9 % confidence on SPyCRID, porosity of unconsolidated soils can now be estimated directly from 2-dimensional resistivity imaging with zero physical samplings.

CHAPTER 1

INTRODUCTION

1.1 Background

Knowledge on ground characteristics is essential in engineering and environmental works such as constructions, slope stability, liquefaction and ground water exploration (Lambe and Whitman, 2008). Two-dimensional resistivity imaging (2-DRI) is a widely employed method to map the ground's resistivity distribution and to distinguish between different ground materials, which are made of varying soil and rock types (Yao *et al.*, 2017). The method's high sensitivity to slight electrical resistivity variations puts it as one of the most powerful tools in engineering and environmental studies. This includes the usage of 2-DRI in porosity estimations of soils and rocks.

Porosity is the voids occupying a material and is one of the principal variables in assessing ground physical properties such as density, strength, friction, and permeability (Lawrence and Jiang, 2017; Lee and Yoon, 2017; Saidian *et al.*, 2016). The variable is also an indicator of effective stress in soils (Henry, 1997). Conventionally, porosity is obtained through physical samplings or pumping tests, which are costly and time consuming. Drilling a single borehole takes weeks for complete installation. Based on quotations by local vendors, borehole drilling costs tens to hundreds of thousands in Malaysian Ringgit depending on type of borehole, ground formation, targeted depth and type of sample collections (Cha *et al.*, 2018; Martínez-Santos *et al.*, 2017). In consideration to their high expenditures, the number of theoretical models to represent the electrical resistivity distribution of the ground for porosity estimations has dramatically increased as new theories and modern technologies were introduced (Cai *et al.*, 2017). Nevertheless, no universal

methodology has yet been formulated that can accurately model the electrical resistivity distribution of heterogenous materials; therefore, limits the effectiveness of porosity estimations (Cai *et al.*, 2017; Kura *et al.*, 2019). Furthermore, the models are also dependent on physical samplings in order to accurately estimate porosity.

Electrical resistivity of saturated porous ground is often studied in a wide field of area; environmental engineering, reservoir study and petrophysics (Mogaji and Lim, 2017). In fact, assessment on ground porosity is commonly conducted using Archie's equation as the equation associates electrical resistivity of the ground to its porosity (ϕ), resistivity of the pore-fluid in its pores (ρ_w), bulk resistivity of the saturated ground (ρ_o) and cementation exponent (m) (Glover, 2009). Archie's equation has set a strong practical foundation for understanding and calculating porosity of saturated ground materials based on the behaviour of electrical resistivity. This leads to the equation's extensive usage in estimating ground porosity in both land and offshore surveys. Such occurrence encouraged researchers to modify Archie's law to induce many Archie's equation derivatives to estimate porosity, although none are completely free of physical samplings to solve their respective equations. In order to tackle this issue, recommendations and gaps in literature were thoroughly reviewed. This study comes up with a novel approach to effectively estimate porosity of shallow unconsolidated soils using 2-DRI data solely with zero dependency on physical samplings.

1.2 Problem statements

In light of porosity's great impact on various ground attributes (permeability, bearing capacity, compressibility and shear strength), evaluation on the variable could assist many ground assessments (Qin *et al.*, 2018; Rashid *et al.*, 2016). Aside from direct measurement of porosity from in-situ physical samples, Archie's equation or its

derivatives is often employed, albeit, the equation's variables also demand physical measurements to be fully functional such seen in studies by Soupios *et al.* (2007), Chukwudi (2011), Li *et al.* (2013) and Glover *et al.* (2000).

By reviewing researches by Cai *et al.* (2017), Byun *et al.* (2019) Niwas and Celik (2012) and Qin *et al.* (2018), it was observed that many Archie's derivatives were formulated to curtail the problem on physical sample dependency. However, no equations nor models are able to be rid of physical samplings completely. Laboratory measurements on samples are still widely conducted to obtain the values for Archie's variables such as m and ρ_w before utilising the measured variables in Archie's equation to calculate porosity. Moreover, the variables' values, which are often measured/collected down boreholes and are one-dimensional, are then extrapolated onto the whole targeted area (Cai *et al.*, 2017; Niwas and Celik, 2012). Such approach is undesirable as extrapolation of the measured values from a singular 1-D location onto a vast area could lead to erroneous porosity calculations, especially if there is a change in ground attributes in the area.

This demonstrates that no method is yet available for practitioners to accurately calculate porosity without being dependent of physical sampling that is expensive and requires tedious laboratory works. This forms the major research gap in this study. Furthermore, Archie's equation requires clay-free and full saturation conditions; therefore, nullifies the law as it introduces substantial errors in porosity calculations for cases where the ground material does not meet the conditions. Different temperatures during measurements of Archie's variables also introduces more errors in porosity estimations (Iravani *et al.*, 2020). These issues are clear indications that there are many potential errors during measurements of Archie's variables prior to

calculating porosity of a ground material that could affect ground evaluations and reservoir estimations.

With regards to the current practice in obtaining porosity of ground materials, which include the reliance on physical samples and erroneous measurements of Archie's variables, this study presents the development of a novel method called Sample-Free Porosity Calculation from Resistivity Imaging Data (SPyCRID). SPyCRID is a method that only requires 2-dimensional resistivity data to estimate porosity using Archie's equation; therefore, resolves the said issues.

1.3 Research objectives

The main outcome of this research is to develop a new method called SPyCRID; an approach that allows practitioners to be free of physical samplings in accurately calculating porosity of unconsolidated soils. Therefore, several objectives were outlined to achieve the primary objective and to test its efficiency, which include;

i. to establish soil characteristics (soil profile, percentage of fine grains and porosity) and type of pore-fluid for SPyCRID's validations,

ii. to develop SPyCRID's workflow for unconsolidated soils without the need of any physical samplings,

iii. to validate SPyCRID's efficiency on varying percentages of fine grains in soil (silt and clay), and

iv. to validate SPyCRID's efficiency on fresh and brackish pore-fluid conditions.

1.4 Scope of study

Focusing only on unconsolidated and saturated soils, this study demonstrates the development of SPyCRID to estimate unconsolidated soil's porosity (calculated porosity) prior to assessing its efficiency. By integrating Archie's and Waxman-Smits' equations, this study utilised only 2-DRI data in SPyCRID. The new method was calibrated using two models of different soil and pore-fluid types to represent different geological conditions. Archie's variables are ρ_w, ρ_o, m and ϕ. By employing a fixed m value, SPyCRID only requires 2-DRI data to fulfil ρ_w and ρ_o values before calculating ϕ of the targeted unconsolidated soil. The 2-DRI array configuration that was considered in the study is Wenner-Schlumberger due to its good vertical and horizontal resolutions. Wenner-Schlumberger's low percentage of error and dense data population also contributes to the reasons for its selection.

In the process of calibrating SPyCRID, several data constraints were established to ensure that SPyCRID effective calculates porosity of the soil. SPyCRID's efficiency in calculating porosity under different geological conditions, which are varying presence of fine grains and pore fluid types (represented by the two models) was also validated from measured porosity values of in-situ soil samples. In consideration to the factors that influence resistivity values, physical samplings on pore-fluid properties and particle size distribution analysis (PSD) of the unconsolidated soil were also conducted. Furthermore, data transformation is also considered as it improves the quality of the data set. Upon the establishment of SPyCRID's workflow and data constraints, the method was tested onto five test sites with varying geological settings around Penang and Kedah regions (Malaysia). The calculated porosity by SPyCRID at each test site was validated from measured porosity

values of in-situ soil samples to ascertain SPyCRID's accuracy and reliability in calculating porosity of unconsolidated soils.

1.5 Research significance and novelty

This research owes its originality to the fact that it is the first of its kind to provide a novel means that fully eliminates physical samplings while maintaining high porosity accuracy. Seeing that current methods in acquiring porosity of soil are still reliant on physical samplings, this study develops SPyCRID to assist practitioners to economise the initially high expenditures of acquiring undisturbed soil samples for porosity measurements, which requires drilling boreholes before proceeding to the taxing laboratory works. Pumping test and wells are also no longer needed to satisfy ρ_w variable, as SPyCRID only requires 2-DRI data.

1.6 Thesis layout

The thesis consists of five chapters which are arranged as follows. Chapter two discusses the literatures concerning porosity and the current methods in obtaining physical information on the variable. Potential ways to address the problem on porosity calculations are also presented here. In the process of reviewing literatures, research gaps were identified to form the basis for this research.

Chapter three presents methods and materials of this research. This chapter is itemised into several subchapters, which thoroughly explains the process of developing SPyCRID for porosity calculations of shallow unconsolidated soils. The geology of each sites is discussed before delving into the study's methodology. The designed methodology is targeted to resolve the delineated objectives; hence, this is also explained. Starting from 2DRI data acquisition, this chapter also discusses on theories and equations used in SPyCRID. Upon establishing SPyCRID, the new

approach's efficiency was tested on several test sites. Along the process of developing and testing SPyCRID's effectiveness, the calculated porosity by SPyCRID for each test site was validated using in-situ samples. The measured values of soil's porosity and pore-fluid's properties are placed under materials' subchapter.

Chapter four marks the most important part of the research as the acquired results are illustrated and discussed here. The performances of SPyCRID during its development under two different geological settings are shown. Upon completion of SPyCRID's development, the method is employed on five other test sites to assess the effectiveness of SPyCRID in terms of percentage of error when compared to measured porosity values from physical samples.

Finally, chapter five concludes the major findings of the research and suggests recommendations for future studies.

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

In consideration to the numerous issues on engineering and environmental that are strong related to porosity, this study came up with an approach to calculate porosity of shallow unconsolidated soils without any physical samplings. In this chapter, the literatures are divided into five main sections; importance of porosity in engineering and environmental works, Archie's law, porosity of ground materials, data transformations and data constraints. The main outcome from the literature reviews gave rise to the research gap that is to be fulfilled by this research. Literatures also provided some ideas to develop SPyCRID and enhance its performance.

2.2 Importance of porosity in engineering and environmental works

Porosity is defined as empty spaces (voids) that occupy soils and rocks, and is often expressed in percentage (Lawrence and Jiang, 2017). It is the principal variable needed to estimate ground water yield and its sustainability (Saidian *et al.*, 2016). As the pores could be filled with air or pore fluid, the variable also plays a pivotal role on bearing capacity of the ground. Highly porous soils frequently have low bearing capacity to support foundations and buildings due to the soils' tendencies to settle via compaction after the structures are erected; therefore, inducing structural failures. Contrarily, non-porous soils such as clay or silt have low permeability for pore-fluid to seep through, which in return, weakens the soils' bearing capacity as the pore-fluid is trapped in the pores with little movement (Sridharan and Choudhury, 2008).

Building constructions face numerous problems during foundation emplacement, especially when the ground is made of soft soil. About one-third of the

world's land is comprised of clayey materials; shales, mudstones, claystones, and siltstones where they are often responsible for land instabilities as a result of their high susceptibility to disintegrate when in contact with moisture (Diaz-Perez *et al.*, 2007; Tan, 2001). Such non-durable characteristic induces slaking; hence, is accountable for many landslides and ground excavation problems. Furthermore, porosity is also one of the variables needed in slope stability equations (Damiano *et al.*, 2017; Wu *et al.*, 2017). For these reasons, engineers always examined the porosity of soils in their works such as constructing dams, bridges, buildings and slope protective measures (Lambe and Whitman, 2008; Rehman and Abouelnaga, 2016).

The importance of porosity can also be highlighted in fresh ground water explorations, now that surface water is decreasing in quality and quantity as a backlash from severe changes in precipitation and temperature (Conti *et al.*, 2016; Prathumratana *et al.*, 2008). Almost half of the readily accessible fresh surface water are already being consumed for agricultural and industrial activities (Whitehead *et al.*, 2009). This condition is further worsen by the fact that many regions are already facing shortages in fresh water supply and a lack in water treatment services (Jury and Vaux, 2005). At this rate, future fresh water sustainability does not look promising if dependency on surface water persists; hence, making fresh ground water to become imperative for sustainability of life (Villholth and Signs, 2019). The main attribute of a sustainable fresh ground water reservoirs is the ability to store large amounts of extractable water by having high porosity (Qin *et al.*, 2018). This emphasises the significance of porosity to efficiently estimate the volume of water trapped in its pores.

Although many studies have come up with equations to accurately model the saturated medium's electrical conductivity systems for porosity estimations, thus far,

none of the equations are fully independent of physical samplings (Cai *et al.*, 2017). They demand certain information that could only be obtained from samplings such as pore-fluid (water occupying the pore spaces of soil/rock) resistivity. Hence, physical samplings from wells, boreholes or core samples are imperative despite often being far from the site location or entirely unavailable in the area (Glover, 2016; Winsauer *et al.*, 1952).

Niwas and Celik (2012) conducted a study on Ruhrtal saturated ground to estimate porosity and hydraulic conductivity using vertical electrical soundings (VES). Archie's equation plays a large role in the study's success as hydraulic conductivity was obtained from porosity calculation using the equation. Kozeny–Archie and Ohm's–Darcy's laws were applied for better approximation of VES values where the values were then used to calculate hydraulic conductivity. By measuring 20 points of the saturated ground bulk resistivity (ρ_o) and pore-fluid resistivity (ρ_w) down available wells, good porosity and hydraulic conductivity values of 11 – 19 % and 0.0076 – 0.19 m/s respectively were successfully determined. With these values, a porosity-hydraulic conductivity model for Ruhrtal's saturated ground was produced to demonstrate that resistivity imaging in conjunction with Archie's equation could be used to calculate porosity with high accuracy. However, the model is dependent on wells for ρ_w values. The measured ρ_w and ρ_o values (from VES) also must have substantially low errors for the model to work accurately. Furthermore, the equation produced is only viable for aquifers with little to none clay materials.

Soupios *et al.* (2007) also did a study using Archie's equation in an attempt to assess and manage shallow sedimentary ground water in Keritis Basin (Greece). Hydraulic properties of the saturated ground were estimated using VES with support

from pumping tests data at available sites. The cementation exponent was estimated from a wide range of literatures for porosity calculation due to lack of core samples. Hydrogeological maps were successfully generated using calculated hydraulic properties from VES (for ρ_o) and pumping test data (for hydraulic properties). The hydraulic properties of the aquifer were extrapolated onto sites that do not have wells. This further proves that Archie's equation is credible in approximating aquifer porosities although this approach is still reliant on physical samplings.

2.3 Archie's law

Archie (1942) has produced one of the most frequently applied theories in estimating porosity of ground materials, which was first introduced for hydrocarbon explorations but was later improvised for ground water explorations. This section demonstrates the derivation of the equation for porosity calculation and the conventional methods in acquiring each variable needed for the equation to operate.

2.3.1 Archie's law and its derivations

Initially adopted from resistivity data log, Archie's law was employed to study the characteristics of saturated granular media. The law associates bulk resistivity of the saturated aquifer (ρ_o) with porosity (ϕ) and pore-fluid resistivity (ρ_w) as shown in Equation 2.1 where m is cementation exponent of the medium. From unconsolidated soils to hard rocks, c varies from 1.2 – 3.5 where the coefficient increases with higher degree of cementation within the material. Intrinsic resistivity formation factor (F_i) is the ratio between ρ_o and ρ_w as portrayed in Equation 2.2.

$$\rho_o = \rho_w \phi^{-m} \tag{2.1}$$

$$F_i = \frac{\rho_o}{\rho_w} \tag{2.2}$$

Worthington (1993), Yadav (2015) and Yue (2019) had pointed out that calculated porosities using Archie's law did not tally with the measured porosity of samples due to the presence of error in F_i for the cases where the material is not clay-free. As a result, the law is rendered invalid as the measured bulk resistivity, ρ_o, is distorted by surface conductivity of clay, which simultaneously causes F_i to become apparent resistivity formation factor, F_a, instead of F_i that is needed in Archie's equation. The F_a, which is the distorted value of F_i, needs to be corrected; therefore, an extra corrective step is added to the Archie's equation in order to compensate for the effects of surface conduction of clay.

Waxman–Smits model (Equation 2.3) was introduced where the model takes into account the effects of surface conductivity, which is denoted by BQ_v (Revil *et al.*, 2017; Vinegar and Waxman, 1984). Equation 2.3 was rearranged to produce a linear relationship between $\frac{1}{F_a}$ and ρ_w (Equation 2.4) where $\frac{1}{F_i}$ represents the y-intercept of the regression's linear plot and BQ_v/F_i is the gradient as depicted in Figure 2.1.

$$F_a = F_i(1 + BQ_v\rho_w)^{-1} \tag{2.3}$$

$$\frac{1}{F_a} = \frac{1}{F_i} + \left(\frac{BQ_v}{F_i}\right)\rho_w \tag{2.4}$$

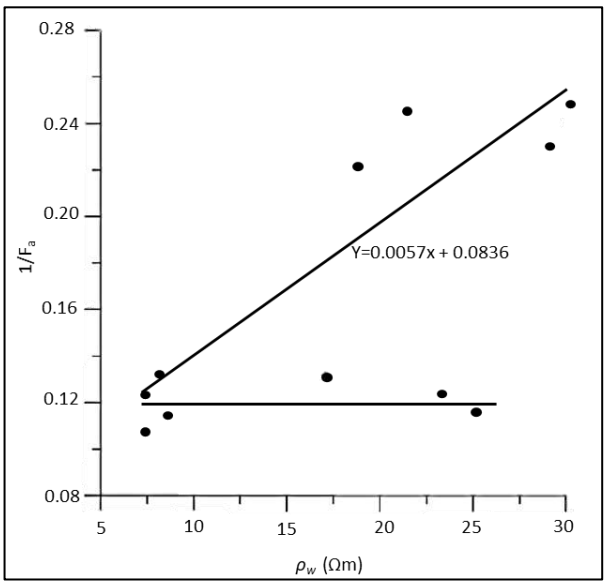

Figure 2.1 Waxman-Smits' regression where y-intercept is the $\frac{1}{F_i}$ (Glover, 2016)

The ρ_o values in the equation are commonly taken from 1-Dimensional (1-D) resistivity inversion whereas ρ_w values are measured at wells. Presence of two fitted lines in Waxman-Smits' regression such as illustrated in Figure 2.1 indicates the presence of two different ground materials (Glover, 2016). After the elimination of surface conduction effects in F_a, porosity can now be calculated using the newly acquired $\frac{1}{F_i}$. By rearranging Equation 2.1 with respect to porosity (Equation 2.5), $\frac{1}{F_i}$ can be directly employed. With strong understanding of the derivation of Archie's equation, conventional methods in acquiring each variable can now be looked into.

$$\phi = e^{\frac{1}{m}\ln\left(\frac{1}{F_i}\right)} \tag{2.5}$$

2.3.2 Saturated bulk resistivity

Resistivity of saturated bulk ground materials (ρ_o) can be easily acquired from VES, 2-Dimensional Resistivity Imaging (2-DRI) or electrical logging. A study by Chukwudi (2011) had demonstrated a good way in approximating porosity of saturated ground in Enugu State (Southeastern Nigeria) via hydraulic properties calculation. By utilizing 322 VES data, porosity of the ground was calculated using hydraulic conductivity, resistivity formation factor, transmissivity and specific yield. The study had successfully mapped the ground's potential for high ground water yield. Ground materials with high transmissivity and specific yield values indicate moderate porosity. This study had also generated regional maps of specific yield and porosity from the combination of VES and pumping tests data where the $\frac{\rho_o}{\rho_w}$ relationship was estimated via linear function. Similar to the work conducted by Soupios *et al.* (2007), the accomplished relationship was then extrapolated onto areas with unavailable wells. The results further proved that investigating aquifers' hydraulic properties using VES

is a viable method, although the technique is still dependent on wells in estimating porosity.

Worthington (2011) had also confirmed the effects of conductive materials during an interpretation of an electrical log recognized by Patnode and Wyllie (1950). Based on retrieved cores and slurries, the presence of conductive wet clay in the cores caused the calculated F_a to be less than F_i. The F_i is defined as values that are not influenced by conductive materials, unlike F_a. Therefore, F_i is a constant value for a particular material but F_a value changes when there is a change in ρ_w or presence of fine materials. This explains the occurrence of $F_a<F_i$ as F_a values decline with increasing ρ_w values that is saturating the ground material. It was also proven that the overall core ρ_o measured are a mixture of resistivity influenced by F_i, ρ_w and surface conductivity of fine-grained materials' presence. Moreover, it was also observed that salinity of the pore-fluid appears to have a negligible effect on the surface conductivity of fine-grained materials.

Another study by Nicaise *et al.* (2012) had produced ρ_o of saturated soils near coastal zone in the South of Benin for better management of water resources in the region. VES and electromagnetic surveys were performed across 350 m^2 area with support from 11 wells available in the area. The spatial distribution of fresh water/sea water boundary was mapped by measuring directly into the wells. A supplementary result in the form of ρ_o of saturated soils with different pore-fluid salinities was also generated (Table 2.1). In correlation to this study, Loke (2004) also made a list of ρ_o for several soils and pore-fluid types, which was included in Table 2.1 for comparison.

Table 2.1 Soil resistivity values produced by Nicaise *et al.* (2012) and Loke (2004)

Soil types	Bulk resistivity, ρ_a (Ωm)	
	Nicaise *et al.* (2012)	Loke (2004)
Sand saturated with fresh water	50 – 500	-
Sand saturated with salt water	0.5 – 5	-
Clay (Undefined saturation degree and type of saturating pore fluid)	2 – 20	1 – 100
Silty sand	10 – 50	-
Alluvium	-	10 – 900

Meanwhile, Braga *et al.* (2006) had tabulated ρ_o values for unsaturated and saturated unconsolidated soils in a research for aquifer protection by employing resistivity method. A variety of soil types were measured for its ρ_o values, which include clayey, sandy-clay, clayey-sand and sandy soils (Table 2.2). As expected, unsaturated soils comprise of the highest range of ρ_o values (up to 1440 Ωm) as compared to all saturated soils due to less moisture content to facilitate current flow. This is the perfect example to showcase that pore-fluid have the greatest weight on ρ_o values such as stated by Kižlo and Kanbergs (2009). Comparisons between saturated soils show that the lowest ρ_o values are clayey soil type (≤20 Ωm) and smaller presence of fine-grained soil (replaced by sand) causes ρ_o values to increase. This portrays that fine-grained soils reduce the overall ρ_o values of a material due to its surface conduction property, which is a critical factor in Archie's equation as the current flow must not be influenced by this effect.

Table 2.2 Bulk resistivity values of sedimentary aquifers based on correlation between resistivity models and wells (Braga *et al.*, 2006)

Type of soil	Lithology	Bulk resistivity, ρ_o (Ωm)
Unsaturated soil	Uncertain	7 – 1440
Saturated soil	Clay	≤ 20
	Sandy-clay	20 – 40
	Clayey-sand	40 – 60
	Sand	≥ 60

Pandey (2015), Kibria and Hossain (2012) also illustrated the correlation between ρ_o of clayey soils with pore-fluid content where the ρ_o increases with decreasing pore-fluid content. However, the ρ_o values remain approximately constant when pore-fluid content approaches the clay's plastic limit that ranges around 15 – 28 % or at 40 % pore-fluid content. This proves that the effect of pore-fluid is reciprocal to soil's ρ_o values. The ρ_o values are also affected by porosity where ρ_o decreases with increasing porosity as a result of pore-fluid dominating the voids of the material.

Cementation exponent (m) is a variable posed by Archie and is generally acknowledged as the degree of pore-fluid connectivity in a ground material (Kadhim *et al.*, 2013). Archie's equation was the only model which inserts a variable that illustrates pore-fluid connectivity, but unfortunately, the law becomes inapplicable in a material with considerable amount of clay. For this reason, Glover *et al.* (2000) made a few alterations on Archie's equation for reservoirs that consist of two conductive materials where the modified model could be used for any ground materials and pore-fluid conductivities (σ). The model added a new exponent expressing the material's connectivity (p) while retaining the original cementation exponent (m) that expresses the pore-fluid connectivity. The equation links both the exponents by adopting fraction volume (χ) of each of the conductive materials such as illustrated in Equation 2.6 –

2.9. The term σ_{eff} represents the effective conductivity of the material with *N* different phases where each conductivity has its individual fraction volume.

$$\sigma_{eff} = \sigma_1 \chi_1^p + \sigma_2 \chi_2^m \tag{2.6}$$

$$1 = \chi_1^p + \chi_2^m \tag{2.7}$$

$$p = \frac{\log(1-\chi_2^m)}{\log(1-\chi_2)} \tag{2.8}$$

$$\sigma_{eff} = \sigma_1 (1-\chi_2)^{\left(\frac{\log(1-\chi_2^m)}{\log(1-\chi_2)}\right)} + \sigma_2 \chi_2^m \tag{2.9}$$

The conventional and modified Archie's formula were tested on 10 granulated porous media having different degrees of porosities and conductive materials that were saturated with pore-fluid of varying ρ_w due to different salinities. The modified model fits the measured ρ_o of the saturated material exceedingly well, which was a significant improvement to the original Archie's law and simultaneously verifies the correlation between *p* and *m*. Table 2.3 shows the values of ρ_o and ρ_w obtained from the modified model. The modified model has significantly improved the conventional Archie's formula for conductive materials (clay-rich) while maintaining the performance for analysis of clean media. However, the produced equations are still reliant on information attained through physical samplings; thus, the target to be independent of costly physical samplings has yet to be resolved.

Table 2.3 Values of Archie's variables from the modified model of Glover *et al.* (2000)

Phase in model	Resistivity (Ωm)
Phase 1 (saturated Earth material) resistivity, ρ_o	32.5
Phase 2 (fluid) resistivity	0.07 – 32.5

Another important information that was found is that ρ_o for saturated sandy soil becomes insensitive when the pore-fluid content is high. According to a study by Pandey (2015), ρ_o values decrease as pore-fluid content increases at a high rate before declining in rate when approaching saturated condition. This suggests that at saturated condition, a medium has similar ρ_o values with narrow ρ_o range such as stated by Hayley *et al.* (2010) and (Iravani *et al.*, 2020) whom considers electrical flow in soils and rocks are greatly dependent on degree of pore-fluid saturation and ionic concentration in the pore-fluid.

2.3.3 Pore-fluid resistivity

Pore-fluid resistivity (ρ_w) is one of the most important variables in Archie's equation but is often misunderstood, which subsequently causes erroneous measurement of the variable. Many researchers assumed the resistivity of bulk pore-fluid sample collected from a medium to be similar to the resistivity of the pore-fluid in the pores; hence, the former values are frequently used in their calculations using Archie's equation (Walker *et al.*, 2014). However, this introduces a major error in the calculation as the collected pore-fluid samples no longer have rock-water interaction. Dissolution and precipitation of ground materials in the pore-fluid continuously occur until they reach physicochemical equilibrium. Boise sandstone samples of insignificant clay presence show that the pore-fluid pH can lower ρ_w as far as 100 % than the bulk pore-fluid when it has been in interaction with lithologies. This reaction is even more prominent in low-salinity bulk pore-fluid with < 100 Ωm where the ρ_w could decrease by 100 % after achieving equilibrium with the ground. Similarly, in high salinity bulk pore-fluid, a reduction of salinity and ρ_w upon equilibrium could also occur as a consequence of salt precipitation out of the pore-fluid or dissolution into the pore-fluid. These experiments clearly show that the typical measurements of ρ_w carry

serious systematic error; therefore, an alternative method for ρ_w acquisition needs to be devised to counter this problem.

Strong understanding of electrical current flow's behaviour is also crucial. Glover (2009) states that for hydrocarbon reserves estimations using Archie's law, F_i increases ($F \rightarrow \infty$) with decreasing porosity ($\phi \rightarrow 0$) as a result of diminishing pore-fluid presence in non-porous materials. It is extremely rare for F_i to be less than one, which only occurs when the material is less resistive than pore-fluid. In addition, Corwin and Lesch (2005) states that there are three possible pathways for current to flow in soil/rock, which are through pore-fluid in large pores, solid–pore-fluid phase pathway (through exchangeable cations between clay and pore fluid) or via solid pathway where the soil particles are in direct contact with each other (Figure 2.2). Considering that ground material usually acts as insulator, the measured ρ_o of saturated soils are mainly dependent on the saturated pores via fluid conduction mechanism (Byun *et al.*, 2019; Hersir and Árnason, 2009). For ρ_o values of ≤2 Ωm, the electrical flow is dominated by pore-fluid while ρ_o values of >2 Ωm are starting to be affected by temperature, ground material and surface conduction of fine-grained soils (Flóvenz *et al.*, 1985; Hersir, 2013; Hogg *et al.*, 2017). Nevertheless, Archie's law is still a reasonably good estimation when the saturating pore-fluid predominates the current flow (Árnason *et al.*, 2000).

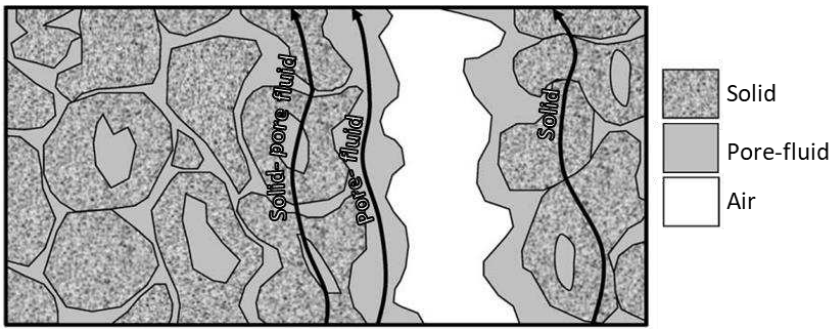

Figure 2.2 Possible pathways of current flow in saturated ground (Corwin and Lesch, 2005)

Another systematic error that is frequently made by researchers is temperature error where ρ_w undergoes changes of roughly 2.3 % per °C before reaching 100 °C. An empirical model was proposed to predict the ρ_w of sodium chloride (NaCl) solution with varying temperatures and salinities (Hayley *et al.*, 2010; Sen and Goode, 1992). This work proved that fluctuating temperatures during ρ_w and ρ_o measurements will introduce errors of -12.03 % to 12.34 % with substantial errors occurring at high salinities. Using these erroneous values in Archie's equation without corrections beforehand will lead to miscalculations of porosity and yield estimations. To avoid this, ρ_w needs to be measured immediately after emerging from the ground while concurrently taking ρ_o reading of the ground to ensure both variables are measured at the same temperature. However, this involves tedious work in laboratories using specialized equipment to circulate pore-fluid in the soil/rock sample or wells and boreholes. Another way in obtaining ρ_w is by measuring the variable down an existing well/borehole (Nicaise *et al.*, 2012), but not all sites have them at disposal. Furthermore, the well/borehole is often far from the site; thus, rendering the measured ρ_w values to misrepresent the actual ρ_w on site.

Ravindran *et al.* (2013) produced a few ranges of ρ_w and total dissolved solids (TDS) values in delineating sea water and fresh water boundary in coastal ground water study at Tamilnadu using 2-DRI. Similarly, Zein *et al.* (2005) also tabulated ρ_w of different water types; fresh water, brackish and saline water along the coastal plain of North Kedah and Perlis (Table 2.4). Using Wenner and Wenner-Schlumberger electrodes configurations for 2-DRI surveys, the transition boundary between saline and fresh ground water was mapped in order to identify consumable fresh water aquifer zone. With classification on the range of ρ_w as a function of pore-fluid types, this tremendously facilitates the prediction of the type of saturated ground in an area.

Table 2.4 Resistivity values of pore-fluid (Ravindran *et al.*, 2013; Zein *et al.*, 2005)

Pore-fluid type	TDS (mg/L)	Pore-fluid resistivity, ρ_w (Ωm)	
		Ravindran *et al.* (2013)	Zein *et al.* (2005)
Fresh	100 – 400	25 – 200	10 – 25
Brackish	400 – 7500	7 – 25	5 – 10
Saline	7500 – 30000	1 – 7	< 5

Lastly, Loke (2004) had created a detailed guideline on electrical resistivity method alongside a software (RES2Dinv) that are being actively used globally for environmental and engineering projects. With this guideline, two types of pore fluid's ρ_w values; sea water and fresh ground water were established. The fresh ground water could be further divided into igneous and sedimentary pore fluids, which was ratified by Keller and Frischknecht (1966). Values of ρ_w by the two researchers can be referred to in Table 2.5.

Table 2.5 Range of pore-fluid resistivities

Type of pore-fluid	Pore-fluid resistivity, ρ_w (Ωm)	
	Keller and Frischknecht (1966)	Loke (2004)
Ground water in igneous rock	30 – 150	5 – 900
Ground water in sedimentary rock	> 1	
Sea water	≈ 0.2	0.2

Additionally, a study by Worthington (2011) pore fluid salinity pore-fluid appears to have a negligible effect on the surface conductivity of fine-grained materials. Brindt *et al.* (2019) have shown that pore fluid salinity causes little (if any) influence on electrical flow as electrical flow is preferential towards pore-fluid content. Based on these reports, brackish pore fluid causes electric current to be more prominently influenced by pore-fluid, instead of the soil grains. However, effect of salinity in calculating porosity using Archie's equation could be tested.

2.3.4 Cementation exponent

Cementation exponent (m) illustrates the degree of compaction in a ground material, where it is influenced by mineralogy, pore structure, specific surface area, pore geometry, degree of saturation, reservoir pressure and temperature conditions (Shahi *et al.*, 2018). Although the c ranges from 1 to infinity, but the ceiling is typically around 5 (Glover, 2009). The variable increases in value as the degree of pore connectedness diminishes.

Byun *et al.* (2019) proved that for unconsolidated soils at 100 % saturation, m ranges from 1.12 – 1.25 but increases to 1.58 – 171 when consolidated as a result of higher density due to compaction. Full saturation could only be assumed when the targeted layer is below water table. The m values produced by researchers for unconsolidated soils are tabulated in Table 2.6 where 1.2 is the average c value.

Table 2.6 Cementation exponent value for various unconsolidated soil types

Unconsolidated material	Cementation exponent, m	Researchers
Silty sand	1.3	Brindt *et al.* (2019)
Glass beads	1.12 – 1.25	Byun *et al.* (2019)
Coastal sands	1.08 – 1.29	Pinas and Acosta (2017)
Clayey sand	1.33	Pinas and Acosta (2017)
Loose sand	1.19	Li *et al.* (2013)
Soft clay	$m = 1.19exp^{\frac{0.033}{\rho_w}}$	Kelly *et al.* (2016)

2.4 Porosity of ground materials

Established data on porosity of soil/rock can serve as a guideline to estimate the variable's values. It is known that effective porosity is difficult to measure as it illustrates the connected pores inside a material. Therefore, most studies approximated effective porosity from total porosity (Wang *et al.*, 2004).

Li *et al.* (2013) conducted a study to calculate degree of saturation of a medium using electrical resistivity method where resistivity formation factor (F) and porosity of samples were obtained from experimental measurement. The ρ_w and ρ_o of each of the lithologies were obtained from the samples. The parameters that were used for Archie's equation include degree of saturation, porosity, resistivities of pore-fluid and bulk in order to calculate m. The porosities from soil/rock samples of loose sand, medium sand and tight sand from various oil and gas environments in China were tabulated as shown in Table 2.7.

Table 2.7 Porosity and cementation exponent of saturated sand (Li *et al.*, 2013)

Type of saturated sand	Porosity, ϕ (%)		Cementation exponent, m		Pore-fluid resistivity, ρ_w (Ωm)	Bulk resistivity ρ_o (Ωm)
	Range	Mean	Range	Mean		
Loose sand	22.6 – 40.4	36.8	1.05 – 1.46	1.19	0.05	0.8
Medium sand	11.0 – 24.6	17.0	1.6 – 2.1	1.84	0.03	8.7
Tight sand	7.5 – 13.3	10.4	1.79 – 2.0	1.89	0.05	40.83

Geotechdata.info (2013) had conveniently produced a compilation of porosity values for various soil types from published literatures including by Das (2013), Swiss Standard (1999) and Hough (1969). The compilation provides geotechnical engineers a comprehensive and high-quality geotechnical information for scientific and practicing purposes. Various information on geotechnical properties of soils and rocks could be retrieved from the database such as void ratio, permeability coefficient, bearing capacity, angle of friction and porosity of soils. The porosity compilation data is a representative of a wide range of USCS soil types at normally consolidated state such as depicted in Table 2.8.

Table 2.8 Porosity range for various soil types obtained from Geotechdata.info (2013)

USCS Soil type	USCS group symbol	Porosity (%)		Reference
		Min	Max	
Well graded gravel, sandy gravel, with little or no fines	GW	21	32	Swiss Standard (1999)
Poorly graded gravel, sandy gravel, with little or no fines	GP	21	32	
Silty gravels, silty sandy gravels	GM	15	22	
Gravel	(GW-GP)	23	38	Das (2013)
Clayey gravels, clayey sandy gravels	GC	17	27	Standard (1999)
Well graded sands, gravelly sands, with little or no fines	SW	22	43	Das (2013); Standard (1999)
Coarse sand, fine sand	(SW)	25	49	Das (2013)
Poorly graded sands, gravelly sands, with little or no fines	SP	23	43	Das (2013); Swiss Standard (1999)
Silty sands	SM	25	49	
Clayey sands	SC	15	37	Das (2013)
Inorganic silts, silty or clayey fine sands, with slight plasticity	ML	21	56	
Uniform inorganic silt	(ML)	29	52	Hough (1969)
Inorganic clays, silty clays, sandy clays of low plasticity	CL	29	41	Swiss Standard (1999)
Organic silts and organic silty clays of low plasticity	OL	42	68	Hough (1969); Swiss Standard (1999)
Silty or sandy clay	(CL-OL)	20	64	Hough (1969)
Inorganic silts of high plasticity	MH	53	68	Swiss Standard (1999)
Inorganic clays of high plasticity	CH	39	59	

2.5 Data transformations

Data clustering and data transformation are essential steps in data mining. Data clustering is often exploited to enhance the quality of a data set where the most common algorithm used is iterative refinement technique (Gupta and Chen, 2011). Distinguished algorithms include Lloyd's algorithm technique where the technique randomly assigns a cluster to the data set prior to the update step (re-assignment step), thus making the initial mean of the cluster to be from the randomly assigned data points