GEOTHERMAL RESERVOIR EVALUATION BASED ON INTEGRATED WELL TESTING AND RESISTIVITY DATA USING ARTIFICIAL INTELLIGENCE

SOLOMON WANGILA NAMASWA

DOCTOR OF PHILOSOPHY

(Physics)

JOMO KENYATTA UNIVERSITY

OF

AGRICULTURE AND TECHNOLOGY

2022

Geothermal Reservoir Evaluation Based On Integrated Well Testing and Resistivity Data Using Artificial Intelligence

Solomon Wangila Namaswa

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Physics of the Jomo Kenyatta University of Agriculture and Technology

2022

DECLARATION

This thesis is my original work and has not been presented for a degree in any other university.

Signature: _____ Date: _____

Solomon Wangila Namaswa

This thesis has been submitted for examination with our approval as the university supervisors.

Signature: _____ Date: _____

Prof. John Gitonga Githiri, Ph.D.

JKUAT, Kenya

Signature: _____ Date: _____

Prof. Nicholas Obuuya Mariita, Ph.D.

DeKUT, Kenya

Signature: _____ Date: _____

Dr. Maurice Odondi K'Orowe, Ph.D.

JKUAT, Kenya

DEDICATION

To my mother Juliana Wangila.

ACKNOWLEDGEMENTS

I pass my gratitude to God almighty for the life I have lived especially the good health I had during the entire study of my PhD. To my supervisors Prof. John Gitonga Githiri and Dr. Maurice Odondi K'Orowe of Jomo Kenyatta university of Agriculture and Prof. Nicholas Obuuya Mariita of Dedan Kimathi University of Technology for influencing and instilling in me the desire to produce high quality research work especially their authentic and genuine reviews. I express my most sincere appreciations to the efforts and contribution of Kenya Electricity Generating Company (KenGen) for providing all the geophysical and reservoir data that was used during this research work. The effort of every other person not mentioned and who has contributed in one way or the other to the success of this thesis is also appreciated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ABBREVIATIONS AND ACRONYMS

AdaBoost:	Adaptive Boosting
AI:	Artificial Intelligence
CSV:	Comma Separated Value
DC:	Direct Current
DTR:	Decision Tree Regression/Regressor
EDI:	Electronic Data Interchange
GDC:	Geothermal Development Company
ISOR:	Iceland GeoSurvey
KenGen:	Kenya Electricity Generating Company
lpm:	Litres per Minute
MAE:	Mean Absolute Error
m a s l:	metres above sea level
m b s l:	metres below sea level
ML:	Machine Learning
MT:	Magneto-telluric
MWe:	MegaWatts electric
NaCl:	Sodium Chloride
Na-K:	Sodium-Potassium
NE:	North-East
N-S:	North-South
PIL:	Python Imaging Library
RBF:	Radial Basis Function
R^2:	R-squared
SVM:	Support Vector Machine

SVR: Support Vector Regression

S-W: South-West

TEM: Transient Electromagnetics

3DPM: Data-Driven Discovery Predictive Model

ABSTRACT

In tectonic regions, the flow of fluid within the subsurface is primarily influenced by subsurface temperature, pressure, porosity and permeability. It is therefore necessary to characterize these properties which help in estimating the ultimate productivity of a geothermal reservoir. These properties are used in volumetric calculation of fluids in the reservoir, calculation of fluid saturations and clustering of the reservoir in terms of aquifers, water confining stratum, hydro- thermal zones, lithological horizons, faults and fracture zones. Some of the methods commonly used for estimating these reservoir properties are time consuming and costly. A decisive method for the reservoir estimation is therefore desirable. Geophysical methods including seismology, gravity, magnetics and resistivity have been put in use for geothermal resource mapping at the Olkaria geothermal field for decades. By applying all the necessary geophysical study techniques and data integrated during interpretation, deeper wells producing up to 30 MWe have been drilled. However, despite all the advancements, geophysical integration of multiple datasets is mostly achieved through manual visualization. Machine vision of Artificial Intelligence is therefore desired. Reservoir temperature distribution and the electrical conductivity of rocks mainly depend on permeability, porosity and fluid chemistry. Machine Learning is needed to establish correlation between the temperature distribution and the electrical conductivity of rocks. This research focused on the integration of Olkaria Domes geothermal well testing and geophysical electromagnetic resistivity data. The aim was to establish an alternative estimation method for reservoir temperature through Machine Learning and application of machine vision perceptions and better visualization of images. To achieve these, Data Driven Discovery Predictive Model and multiple image stacking technique using Pivotal Focus Algorithms were built using Python programming language on Anaconda framework. The open-source web-based application Jupyter Notebook for coding and visualization was used. Different Regression models such as Polynomial Regression, Decision Tree Regression, Adaptive Booster Regression, Support Vector Regression and Random Forest Regression were attempted. The performances of the models were compared using R^2 (R-squared) and Mean Absolute Error (MAE). Based on the performance score, best performing model was suggested to predict subsurface temperature from resistivity. From the well recovery results, the Olkaria Domes reservoir can be classified as a convective heat flow system. Two main heat sources were inferred: One to the Northwest and the other to the Eastern side of the field. The two heat sources are separated by a NE-SW trending fault that is believed to control the fluid flow with natural recharge to the reservoir coming from the SW direction. The reservoir had two major feed zones at depths of (900-1300) m a s l and (250-0) m a s l. Step rate injection results indicated that both injectivity index and

transmissivity are higher in the north east regions of the reservoir and decreases towards southwest. The resistivity structure of Olkaria domes at selected depths revealed three main resistivity regions; one low resistivity to the Northwest, the second low resistivity was observed to the Eastern side of the field. The two low resistive regions are separated by a third NE-SW trending high resistive region. Resistivity decreases with depth up to a depth of 500 m a s l then it increases with depth. From the resistivity cross sections, the results reveal three main resistivity zones. The first zone was characterized by a narrow layer of higher resistivity near the surface, likely to represent unaltered region. Underlying this layer was another broader layer of high conductivity that was interpreted to be due to high conductive hydrothermally altered mineralogy such as zeolites and smectites. A relatively higher resistive zone follows whose resistivity may be due to the formation of high temperature mineralogy at depth such as epidote. The image stacking of temperature and resistivity narrowed down Olkaria dome geothermal field into three main regions of interest. The first is located on the Northwest side of the field; the second is on the eastern side of the study area trending north-south. These were high temperature and high conductivity structures regions within the field. The third region is located in the southwest where both temperatures and conductivity were low. This could be the recharge zones where cold fluid is entering the reservoir. Decision Tree Regression (DTR) Machine Learning model was able to learn the trend of resistivity change and the predicted temperature graph matched well with the actual temperature graph. Training the model using the DTR algorithm approach provided superior outputs with the R^2 of 0.81 and MAE of 29.8. DTR being the best algorithm based on the regression model employed was tested then with new data and the R^2 was 0.835 with MAE at 21.7.

CHAPTER ONE

INTRODUCTION

1.0 Background

The earth's subsurface contains large volume of heat bearing formations that has the ability to supply a considerable amount of energy if suitable extraction techniques are applied for exploitation (Hu, 2016). This heat could be used for electricity generation or enjoyed simply as a hot spring. It has been put in use in several parts of the world including Kenya as a green, replenishable and reliable source of energy to curtail the immense dependency on coal, fossils and other conventional sources of energy. It has also reduced public health risks and effects of global warming resulting from the use of conventional sources of energy (Kahlen et al., 2019).

For proper exploitation of this resource, comprehensive and integrated reservoir analyses need to be implemented (Aragón et al., 2019). Integration knowledge is needed in the prediction and mapping of subsurface reservoir properties as a result of recognition of the importance and advantages of integrated reservoir studies (Zhu et al., 2020). Even though data integration benefits have long been recognized, the epistemological barriers that obstruct the flow of knowledge and information between different disciplines have not (Ortiz et al., 2020).

For better determination of reservoir quality and its performance and correct prediction of complex reservoir behaviours, different techniques must be successfully integrated. This is because successful geothermal energy exploitation depends on the availability of a suitable reservoir based sufficient temperature associated with the reservoir, adequate permeability and porosity of the reservoir rocks (Mania, 2017).

The approaches employed in determination some of subsurface properties such as well drilling for subsurface temperature measurements are time-consuming and expensive. Relating and integrating these reservoir properties with the geophysical rock resistivity

can be highly effective. This is because reservoir temperature distribution and the electrical conductivity of rocks mainly depend on the same parameters such as porosity, permeability, tortuosity and pore geometry (Adebayo et al., 2019). The electrical conductivity of reservoir rocks and aquifers can be measured at the surface using geophysical survey methods such as Electromagnetic methods. However subsurface temperature can only be obtained through well logging after drilling the well. Therefore, a method for reliably estimating subsurface temperatures from surface electrical conductivity measurements would be beneficial. This is because geothermal industry operators would wish to have cheap, rapid and dependable techniques by which the underlying properties can be determined. This could be easily achieved by application of knowledge of integration by use of Artificial Intelligence (Chaki et al., 2014). In recent times, Machine Learning (ML) of Artificial Intelligence has been used to solve complicated issues in different scientific fields with the idea of inputting a number of related parameters and use them to envision other output properties (Schmidt et al., 2019).

Different geophysical methods have been put in use for geothermal resource mapping and process validation to help determine the best possible economical exploitation model and minimize reservoir model ambiguity (Chaki et al., 2014). This aids in increasing confidence in siting well location since accuracy of the reservoir characterization relies upon the uncertainty in the available data and also the processes employed during collection, interpretation and assigning reservoir description.

1.1 Study Area

The greater Olkaria geothermal area is situated in the southern Kenyan rift, 127 Km from Nairobi (Figure 1.1). The Kenyan rift constitutes part of the East African rift structure that extends to Mozambique region in the south from Ethiopia (Kandie et al., 2016). It forms a section of the eastern arm that extends to Lake Natron from Lake Turkana, (Omenda, 2010).

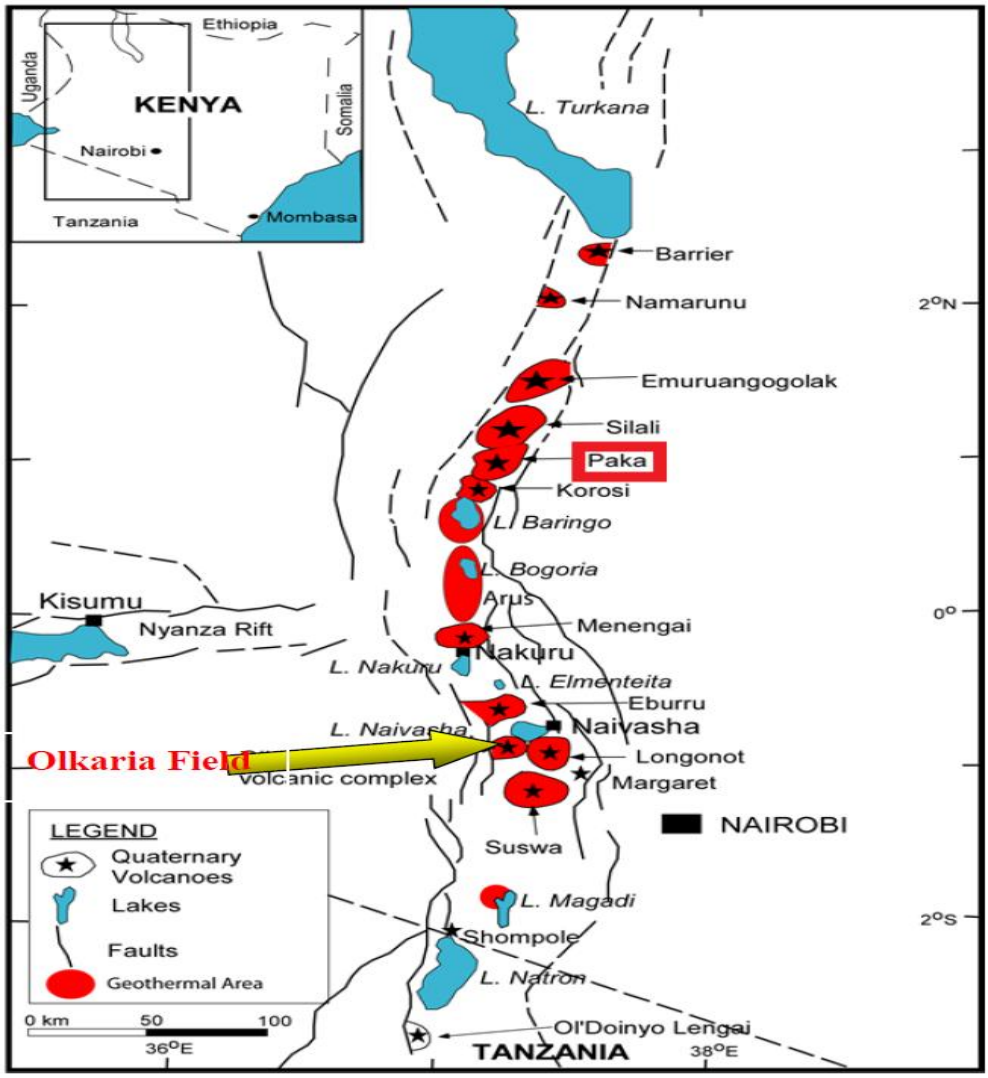


Figure 1.1: Location of Olkaria Field in Kenyan Rift Valley (Wanjohi, 2014)

1.2 Geology of the Study Area

The Olkaria volcanic system (Figure 1.2) is characterized by numerous volcanic centres of Quaternary age and is an exclusive region within the Kenya rift with occurrences of comendite on surface (Nyandigisi, 2020). The Eburru, Suswa volcano and Longonot caldera are the other quaternary volcanic regions neighboring the Olkaria (Okoo et al., 2017). Although calderas are associated to other volcanic centers, an elaborate caldera is not linked to the Olkaria volcanic complex. The existence of a covered caldera theory

has been supported by the visibility of a ring of doming volcanoes in the southwest, south and east (Mwangi et al., 2018). The study seismicity within the Olkaria region has shown an anomalous structure that also agrees with the proposed buried caldera. (Sirma, 2019)

Within the Olkaria, basalt rock formation occupies the Upper Olkaria volcanics to the east of Olkaria Hill sides but they are invisible towards the west ranging from 100 m to 500 m in thickness and are seen as the geothermal system's cap rock (Kibet et al., 2019). From the surface down to a depth of 500 m, comendite lavas and their pyroclastic equivalents and minor trachytes are the dominant rocks (Lagat, 2004). Ololbutot comendite that is tectonically aligned along the ring structure and the N-S faults is the youngest lava (Figure 1.2).

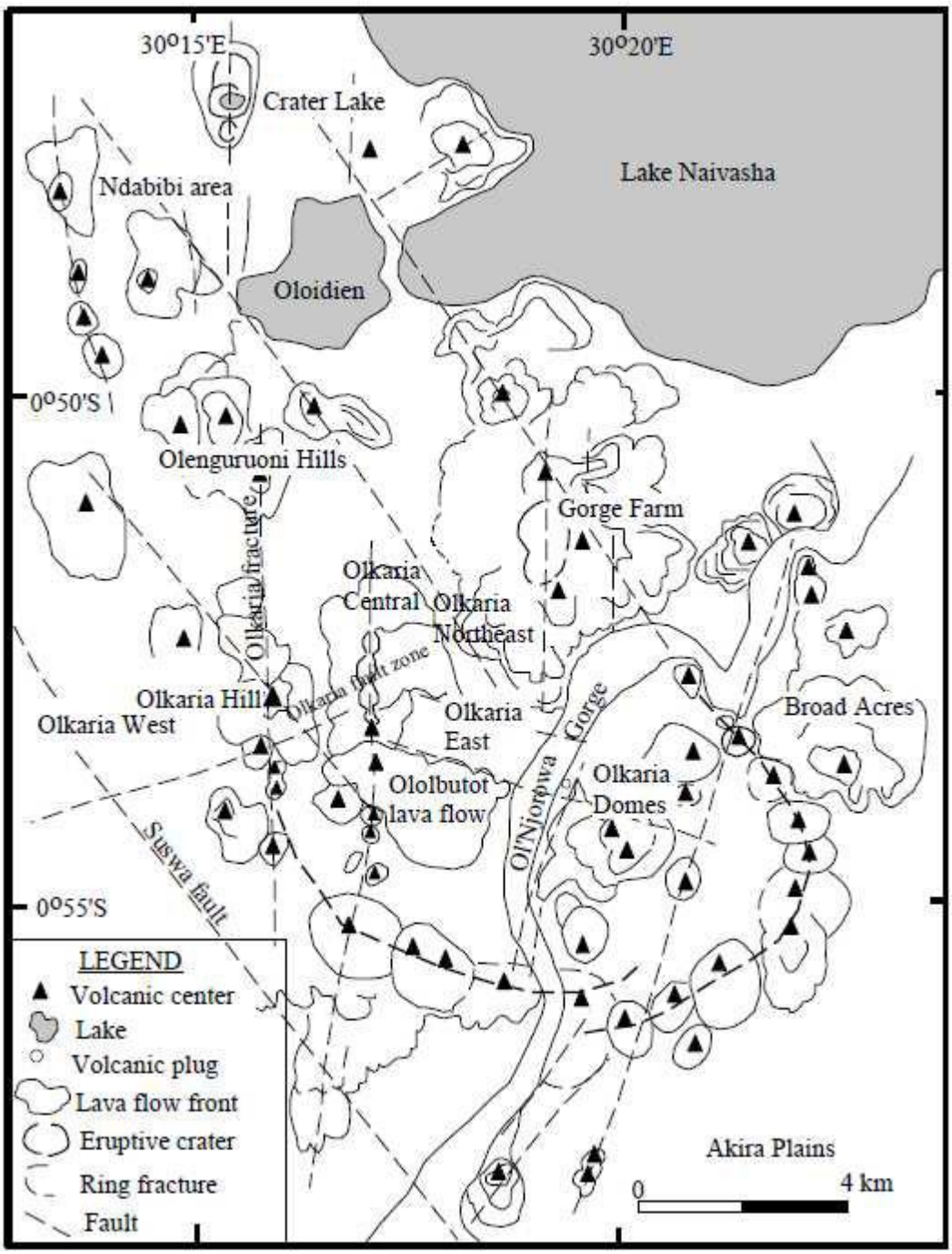


Figure 1.2: Geological Map of Olkaria Volcanic Complex (Okoo et al., 2017)

The NW-SE and WNW-ESE trending faults are more pronounced in the Northeast, East and West of the field with the Gorge Farm fault as the most outstanding as compared to the Olkaria Domes field with scares faults. This could be due to the presence of thick pyroclastics cover in the Domes field. Olkaria Domes is constrained by the Hell‘s Gate – Ol‘Njorowa gorge to the west with swams of dykes exposed trending in a NNE and a ring of domes to the east and south of the area (Mwandigha, 2020).

1.3 Statement of the Research Problem

Geothermal potential is high in regions of active volcanism and high heat-flow such as the East African Rift in Kenya in which Olkaria Domes geothermal field is located. Advancing geothermal energy as a substitute form of heat and electricity usage needs an estimation of the resource and its uncertainty. Improved understanding and visualizing of a geothermal reservoir leads to fewer wells being drilled, fewer well pads and reduced surface disturbance, as well as reduced costs in future exploration and development. For these reasons geophysical methods including seismology, resistivity, gravity, magnetics and Electromagnetics have been put in use for geothermal resource mapping at the Great Olkaria Geothermal field for decades. These methods have aided delineating geothermal fields, siting well locations and locating aquifers (Wanjohi, 2014). They also provide the best ways of establishing at a lower cost the deep-seated subsurface structures as compared to the direct drilling method. Application and integration of different geophysical methods in one field has enabled KenGen to drill good deeper wells producing up to 30 MWe.

However, with Kenya’s increased attention on geothermal development, advanced exploration techniques are desired for the purposes of enhancing general knowledge of geothermal reservoirs, characterize their extent and assess the potential for sustainable utilization. Despite all the advancements, geophysical data processing for different techniques is always done independently and integration of multiple datasets only done through manual visualization. Image processing technique through machine vision of AI

for stacking multidimensional image data sets into one image for perceptions and better visualization is therefore needed.

The adoption of predictive modelling through Machine Learning (ML) in different disciplines has greatly developed in recent times leading to transformation of every field in science, philosophy, psychology and engineering. There have been no attempts to develop a predictive model to determine Olkaria Domes subsurface temperature structures through geophysical Electromagnetic methods.

A model that can result into a cheap, rapid, and dependable method of estimating temperature properties from surface geophysical electrical conductivity measurements through the application of Machine Learning algorithms. The main aim being the desire to establish the unseen pattern behind these multiple data sources through data driven discovery predictive modelling.

1.4 Justification

Destruction of the world ozone layer by pollution from industries and carbon emission from green houses, driven by the demand of energy over the years pushed by the world developing economies has necessitated the need for research for alternative source of energy. Among the leading sources of the alternative green energy production is geothermal energy. Kenya's population is rapidly growing (World Bank, 2016) and therefore the need for economical and dependable energy for provision of fundamental services such as lighting, heating, cooking, mobility and communication as well as spearheading industrial growth is of paramount importance. The need for reliable and sufficient energy source is stipulated in Vision 2030 (blue print) in which the Kenyan government aims at achieving a middle class economy, partly through provision of cheap reliable and clean energy at household and national level. However, there are concerns in the sustainability of the energy resource base in supporting the needs as the high population pressure and industrialization has exerted pressure on the available energy sources.

In a move to avert this effect, the Kenyan government through the ministry of energy and other energy stakeholders have ventured into geothermal energy that is vast, available, but largely underexploited alternative energy source potential for boosting energy supplies in Kenya. This is evidenced in the steadfast production of geothermal energy as shown in Figure 1.3.

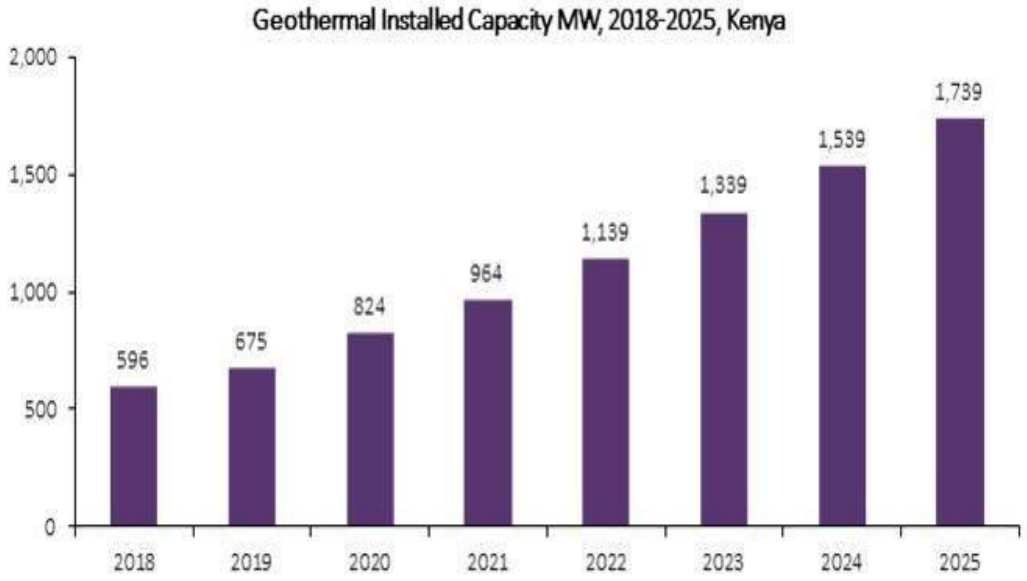


Figure 1.3: Trend of Geothermal Energy Production in Kenya (Power Technology, 2018)

The country's geology and hydrogeology favour economic exploitation of geothermal resources. Kenya being located on the East African Rift, boasts enormous geothermal potential, estimated at 10,000 MW (Mwawasi, 2018). Unfortunately, currently only 861 MW (about 8%) of geothermal energy is installed. This resource, being among the most reliable and inexpensive energy source is therefore foreseen to have a greater impact in generation of alternative energy source for generation of electricity in Kenya. This research was important in improving the knowledge of the geophysical depth structure in geothermal reservoirs in Kenya. The cheaper method of reservoir temperature

determination based on geophysical resistivity was established. Also, application of multiple images stacking shall improve reservoir visualization. From the study, geothermal industry operators in Kenya such as KenGen, GDC and the Kenyan government as well as all geothermal industry stakeholders worldwide who desire cheap, rapid, and dependable techniques by which the underlying subsurface conditions can be determined shall benefit. The new recommendations may also be put in use as a guide for more studies.

1.5 Objectives

1.5.1 Main Objective

To evaluate Olkaria domes geothermal reservoir based on integrated well testing data and electromagnetic data by use of Artificial Intelligence.

1.5.2 Specific Objectives

1. To evaluate Olkaria domes geothermal reservoir sub surface structures using temperature heat up and pressure transient tests
2. To image resistivity structures in high-temperature Olkaria domes geothermal field by use of MT and TEM resistivity methods
3. To map the Olkaria Domes geothermal reservoir by use of multiple image stacking techniques
4. To predict subsurface temperatures from resistivity data in Olkaria Domes geothermal fields using Machine Learning (ML) algorithms

1.6 Research Questions

i. Can geophysical resistivity be used to develop a model to predict subsurface temperature in a high temperature geothermal system?

ii. Can image stacking be used to generate one image obtained from resistivity and temperature contours for better visualization of a geothermal reservoir?

1.7 Scope of the Study

This research focused on application of temperature heat up and pressure transient tests to evaluate reservoir sub surface structures. Resistivity method by use of MT and TEM was used to image resistivity distribution in high-temperature Olkaria domes geothermal field. From the temperature and resistivity contours, image processing was performed to generate an image that was a representation of resistivity and temperature parameters. This aided in identifying areas with high temperature and high conductivity in the image since these are parameters required for high well productivity. Regression algorithms of Machine Learning for predicting subsurface temperature from resistivity data in high temperature geothermal fields was realized by use of Python Programming on Anaconda framework.

1.8 Limitations

Due to processes required in collection of data such as shutting the well and the cost of the equipment, it was not possible to collect primary data. The research used the already acquired electromagnetic, pressure transients and well recovery temperature data to which the accuracy of the data collection cannot be ascertained by the researcher.

CHAPTER TWO

LITERATURE REVIEW

2.0 The Greater Olkaria Geothermal Field

In the early seventies, broad and comprehensive geothermal exploration was carried out after which drilling for exploration purpose began in the year 1974 and extended into 1977 (Ouma 2012). In 1977, a feasibility report was produced based on the assessment of the first drilling outcomes (Nzioka, 2017). Production drilling commenced the following year and extended into 1983. Between 1981 and 1985, the first power plant of 45 MWe was installed after confirmation of satisfactory resource capacity (Nzioka, 2017).

Additional geothermal resource capacity in Olkaria geothermal field was confirmed after further exploration from which the size of the field was estimated to be about 140 km^2 (Omenda et al., 2021). Owing to the expansiveness of the field and for ease of management and development, it was then partitioned into seven zones. These are Olkaria East, Olkaria North East, Olkaria Central, Olkaria North West, Olkaria South West, Olkaria South East and Olkaria Domes (Rogei, 2021) as shown in Figure 2.1.

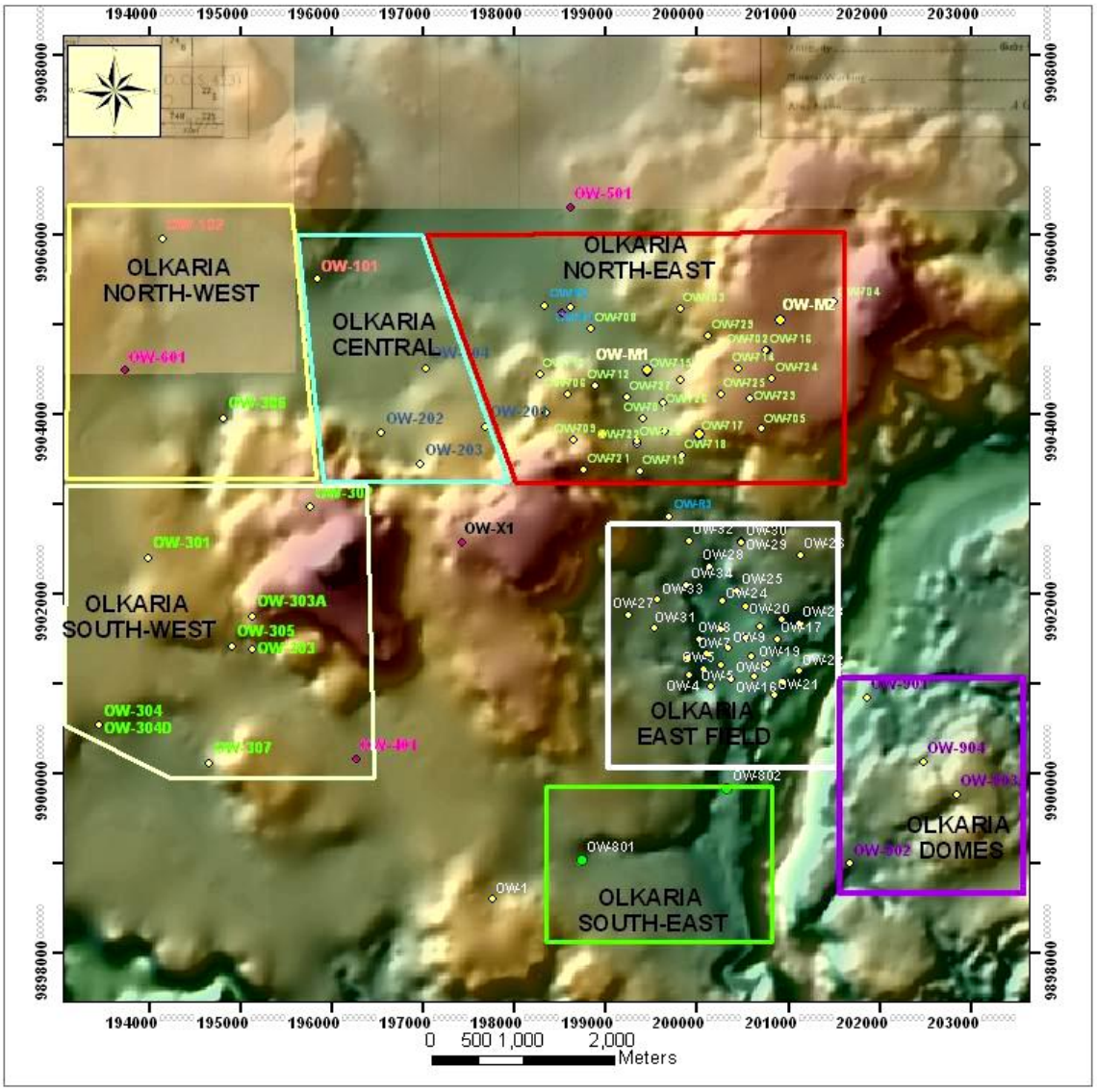


Figure 2.1: Seven Fields of Olkaria Geothermal Area (Okoo et al., 2017)

2.1 Domes Field Development

The high-temperature Olkaria Domes field is based at the SouthEast side of the Greater Olkaria geothermal area with wells producing two-phase fluid. In 1993, surface exploration was completed (Nyandigisi, 2020). Drilling of the first three wells for exploration was conducted between 1998 and 1999 (Warega, 2019). In 2007, six

evaluation wells were drilled and from the results, an updated conceptual model was developed. The model resulted into siting and drilling of the production wells. The Olkaria Domes field is approximately 27 km^2 based on available data with 10 to 15 MW/km^2 power density and between 270 and 405 MW estimated resource capacity (Kandie et al., 2016). KenGen is still drilling and testing wells in the field for the purpose of providing steam for new power plants (Mbithi, 2016).

2.2 Theory of Well Testing

A geothermal energy process from exploration to exploitation is divided into four major series that include surface exploration, well drilling, reservoir analysis and operation stage that entails harvest of heat from the earth's subsurface (Parada, 2016). Dependable knowledge for the reservoir under study is paramount in determination of the best and economical exploitation method. Reservoir properties must be known in order to reach this decision. This is achieved by well testing in combination of other disciplines such as geology, geochemistry and geophysics through development of a comprehensive conceptual model (Mortensen & Axelssion, 2013).

Geothermal reservoir performance predictions in most cases depend on the knowledge of the reservoir's natural conditions established from temperature and pressure measurements. These two parameters are the most crucial parameters required for geothermal resource assessment since they control the movement and distribution of the reservoir fluids (Franco & Vaccaro, 2020). They also form the basis for development of reservoir conceptual models (Jansen & Miller, 2017).

Well testing as a tool in geothermal reservoir engineering plays a role in determination of reservoir quality (Rionomakal et al., 2018). It involves use of pressure and temperature data (Mwaura, 2018). Well testing after well completion and during discharge is conducted so that reservoir properties such as transmissivity, injectivity, storativity and initial reservoir temperature parameters are deduced (Biru, 2016). The reservoir's response of temperature and pressure as a function of injection, warm up or

drawdown are monitored during well testing by use of pressure and temperature gauge in a well.

2.2.1 Analysis of Temperature Profiles

Temperature profiles measured at depth provide crucial data about reservoir physical conditions and are primary measurements performed in wells (Kargarpour, 2017). In order to obtain meaningful results, these measurements should be run over the full depth of the wellbore and are done when the well is flowing, shut in or under injection. This is performed with the aim of inferring reservoir properties. Temperature profile analysis amid well warm up helps in identification of feed zones and initial reservoir temperature (Gebru, 2018).

From the warm up temperature profiles, location that are directly above or close to the main heat resource should project the highest temperature. In these locations, temperature should increase with depth. High values are observed closer to the up-flow zones while zones further away will be expected to be at lower temperatures (Vehling et al., 2020). Thus temperature variation may indicate proximity to the up-flow zones and fluid movement from fluid up flow zones laterally outwards. Zones with consistent high temperature and increasing consistently with depth will define the probable or even possible reservoir.

It is standard practice to use temperature contours drawn at selected elevations to infer directions of fluid movement, generally flowing from high temperature areas outwards. Since temperature can be inverted by cold water inflow, temperature decline could also be due to cold water inflow (Vihma et al., 2011). Also temperature variations can define reservoir limits or boundaries. These zones may have low temperature or may show temperature decline from central locations. It is also possible to have situations where higher temperature caused by adverting hoter fluid that is lying above a cold zone below within a well, referred to as inversion. Decreasing temperature with depth after an initial

increase should indicate this (Zarrouk & McLean, 2019). The inverted temperature profile shown in Figure 2.2 is most likely an out flow zone.

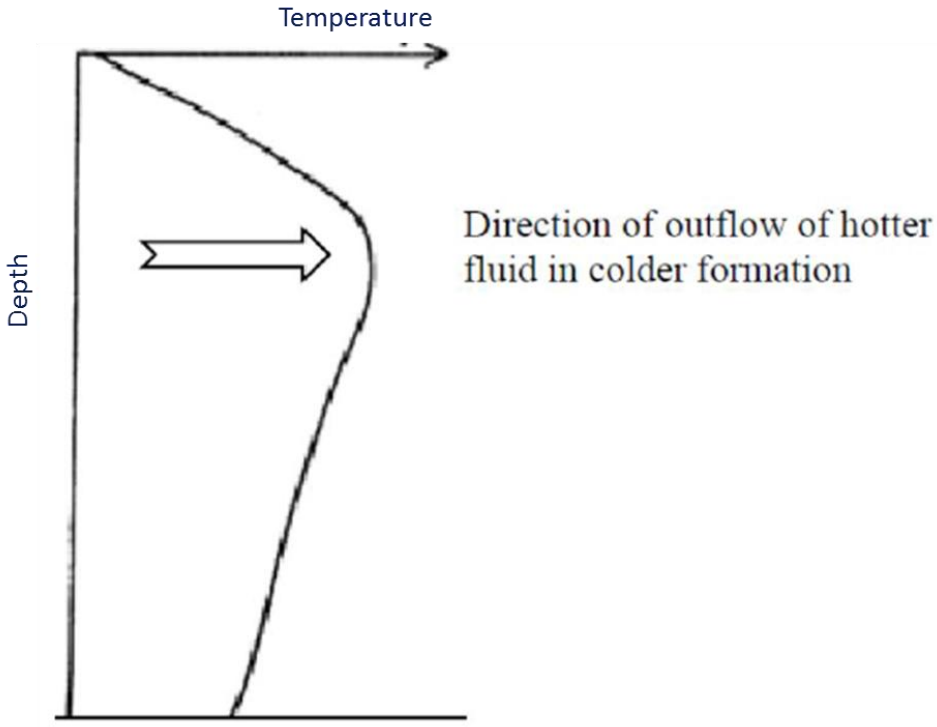


Figure 2. 2: Inverted Temperature Profile (Vihma et al., 2011)

Temperature recovery in a well after cold water injection or after drilling may also offer valuable data on intervals with permeability or feed zones. Such zones will generally accept water during injection or drilling and will start of cooler than less permeable zones where injection or drilling fluid will be passing largely on the surface and therefore will be less affected (Schölderle et al., 2021).

2.2.2 Pressure Transient Well Testing Equation

Pressure transient test is done in a well after completion of drilling to access reservoir conditions with the pressure diffusion equation being applied in obtaining reservoir pressure at a distance (r) from the well having flow rate (q) as a function of time (t) (Biru, 2016). The derivation of this equation is guided by the following:

1. Mass conservation inside a given control volume:

The rate of change of mass inside a control volume = Mass flow in – Mass flow out

Darcy's law of momentum conservation (Tahir, 2019):

$$q = 2\pi rh\frac{k}{\mu}\frac{\partial p}{\partial r} \qquad 2.1$$

2. Equation of the state of the fluid:

$$\rho = \rho(PT) \qquad 2.2$$

3. The fluid compressibility equation:

$$c_f = \frac{1}{\rho}\frac{\partial \rho}{\partial p} \qquad 2.3$$

Combining equations (2.1), (2.2) and (2.3) and defining

$c_t = \phi c_f + (1-\phi)c_r$ as total compressibility of rock and water

$$\frac{1}{r}\frac{\partial}{\partial r}\left\{\frac{r\partial p(r,t)}{\partial r}\right\} = \frac{\mu c_t}{k}\frac{\partial p(r,t)}{\partial t} = \frac{S}{T}\frac{\partial P(r,t)}{\partial t} \qquad 2.4$$

ϕ = porosity

$s = c_t h$ is the storativity

$c_r = \frac{1}{1-\phi}\frac{\partial \phi}{\partial p}$ is the compressibility of porous rock

$T = \frac{kh}{\mu}$ is the transmissivity 2.5

$p(r,t) = p_i$ for $t = 0\ r > 0$ 2.6

An integral solution for this equation was proposed by Theis:

Initial conditions:

$p(r,t) = p_i$ for $t = 0\ r > 0$ 2.7

Boundary conditions:

(i) $p(r,t)$ for $r \to \infty$ and $t > 0$

(ii) $q = 2\pi rh\frac{k}{\mu}\frac{\partial p}{\partial r}$

Equation 2.8 gives solution to the radial diffusion equation having these boundary and initial conditions:

$$p(r,t) = p_i + \frac{q\mu}{4\pi kh}E_i\left(\frac{-\mu c_t r^2}{4kt}\right) \quad 2.8$$

Where

$E_i(-x) = -\int_x^{\infty}\frac{e^{-u}}{u}$ is the exponential integral function

According to Lewis (2014), if

$t > 100\frac{\mu c_t r^2}{4k}$,

The integral function expressed in exponential form can be expended through series of convergent. The total pressure change in equation 2.9 for a pumping well with skin is the Theis solution:

$$\Delta p_t = p_i - p(r,t) = -\frac{2.303q\mu}{4\pi khk}\left\{\log\left(\frac{\mu c_t r^2}{4kt}\right) + \frac{0.5772-2s}{2.303}\right\} \quad 2.9$$

The additional change in pressure to the normal change in pressure in the wells near vicinity is the skin factor, *s*. Plotting for Δ *Pt* vs. log *t* results into a semi-log straight line with a slope m per log cycle in equation 2.10, a method known as semi-log analysis (Kahuda & Pech, 2020). It is an infinite acting radial flow period of a well response.

$$m = \frac{2.303q\mu}{4\pi kh} \text{ (Pa/log cycle)} \quad 2.10$$

The basis of semi-log analysis is the interpretation of the straight-line response of the semi-log that represents the infinite acting radial flow behavior of the well (Spivey et al., 2020). Putting in mind that an actual wellbore has finite volume, it's important to establish wellbore storage effect duration or the time at which the semi-log straight line begins (Spivey et al., 2020). This can be determined on a log (*ΔP*) vs. log (*t*) graph whereby the semi log straight line is expected to start after about 1½ log cycle from the end of the unit slope line (Wang and Sharma, 2018).

Figure 2.3 shows that plotting the change in pressure against time on semi-logarithmic scale results in an asymptotic straight line whose slope is *m*. if the volumetric rate of flow is *q,* the formation transmissivity, *T*, can be calculated as (Ferroud et al., 2019):

$$T = \frac{kh}{\mu} = \frac{2.303q}{4\pi m} \qquad 2.11$$

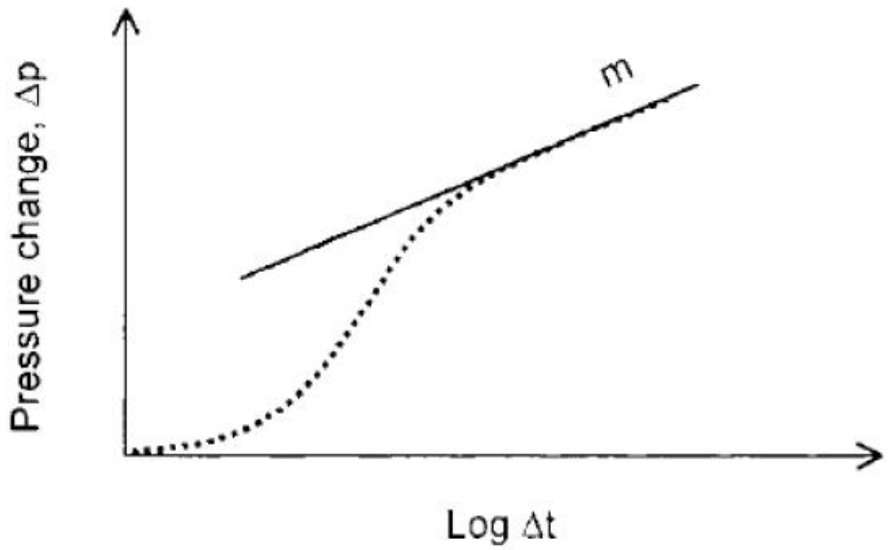


Figure 2.3: Semi-Log Analysis (Gebru, 2018)

Applying drawdown values *Δp* at some time *t, S = cth,* the formation storage coefficient can be established by writing the Theis solution as in equation 2.12 (Naderi, 2019)

$$\frac{\Delta p}{m} = -log\left\{\left(\frac{4kh}{\mu}\right)\frac{1}{c_t h r^2}\frac{t}{r^2}\right\} + 0.251 \qquad 2.12$$

2.3 Resistivity Method

Geophysical method of resistivity is used to show the resistivity changes beneath the earth. This method has found its application in the expedition of geothermal energy since good geothermal reservoir has been associated with low resistivity (Seyedrahimi et al., 2017). Rock electrical conductivity rely on the water salinity, porosity and the pore structure of the rock, water steam content, temperature, alteration, saturation, pressure and interaction between water and the rock (Revil & Gresse, 2021).

The Magneto-telluric and Transient ElectroMagmetic geophysical methods in the earth's subsurface study have contributed immensely to the X-Raying and detailing the buried earth's interior (Jolie et al., 2021). This has been achieved by establishing structures that are deeply seated and are in charge of geothermal system control such as sources of heat and geothermal fluid flow conduits leading to characterizing the potential of the resource and determining the drilling sites. It gives relevant information concerning the resistivity structure of the subsurface earth which when well interpreted may be linked to the presence of the geothermal system (Arthur, 2018). Since they are the best techniques of establishing deep subsurface structures at much lower costs than the most direct method of drilling, geophysical exploration methods play a key role (Wamriew, 2019).

Rocks are regarded as poor conductors and their conductivity can only be achieved if their properties are transformed through alteration by earth processes. Therefore, porosity and fluid conductivity are the major reasons for rock conductivity near earth surfaces (Thanh et al., 2019). These two also depend on the salinity, temperature and pressure as shown in Figure 2.4. Also salinity and temperature varies with depth and regions.

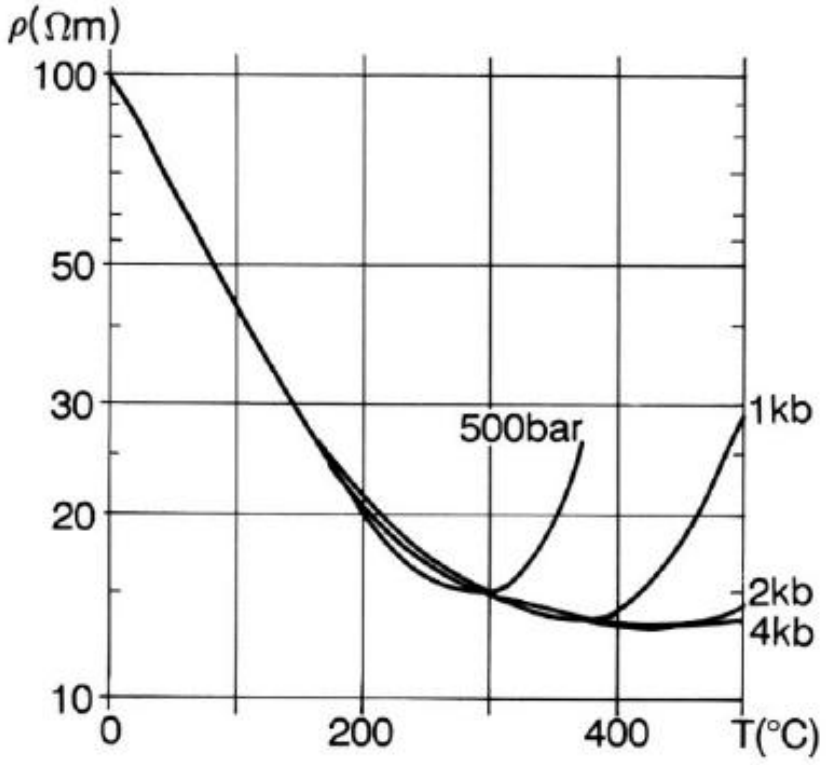


Figure 2.4: Effect of Temperature at Different Pressures on Conductivity of Sodium Chloride Solutions (Yakovlev et al., 2018)

A direct relationship between salinity and conductivity that shows that in a salt solution rock conductivity is by mobile ions is shown in Figure 2.5 (Hu et al., 2022).

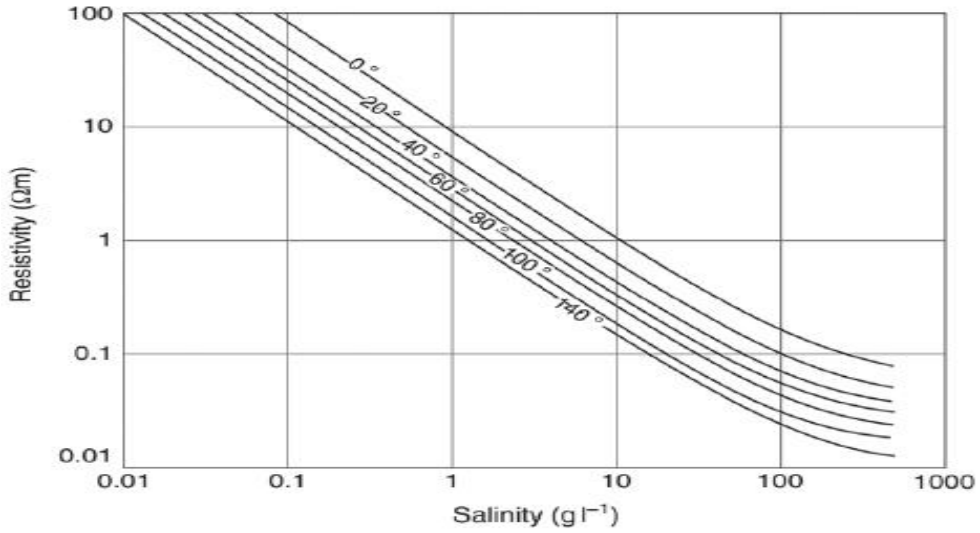


Figure 2.5: Effect of Concentration and Temperature on Conductivity of Sodium Chloride (Flóvenz et al., 2012)

Also, the fluid nature and level of its saturation, alteration intensity and mineralogy influences rock conductivity. Fluids and free ions presence in porous rocks and electrons in minerals at the water-rock interface are the main mechanism of rock conductivity in geothermal system (Manning, 2018). These conduction mechanisms are illustrated in Figure 2.6.

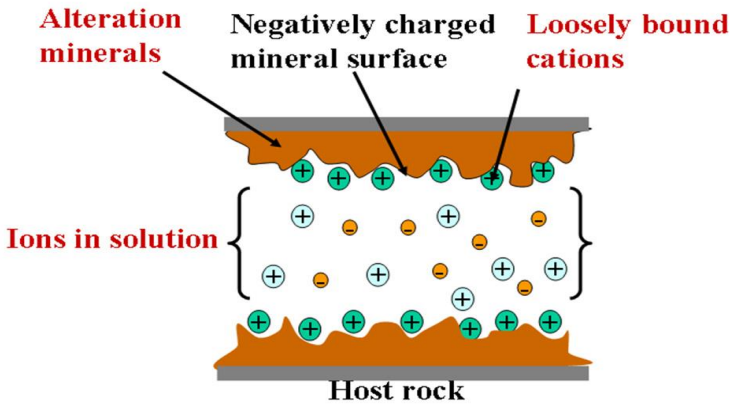


Figure 2.6: Mechanism of Rock Conductivity in Geothermal System (Gylfi & Árnason, 2013)

2.3.1 Mineral Alterations and Resistivity of High Temperature Regions

In hydrothermal systems exhibiting elevated temperatures, rock conductivity varies with depth and temperature. This is because geothermal hot fluids interact with rocks and react to constitute altered minerals at different temperatures (Ojha et al., 2021). In the near subsurface region, the resistivity is high since the temperature is low and therefore the rock formation is un-altered. Pore fluid conduction is the main conduction mechanism in this zone (Mwaura, 2018). As temperature gradually rises with depth, from 100 °C to 220 °C, clay minerals such as smectites and zeolites are formed. This zone is called the smectite-zeolite zone. In this zone, rock resistivity is reduced because of their loosely bound cations (Escobedo et al., 2021). At a fairly higher temperature ranging from 220 °C - 250 °C, resistivity of rocks starts to rise since zeolites fades away and alteration of smectite into chlorite as the prevailing mineral (Weisenberger et al., 2020) . In the region of temperatures beyond 250 °C, resistivity increases again since epidote and chlorite are more prevalent and the cations are constrained in the crystal lattice this is illustrated in Figure 2.7.

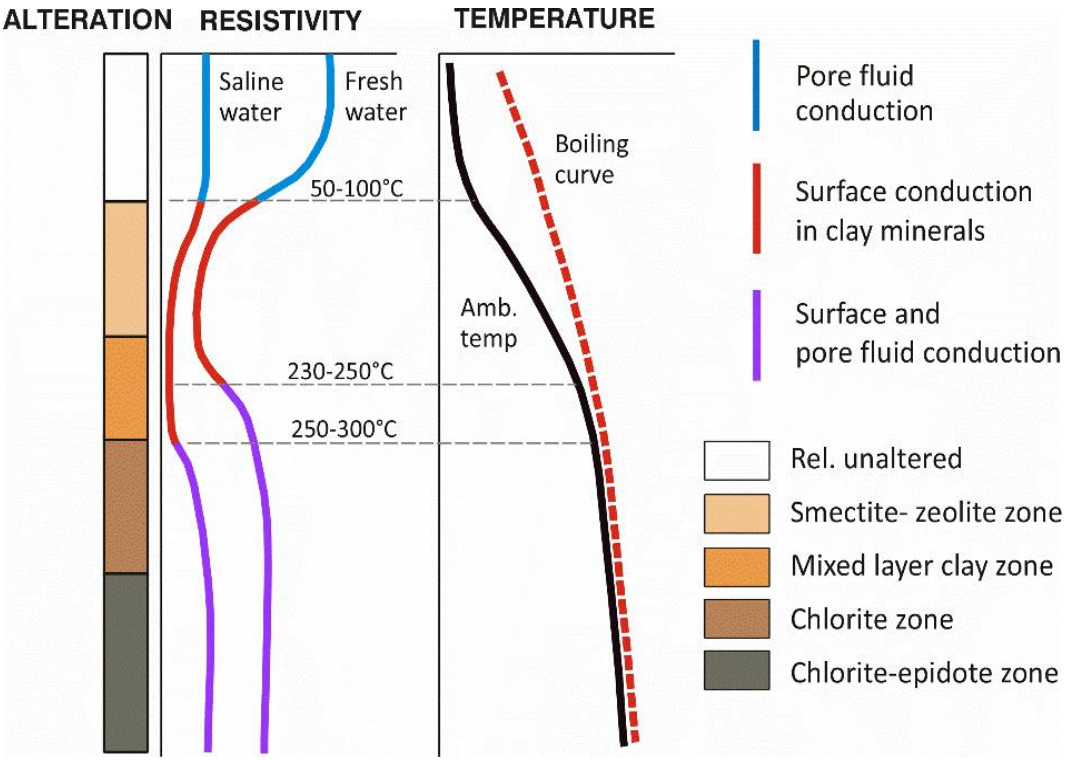


Figure 2.7: High Temperature Geothermal System Resistivity Structure (Gylfi & Árnason, 2013)

A subsurface temperature and mineral alteration relationship analysis is shown in Figure 2.8.

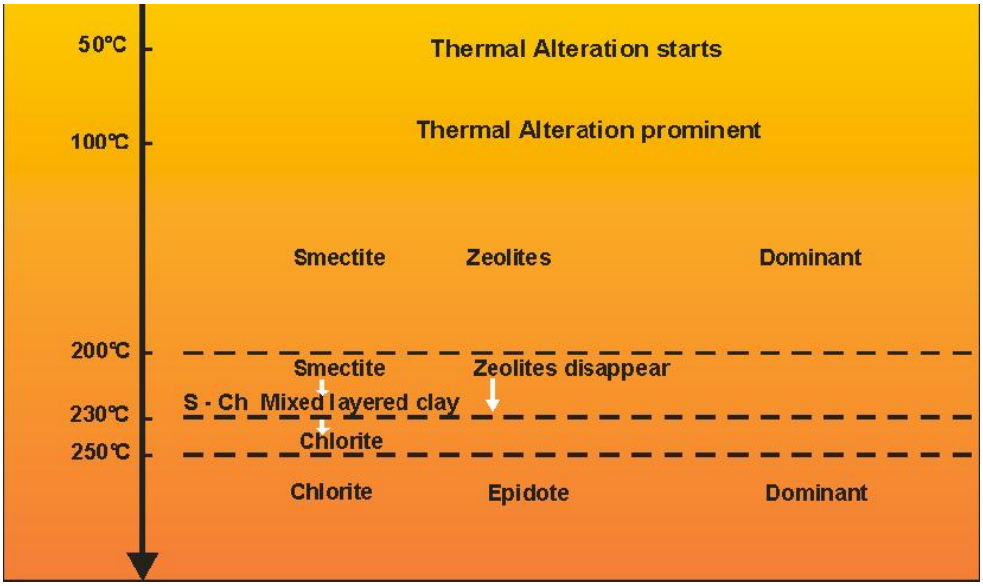


Figure 2.8: Alteration Mineralogy and Temperature (Gylfi & Árnason, 2013).

2.3.2 Principles of the TEM and MT Methods

The principle of operation in the TEM entails the use of a generator or battery pack to transmit a steady current into a large insulated square loop laid on the ground (Figures 2.9 and 2.10) and as a result, a known magnetic field strength is built up (Cai, 2019). In a fraction of a second, the loop current is abruptly turned off and turned on again. In the off current mode, magnetic field developed in the loop decays at the same time inducing in the ground secondary electrical currents (eddy currents) which again induces another secondary magnetic field that deteriorates with time. The distribution of current and the secondary magnetic field rate of decay is dependant to the resistivity subsurface structure of the earth (Mohamud, 2013). By measuring the induced voltage at the transmitter loop's centre in a receiver coil in time, the response is monitored on the ground surface (Nordiana et al., 2016).

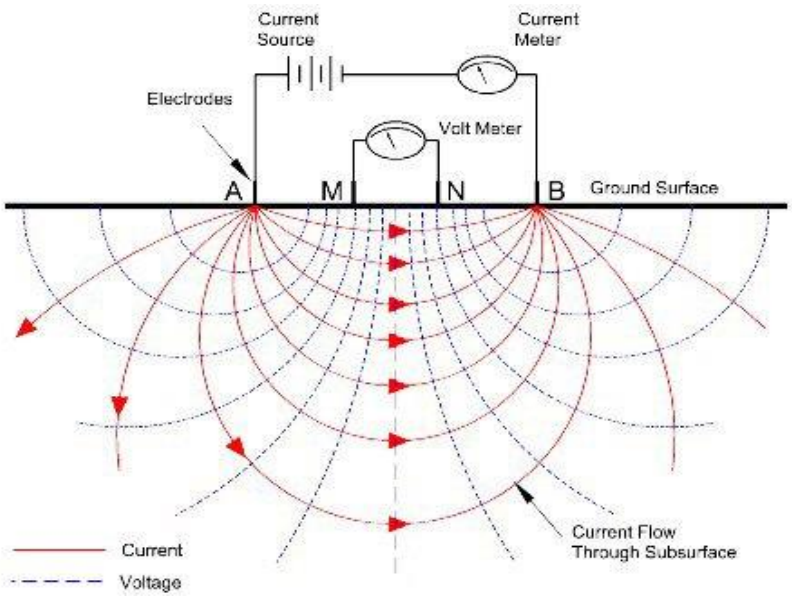


Figure 2.9: Field Layout Schlumberger (Mohamud, 2013)

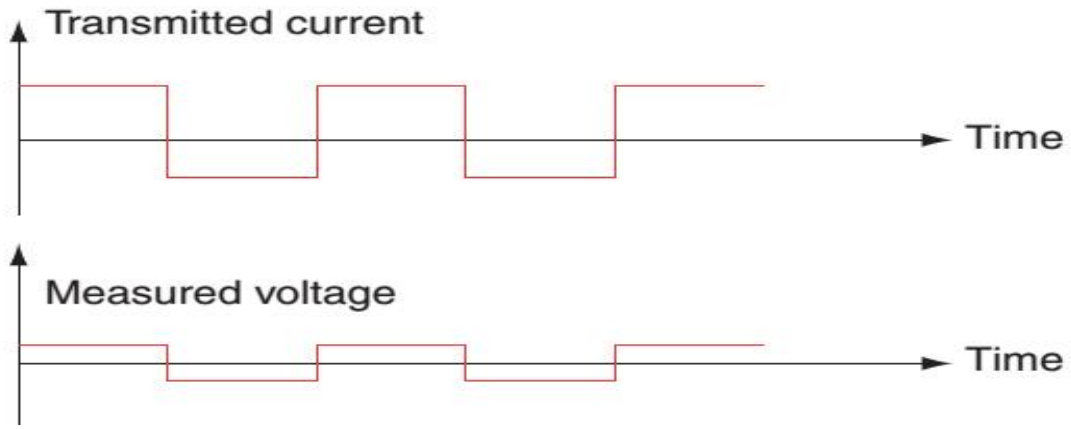


Figure 2.10: 1/2- Duty Current Square Wave and the Analogous Measured Voltage (Mohamud, 2013)

According to Hadush (2018), the depth pervading in the TEM sounding central loop relies on the duration the induction in the coil of the receiver can be traced before the noise drowns it. Wamriew (2019) states that in a homogeneous half space of

conductivity σ , the induced voltage in the coil of receiver at so called late times is given by:

$$V(t,r) \approx I_o \frac{C(\mu_o \sigma r^2)^{\frac{2}{3}}}{10\pi^{\frac{1}{2}} t^{\frac{5}{2}}} \qquad 2.13$$

Where

$$C = A_r n_r A_s n_s \frac{\mu_o}{2\pi r^3}$$

and

A_s= Transmitter loop cross sectional area (m^2)

t = Elapsed time after the transmitter current is turned to zero (s)

A_r= Receiver loop cross sectional area (m^2)

r= transmitter loop radius (m)

n_r= receiver loop number of windings

n_s= transmitter loop number of windings

$V(t,r)$= transient voltage (v)

I_o= transmitter loop current (A)

μ_o= magnetic permeability (H/m)

From equation 2.13, the late times transient voltage is proportional to $\sigma^{3/2}$ and falls off with time as $t^{-5/2}$ after the current in the transmitter loop is abruptly turned off. By solving for the resistivity in Equation 2.13 we get Equation 2.14 as the late time apparent resistivity.

$$\rho_a = \frac{\mu_o}{4\pi}\left[\frac{[2I_o \mu_o A_r n_r A_s n_s]}{5t^{\frac{5}{2}} V(t,r)}\right]^{\frac{2}{3}} \qquad 2.14$$

Magnetotellurics (MT) is an approach that makes use of the naturally occurring Electromagnetic field of the earth imaging the sub surface resistivity structures (Wamriew, 2019) as shown in Figure 2.11. These natural fields are produced by a scope of mechanisms such as high frequency signals that originate from the activity lightening, intermediate frequency signals that originate from the ionospheric resonances and low frequency signals that originate from the sun-spots. The decay rates of these waves as they travel into the Earth's interior depend upon their wavelengths (Sakindi, 2015).

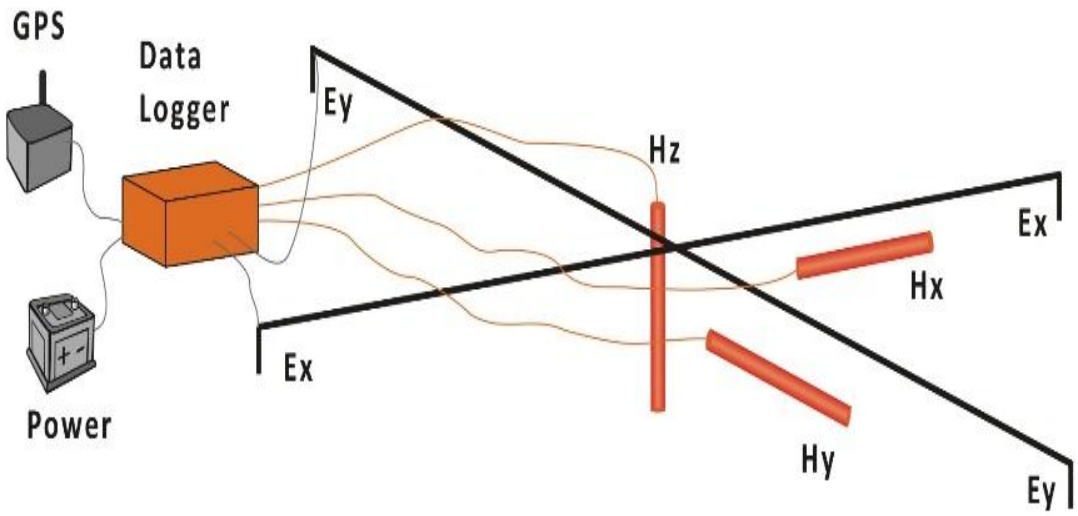


Figure 2.11: MT Field Layout (Mohamud, 2013)

The propagation of EM fields is expressed by the Maxwell's equation:

Faraday's law $$\Delta x \boldsymbol{E} = -\mu \frac{\partial \boldsymbol{H}}{\partial t} \quad (2.15)$$

Ampere's law $$\Delta x \boldsymbol{H} = \boldsymbol{j} + \mathcal{E} \frac{\partial \boldsymbol{E}}{\partial t} \quad (2.16)$$

Gauss's law for the electric field $$div\boldsymbol{D} = \eta \quad (2.17)$$

Gauss's law for the magnetic field $divB = 0$ 2.18

Where E = Electric field (V/m)

H = Magnetic intensity (A/m)

j = Electrical current intensity (A/m^2) and $j = \sigma E$

μ = Magnetic permeability (H/m)

ε = electric permittivity (F/m)

$D = \varepsilon E$

$B = \mu H$

2.3.2.1 Electromagnetic Induction in Homogeneous Earth

The electric to magnetic field intensity ratio is the characteristic impedance, (Equations 2.19 and 2.20)

$$Z_{xy} = \frac{E_x}{H_y} = \frac{i\omega\mu}{k} \quad 2.19$$

$$Z_{yx} = \frac{E_y}{H_x} = -\frac{i\omega\mu}{k} \quad 2.20$$

and

$$Z_{xy} = \frac{E_x}{H_y} = \frac{i\omega\mu}{k} \approx \frac{i\omega\mu}{\sqrt{i\omega\sigma}} \approx \sqrt{i}\sqrt{\omega\rho\mu} = \sqrt{\omega\rho\mu}\,.\,e^{\frac{i\pi}{4}} \quad 2.21$$

$$Z_{yx} = \frac{E_y}{H_x} = -\frac{i\omega\mu}{k} = Z_{xy} \quad 2.22$$

Through equation 2.23, the true earth resistivity is related to the characteristic impedance for a homogeneous and isotropic earth:

$$\rho = \frac{1}{\omega\mu}\left|Z_{xy}\right|^2 = \frac{1}{\omega\mu}\left|Z_{yx}\right|^2 \qquad 2.23$$

According to Mohamud (2013) and Nabi et al., (2020), the apparent resistivity (ρa) formula for a non-homogenous earth can be defined as if the earth was homogeneous. For a homogeneous earth in practical units, the resistivity, ρ, in equation 2.23 can be written as:

$$\rho_a = 0.2T|Z|^2 = 0.2T\left|\frac{E_x}{B_y}\right|^2 \qquad 2.24$$

The phase (θa) and apparent resistivity (ρa) for non-homogeneous earth are functions of frequency and are characterized by the equation 2.25 (Nguimbous-Kouoh, et al., 2018):

$$\rho_a = 0.2T|Z|^2 \; and \arg(Z) \neq 45^o \qquad 2.25$$

2.3.2.2 The Impedance Tensors

Conductivity σ of 1D layered earth only changes with depth (Elizondo, 2019). Therefore, the impedance tensor, ***Z***, can be expressed as

$$Z_{1D} = \begin{bmatrix} 0 & Zxy \\ -Z_{xy} & 0 \end{bmatrix} \qquad 2.26$$

The apparent resistivity ρ_a, and the impedance phase angle Φ for a layered earth the apparent resistivity both depend on the angular frequency = $2\pi f$ and are expressed by equations 2.27 and 2.28 respectively (Rosenkjær, 2011):

$$\rho_a(\omega) = \frac{1}{\omega\mu_o}\left|Z_{xy}(\omega)\right| \qquad 2.27$$

$$\Phi(\omega) = arctan\left(\frac{\Im(Z_{xy}(\omega))}{\Re(Z_{xy}(\omega))}\right) \qquad 2.28$$

Conductivity, σ, for a 2D earth changes with depth and in one horizontal direction while there is no variation in conductivity on the other orthogonal horizontal direction and is generally referred to as the Electromagnetic or electrical strike direction (Hersir, 2018). If the $\boldsymbol{E}$ and $\boldsymbol{H}$ are decomposed into normal and anomalous field components, then the impedance tensor in 2D can be written as in equation 2.29

$$Z_{2D} = \begin{bmatrix} Z_d & Z_{xy} \\ Z_{yx} & -Z_d \end{bmatrix} \quad 2.29$$

$Z_{xx} = -Z_{yy} = Z_d$, the diagonal elements in the impedance tensor are opposite in signs but equal in amplitude while the off diagonal elements Zxy and Zyx are independent values (Bravo- Osuna et al., 2021). The Z_{2D} impedance tensor can be written as in equation 2.30 if the impedance tensor is rotated such that the y direction is perpendicular and x direction is parallel to the Electromagnetic strike direction (Gómez- Treviño et al., 2018)

$$Z_{2D} = \begin{bmatrix} 0 & Z_{xy} \\ Z_{yx} & 0 \end{bmatrix} \quad 2.30$$

Impedance tensor modes; B-polarization or Transverse Magnetic (TM) mode $Z_{TM} = Z_{yx}$ and E-polarization Transverse Electric (TE) mode $Z_{TE} = Z_{xy}$ can then be independently evaluated such that the phase angle and apparent resistivity for each of the mode be obtained from equations 2.31 and 2.32 respectively (Rosenkjær, 2011).

$$\Phi_{TE}(\omega) = arctan\left(\frac{\Im(Z_{TE}(\omega))}{\Re(Z_{TE}(\omega))}\right) \; and \; \Phi_{TM}(\omega) = arctan\left(\frac{\Im(Z_{TM}(\omega))}{\Re(Z_{TM}(\omega))}\right) \quad 2.31$$

$$\rho_{appTE}(\omega) = \frac{1}{\omega \mu_o}|Z_{TE}(\omega)|^2 \; and \; \rho_{appTM}(\omega) = \frac{1}{\omega \mu_o}|Z_{TM}(\omega)|^2 \quad 2.32$$

The conductivity σ , in 3D earth changes in all directions and all the elements in the impedance tensor are non-zero elements. Impedance tensor is represented by equation 2.33 (Elizondo, 2019).

$$Z_{3D} = \begin{bmatrix} Z_{dxx} & Z_{xy} \\ Z_{yx} & Z_{yy} \end{bmatrix} \quad 2.33$$

2.3.2.3 Depth of Penetration (Skin Depth)

According to Brodie 2019, the depth below the surface where the Electromagnetic field has diminished to 1/e of its original value at the surface and is given by equation 2.34:

$$\delta = \frac{1}{Reak\ (k)} = \frac{1}{Real\ (i\omega\mu\sigma)} = \sqrt{\frac{2}{\omega\mu\sigma}} = \sqrt{\frac{2T\rho}{2\pi .4\pi .10^{-7}}} = \frac{10^3}{\pi} \cdot \sqrt{\frac{20}{8}} \cdot \sqrt{T\rho} \quad 2.34$$

This reduces to:

$$\delta \approx 500\sqrt{T\rho} \quad 2.35$$

δ = Skin depth (m)

T = Period (s)

ρ = Resistivity (Ωm)

2.4 Artificial Intelligence

Computers have become immensely powerful over the recent years; therefore, it has become easier to emulate tasks carried out by a human being (Chaki, 2015). Artificial intelligence (AI) is used to refer to intelligence in machines to perform certain functions as illustrated in Figure 2.12 with the ability of rationalizing and taking actions that have the best chance of achieving a specific goal (Karrar & Sun, 2017).

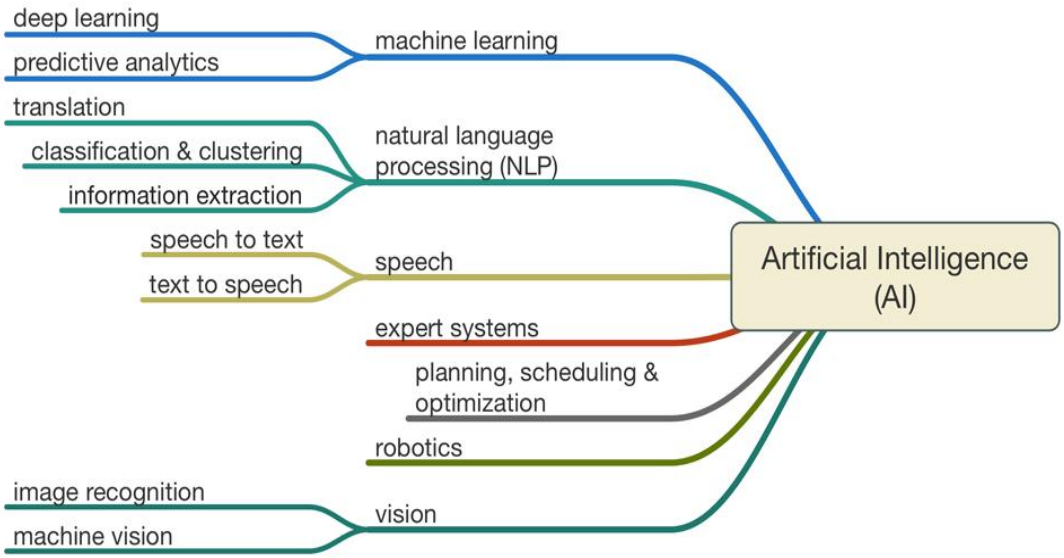


Figure 2.12: Artificial Intelligence Techniques (Gokani, 2017)

2.4.1 Machine Learning (ML)

In order to improve energy productivity and reduce exploration costs, reservoir evaluation by use of integration techniques has attracted attention (Alhakeem, 2018). Machine Learning is the ability of a computer algorithm to learn from data and accordingly through experience make or become better by discovering patterns that can be later used to analyze new data. This area of study comes in handy because of the complexity of tasks to be learned, that traditional methods of programming cannot handle (Dawood, 2019). For this to happen, the issue to be worked out is analyzed, sufficient data for training the machine is then collected and input in the model to evaluate the outcomes. From own experiences, the program learns as more data comes in as illustrated in the Figure 2.13

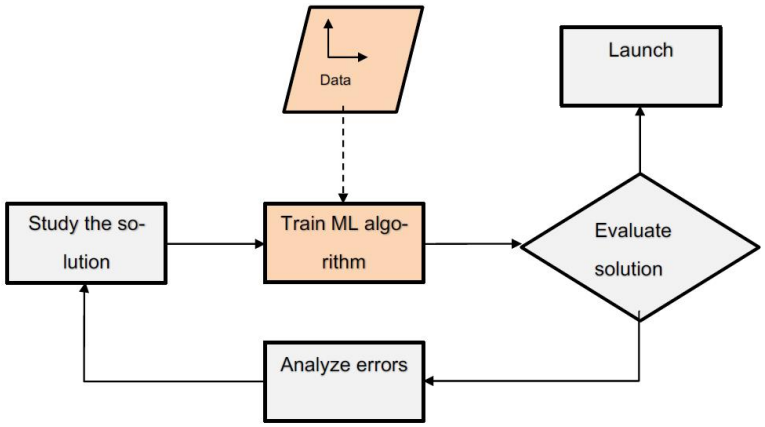


Figure 2.13: Machine Learning Approach (Sapkota, 2019)

Machine Learning (ML) is a research area in computer science Artificial Intelligence, which can be incorporated to develop intelligent models for geothermal systems to improve the learning from the geothermal wells data. Machine Learning can be used in geothermal research in the prediction of geothermal parameters provided with the wells data (Muther et al., 2022) Machine Learning applied in geothermal technology can be used to solve the old problems using the new technology solutions to augment research in the field from data ingestion, data preprocessing, modelling, visualization and interpretations. Artificial Intelligence's Machine Learning, if embraced will aid in the achievement of predictive models that can be used on existing data to predict future unseen geophysical properties in direct exploration methods (Antonopoulosa et al., 2020).

Due to plentiful geophysical and geological data application of Machine Learning techniques or other different methods of Artificial intelligence in data integration for the exploration has attracted much interest (AlHakeem, 2018).

2.4.2 Types of Machine Learning

According to Deland (2018) two types of techniques of Machine Learning are employed: supervised learning, where the model is trained on known input and output data with the

aim of predicting future outputs, and unsupervised learning, where intrinsic structures or hidden patterns are discovered from input data. The two Machine Learning techniques are displayed in Figure 2.14.

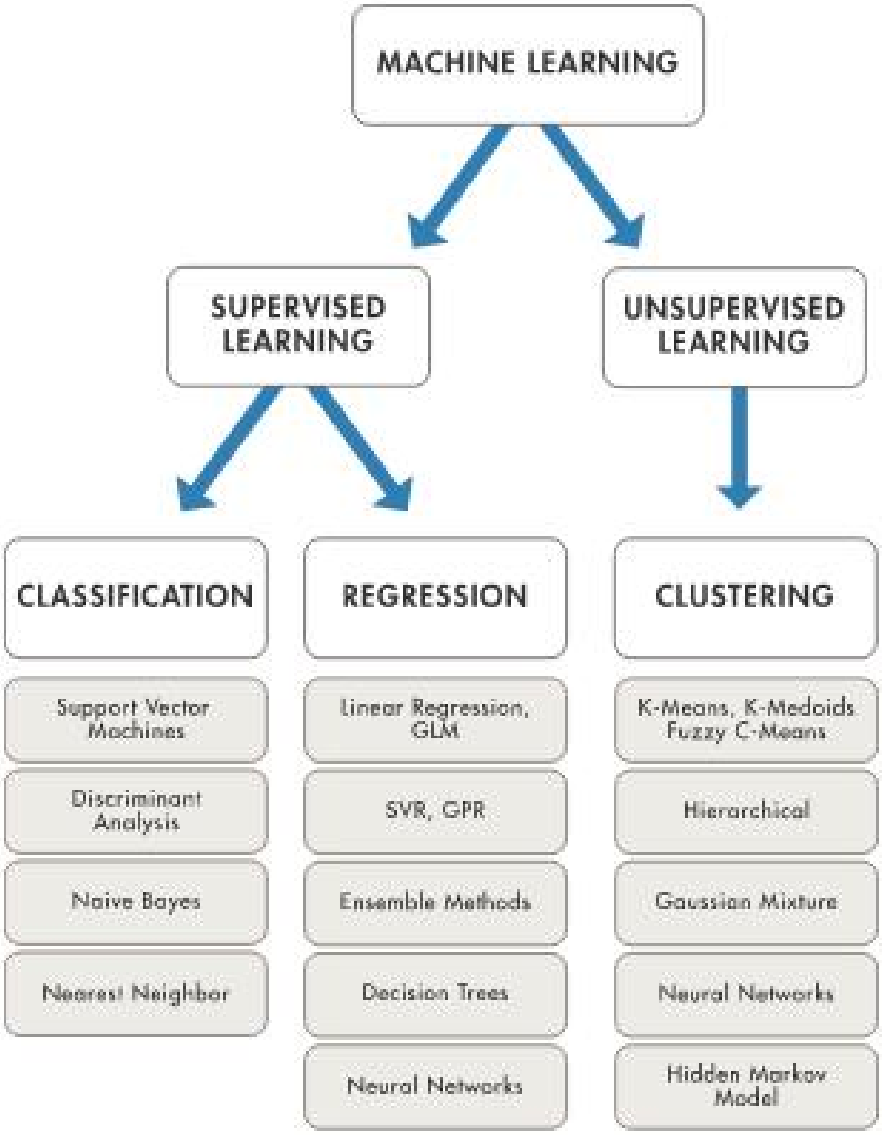


Figure 2.14: Machine Learning Algorithms (Maxhuni, 2017)

2.4.3 Unsupervised Learning

In this category, the available input data does not have the corresponding output target, but the machine tries to deduce a pattern from the data provided as shown in Figure 2.15. The algorithms are left to themselves to identify similar patterns or structures in the data and gather the clusters. In order to achieve this, the machine will require huge amount of data (Neupert et al., 2021).

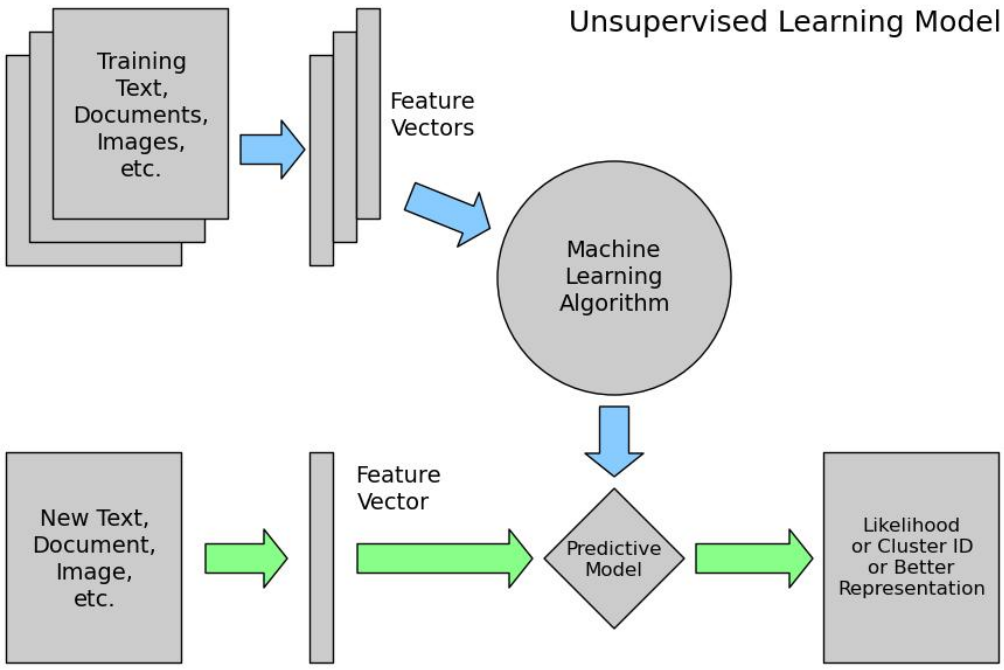


Figure 2.15: Unsupervised Learning Model (Barnadas, 2016)

2.4.4 Supervised Learning

This is where a model that makes prediction based on the known data values is developed. A program will then learn from the input data, learns and improves its prediction ability. For this process to yield a good supervised Machine Learning model, data is prepared, an algorithm chosen, a model developed and validated then applied to new data sets (Neupert et al., 2021). This is illustrated in the Figure 2.16

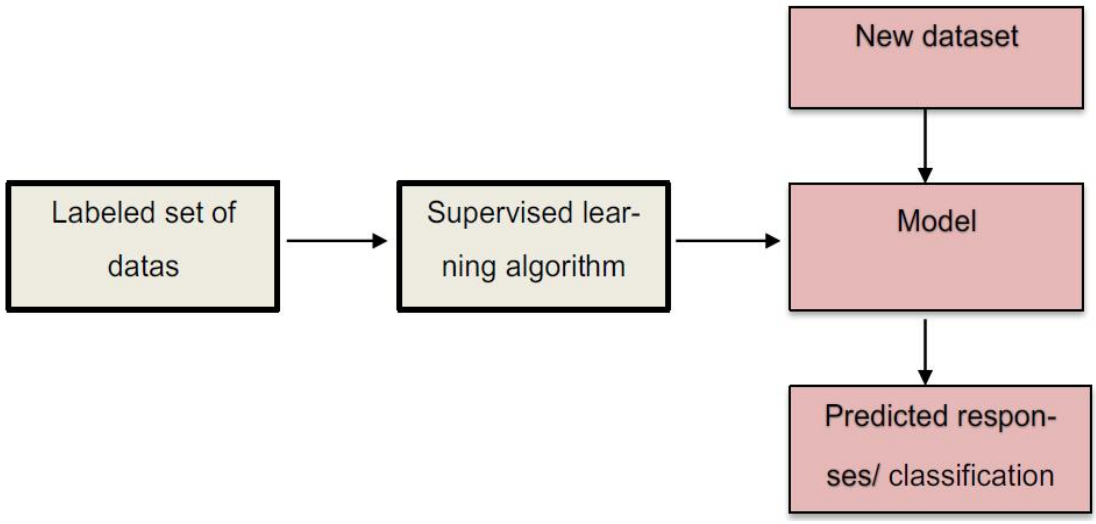


Figure 2.16: Supervised Learning Model (Neupert et al., 2021)

Based on the output to be predicted, the exercise of supervised learning is grouped into two categories: classification and regression.

2.4.5 Classification

It is used for discrete output values such as classifying whether a tumor is malignant or benign, predicting the state of weather as sunny, cloudy, or rainy; or detect whether a movie review is negative or positive statement (Dawood, 2019). To verify that the machine has learned the classification technique, the test data is used and validated.

2.4.6 Regression Model

Regression, a predictive modeling technique has been proven scientifically to predict the future and causal inference. It is a long-established statistical procedure that has been adapted in AI and Machine Learning because its models are understandable (Ahmed et al., 2020). The choice of a Machine Learning algorithm is dependent with the type of data. Regression is adapted in numerical data whose value of prediction is a single value hence the choice of the algorithm in this research.

2.4.7 Linear Regression

In simple linear regression, there is only a single feature for prediction that is dependent on a single input variable and a single output variable (Scowen et al., 2021). In multiple linear regression, there are two or more features or exploratory variables to predict the result of a single response variable. In linear regression data the equation is given by equation 2.36:

$$Y_i = f(X_i, \beta) + e_i \quad 2.36$$

Where

Y_i = dependent variable

f = function

X_i = independent variable

β = unknown parameter

e_i = error terms

Polynomial regression is a special type of a linear regression where the technique of transforming the original features into polynomial features of different degrees and the applying regression onto it (Nisbet et al., 2018). This method is used in case where the explanatory and the response variable do not exhibit linearity in data or are curvilinear in nature. The polynomial model of n^{th} order is given by equation 2.37

$$Y = \beta_o + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon$$

2.37

Where,

β_0, $\beta_1 \ldots \beta_n$ are the coefficients, the computed values by the regression tool reflecting the relationships between the explanatory and the response variables. X_1, X_2, …X_n, ε are the residuals or the bias or the random error term. This is that potion of the response variable not explained by the model. In the consideration of the order or the degree

model, it is important to keep the orders as low as possible and if it does not fit well, other orders are tried.

2.4.8 Support Vector Machine (SVM)

SVM is a technique of finding a hyperplane in the original input space through optimization by correctly separating a given training set while leaving as much distance as possible from the closest instances to the hyperplane on both sides (Pupale, 2018). Support vectors are the data points used in achieving maximal margin in regression estimation. For nonlinear separable training set, construction of nonlinear boundary is necessary by mapping original input space into feature space that is defined by a kernel function. Support Vector Regression (SVR) is the term used when the SVM algorithm is used in regression problems (Orchel, 2011). The SVR is defined by equation 2.38.

$$f(\mathbf{x}) = (\mathbf{z} \cdot \phi(\mathbf{x})) + b, \quad 2.38$$

where

b is a bias value,

z is weight vector, and

ϕ(x) is a kernel function.

The insensitivity loss ε is used in SVR and can be written as shown in equation 2.39 with ε as the region for ε-insensitivity

$$L_{\varepsilon}(f(x) - y = \begin{cases} |f(x) - y| - \varepsilon & if\ |f(x) - y| \geq \varepsilon, \\ 0, & otherwise, \end{cases} \quad 2.39$$

To minimize quadratic programming problem, SVR model can be written as shown in equation 2.40

$$\begin{aligned} min: \quad & \frac{1}{2} z^T + C \sum_i (\zeta_i + \zeta_i^*, \\ subjected\ to\ & y_i - z^T x_i - b \leq \varepsilon + \zeta_i, \\ & z^T x_i + b - y_i \leq \varepsilon + \zeta_i^*, \end{aligned} \quad 2.40$$

$$\zeta_i^*, \zeta_i \geq 0,$$

With i = 1, 2, ….., n representing the training data number,

C = regularization parameter,

$(\zeta_i + \zeta_i^*)$ = empirical risk,

$(\frac{1}{2}z^T z)$ = structure risk

The structure risk hinders overlearning and lack of applied universality. After the constants C, ε and K are properly chosen, Lagrange function can be used to resolve the optimum of each parameter. Linear, polynomial, radial basis function (RBF), or Gaussian and sigmoid kernels illustrated in equations (2.41), (2.42), (2.43), and (2.44) respectively are the Kernels mostly used.

Linear kernel,

$$K(x_i, x_j) = x_i^T x_j \qquad 2.41$$

Polynomial kernel,

$$K(x_i, x_j) = (1 + x_i . x_j)^d \qquad 2.42$$

Where d is the degree of polynomial, x_i, x_j representing the input and outputs respectively

Radial Basic Function (RBF) kernel,

$$K(x_i, x_j) = exp(-\gamma \| x_i - x_j \|^2) \qquad 2.43$$

γ represents the variance

Sigmoid/ hyperbolic tangent/ multiplier perceptron (MLP) kernel

$$(x_i, x_j) = tanh[v(x_i, x_j) + \propto] \quad 2.44$$

According to Alwee et al. (2013), RBF kernel is more effective with fast training process suitable for solving forecasting problems and therefore was employed in this research work.

2.4.9 Decision Tree Regression (DTR)

It is a technique that is used for fitting a sine curve on a model where the continuous values are being sought. It can be used for building classification or regression models. The model observes the features of a dataset to learn the tree structure in the dataset for predicting a real value output. The dataset is broken down into decision nodes until it reaches the leaf nodes (Rifat, 2020). The Decision nodes can further be divided into either one or two more decision nodes. The top decision node is referred to as the root node and is considered the best predictor.

According to Loh (2011), regression tree was established for evaluation of real-valued functions and can be referred to as a variant of decision tree. Regression tree is modelled for solving dependent and continuous variables or ordered discrete values in which the sum of the squared difference between the predicted and observed values is used to measure the prediction error. The datasets for the operation of regression trees consists a single output variable with one or more input variables. The output and input variables are also known as response and predictor variables, respectively, and the output variable is numerical. Generally, the methodology employed in the construction of regression trees allows the input variables to be a combination of continuous and categorical variables. Whenever each decision node in the regression tree contains a test on the values of some input variables, a decision tree is developed and the terminal node of the tree contains the values of the predicted output variable.

The binary recursive partitioning process is an iterative process which splits the dataset into simple partitions and then continues to split every partition into smaller partitions or groups at each stage of the process.

Let $y_1,y_2,\ldots,y_N$ be a collection of observation of the response variable y_i. Each observed value $y_i,i=1,2,\ldots,N$, depends on the explanatory variable $X_1,X_2,\ldots,X_p$. This implies that we divide the predictor space which is the set of possible values for $X_1,X_2,\ldots,X_p$ into J- distinct and non-overlapping regions, $R_1,R_2...,R_j$. Then for every observation that falls into the region Rj, we make the same prediction, which is simply the mean of the response values for the training observations in Rj. The regions can have any shape depending on the user. Nevertheless, we can decide to split the predictor space into j-high-dimensional rectangles or boxes because of the ease and effortlessness in the interpretation of the resulting predictive model. Then we consider all the predictors $X_1,X_2,\ldots,X_p$, and all the possible values of the split for each of the predictors. We choose the predictor and split point that will result into a tree that has the lowest Residual Sum Square (RSS). The goal is to find boxes $R_1,R_2...,R_j$ that minimizes the Residual Sum Square given by

$$RSS = \sum_{j=1}^{J} \sum_{i \in Rj} (y_i - \hat{y}_{Rj})^2 \qquad 2.44a$$

Where $\hat{y}_{Rj}$ is the mean response for the training data set within the jth box. A top-down greedy approach called recursive binary splitting is applied. The process is known as at top-down process since it starts at the top of the tree, that is, the point where all the observations belong to one region and then splits the predictor space. Each split is specified through two new branches further down the tree.

In performing the recursive binary splitting, we first select a predictor and a split point such that splitting the predictor space into two regions results to the greatest possible reduction in RSS. This process is repeated, with the aim of finding the best predictor and best split point in order to split the data further and minimize the RSS within each of the regions. However, instead of splitting the entire predictor space this time, we split one of the previously identified regions. This process is continued until a certain criterion is

reached. Let us consider the tree in Figure 3. The first node asks if X_1 is less than or equal to some threshold t_1 the mean square error, a measure that tells us how much our predictions deviate from the original target and that's the entry-point of mean square error. In case of a yes, we then ask if X_2 is less than or equal to some other threshold t_2. In case of a yes, we go to the bottom left quadrant of the space, in particular $R1$. In case of a no, we ask if X_1 is less than or equal to t_3. And we continue in this same way. The result of those parallel splits of axis is the partition of a $2d$ space into 5 different regions.

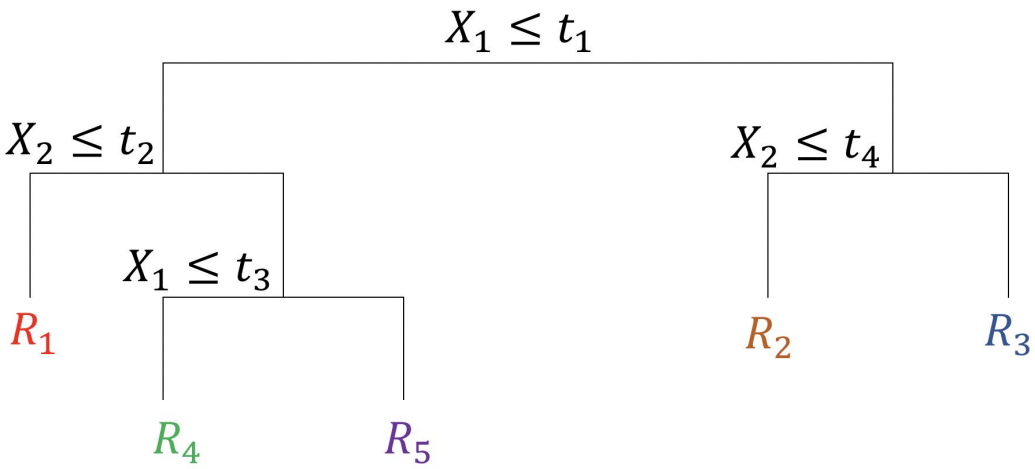


Figure 2.17: DTR Split

The DTR split the node by use of the core algorithm known as Iterative Dichotomizer (ID3) (Gao et al., 2018)

2.4.10 Random Forest Regression (RFR)

From the name, RFR is incorporated with numerous individual decision trees each spitting out a class prediction operating like an ensemble and the model's prediction is the class with the most votes. A random sample with replacement of the training set is selected with response $Y=y_1,...,y_n$ bagging repeatedly (B times) from the training set $X=x_1,...,x_n$ fitting trees to these samples. From all the individual trees regression trees on

x', predictions for unseen samples x' can be made by averaging the predictions after the training (Venkatasubramaniam et al., 2017).

$$\hat{f} = \frac{1}{B}\sum_{b=1}^{B} f_b(x') \qquad 2.45$$

Prediction's uncertainty estimates can be made from all the individual regression trees on x' as the standard deviation of the predictions based on equation 2.46

$$\sigma = \sqrt{\frac{\sum_{b=1}^{B}(f_b(x') - \hat{f})^2}{B-1}} \qquad 2.46$$

B is a free parameter representing the number of samples/trees ranging from a few hundred to several thousand. Cross validation can be used to obtain the optimal number of trees B.

The general algorithm of random forests first selects the number of samples to aggregate, m , and these m prediction models are aggregated to give a stable and lower variance prediction response. However, instead of selecting all the original variables at each split in the bagging trees, a random selection of k variables from all the original variables is performed at each split. Only the variable with best performance within this subset can be selected to split the data. Thus, tree correlation can be de-correlated by introducing this kind of randomness to the tree construction process.

There are two tuning parameters in the Random Forest model: the number of the samples to aggregate, m and the number of the randomly selected variables, k. As the number of trees m increase, the computational burden will also go up. As the intuitive concept of the Random Forest, a forest within a large number of trees (m > 1000) is suggested to use. And typically $k = \sqrt{p}$ where p is the total number of the variables in the original dataset. As the randomly selected variables is only a small part of the original variables, even if the number of trees m in Random Forest is much bigger than that in Bagging Tree, the computation is still more efficient than that of bagging trees.

Apart from the stable, highly accurate and efficient characteristics, Random Forest is also able to deal with the dataset with a large number of variables, and the relative importance of variables can still be estimated even if the correlation among variables and the tuning parameter k have serious influence on the result. It is also a good approach to estimate the missing data and maintain good performance for the dataset with a large number of missing data. The disadvantage of the Random Forest is that it is not able to do the prediction when the predicted response is beyond the range of the observed outcomes in the training data.

2.4.11 Adaptive Boosting (AdaBoost)

AdaBoost is a popular boosting algorithm for which the fundamental principle is combining weak estimators and implementing them on improved data versions. The results of each prediction are then combined with the help of weighted majority voting/hard voting. AdaBoost Regressor is used for solving regression problems (Brownlee, 2020).

Adaptive Boosting (AdaBoost) first gives a weight value to the initial sample, inputs the sample and the corresponding weights into the base learning method, and trains a base learner. The distribution of training sample weights is adjusted according to the base learner's performance, and the next base learner is trained based on the adjusted weights. The process is repeated until the number of base learners reaches a pre-specified value. All base learners are finally combined according to a binding strategy to obtain the final strong learner.

The process of applying the AdaBoost regression algorithm to the grain yield dataset is as follows:

The input grain yield data training set:

$$T = \{(x_i, y_{i)})\}_{i=1}^{m},$$

the base learning algorithm is τ; the number of base learners is K, and the output final strong learner is $f(x)$.

Distribution of weights for initialized training sample samples.

$$\hat{T}(1) = (w_{11}, \ldots \ldots w_{1m}); w_{1i} = \frac{1}{m}; i = 1,2 \ldots.$$

For iterative rounds there is k=1,2,⋯,K

Train the base learner $G_k = \tau(T, \hat{T})$ using a T_k with weights;

Calculate the maximum error of the sample on the training set.

$$E_k = \max |i = 1,2 \ldots . m$$

To calculate the relative error for each sample, here the linear error was chosen.

$$e_{ki} = \frac{|y_i - G_k(x_i)|}{E_k}$$

Calculate the regression error rate of the base learner G_k on the training set, weighting coefficients, update the distribution of weights for the training set sample T_{k+1}.

$$\begin{Bmatrix} \varepsilon_t = \sum_{i=1}^{m} w_{ki} e_{ki} \\ \alpha_t = \frac{\varepsilon_t}{1 - \varepsilon_t} \\ Z_k = \sum_{i=1}^{m} w_{ki} \varepsilon_k^{1-e_{ki}} \\ w_{k+1,i} = \frac{w_{ki}}{z_k} \alpha_k^{1-e_{ki}} \end{Bmatrix}$$

A linear combination of base learners is constructed, and the binding strategy uses the median of the weights of the base learners is taken. The base learner is used as the basic method for the strong learner, and the final strong learner is as follows:

$$f(x) = \sum_{k=1}^{k} \left(ln \frac{1}{a_k} \right) g(x)$$

Where $g(x)$ is the mediam of all $\alpha_k G_k(x), k = 1,2 \dots \dots K$

2.5 Machine Vision

The image processing is an analysed and manipulation of a digitalized image, especially in order to improve the quality of image processing (Lee et al., 2019). Digital Image Processing technique can be applied in variety of different fields such as Diagnostic image analysis, Surgical planning, object detection and matching, background subtraction in video, localization of tumours, measuring tissue volumes, locate objects in satellite images (roads, forests, etc.), traffic control systems, locating objects in face recognition, iris recognition, agricultural imaging, and medical imaging (Habuza et al., 2021).

Modern digital technology has made it possible to manipulate multi-dimensional signals with systems that range from simple digital circuits to advanced parallel computers. Generally image processing consists of several stages: image import, analysis, manipulation and image output. Computer algorithms play a crucial role in digital image processing.

2.6 Data-Driven Discovery Predictive Modelling

Data-Driven Discovery Predictive Modelling is defined as the construction of robust empirical models of real or complex systems with the intention of helping decision makers establish the relationship between the input and output parameters of the system

without having a proper understanding of the physical behavior of the system (Mont´ans et al., 2019). This is illustrated in Figure 2.18.

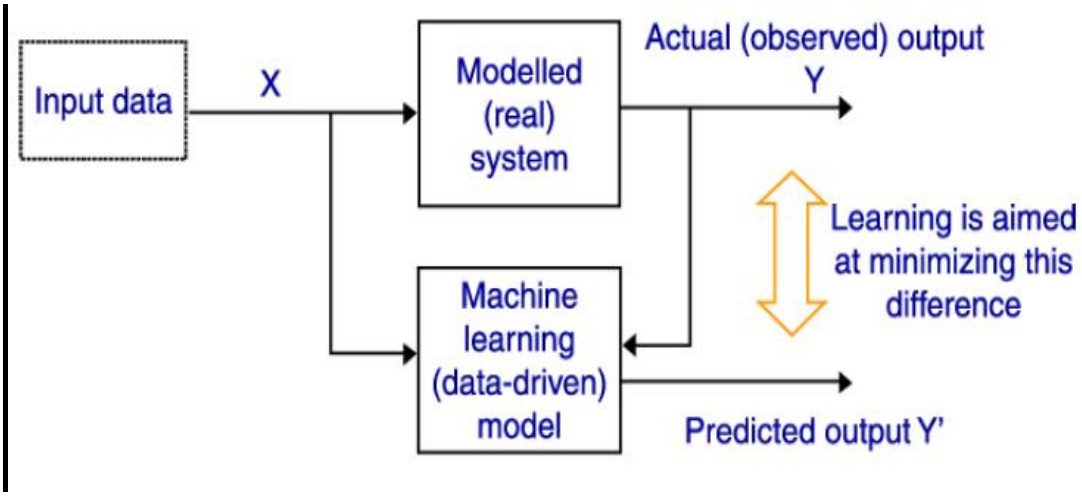


Figure 2.18: Data Driven Modelling Block Diagram (Enyenihi, 2018)

The technique can develop models that characterize the performance of correlating physical processes through the study of relevant data characterizing the systems of interest (Siebert et al., 2021). This is by inferring the dependencies between system inputs and outputs using certain learning algorithms without building the complex physical models (Lee et al., 2019). The detected relations or patterns can be formulated into mathematical statement or laws in the case of inventions in science, which then enables prognosis.

Advancements in computational intelligence especially in Machine Learning area have facilitated rapid growth of empirical modelling and its capabilities. Computational intelligence and Machine Learning are the two main methods used to construct models that Data-driven modelling is relied upon (Sarker, 2021). Input and output relationship of the system is determined by Machine Learning algorithm by use of a training data set that represents all the systems behaviours. After training the model, independent data set

is then used to test the model to establish how well it can predict another output data through testing correlations between different variables and observations. This enables performing of predictions in the absence of specific laws governing the available data sets (Ngarambe et al., 2020). Although data-driven modelling techniques have been widely employed in almost all fields of research, its application in geothermal industry reservoirs and characterizing reservoir is lacking; in particular, the applications of data-driven models for subsurface parameter characterization.

2.7 Previous Work

Geochemistry, geology and geophysical methods of survey such as gravity, seismic, magnetics, TEM, and MT were performed by stakeholders in the Kenya's rift valley. These surveys were motivated by the establishment of surface manifestations with the aim of imaging the subsurface features and subsequently its geothermal potential (Macharia et al., 2017).

2.7.1 Geochemical Data

The Na-K geothermometer was used to determine subsurface temperature and higher temperatures were observed from the central towards the western and the southwestern regions of the Olkaria domes field as shown in Figure 2.19 (Wamalwa et al.,2016)

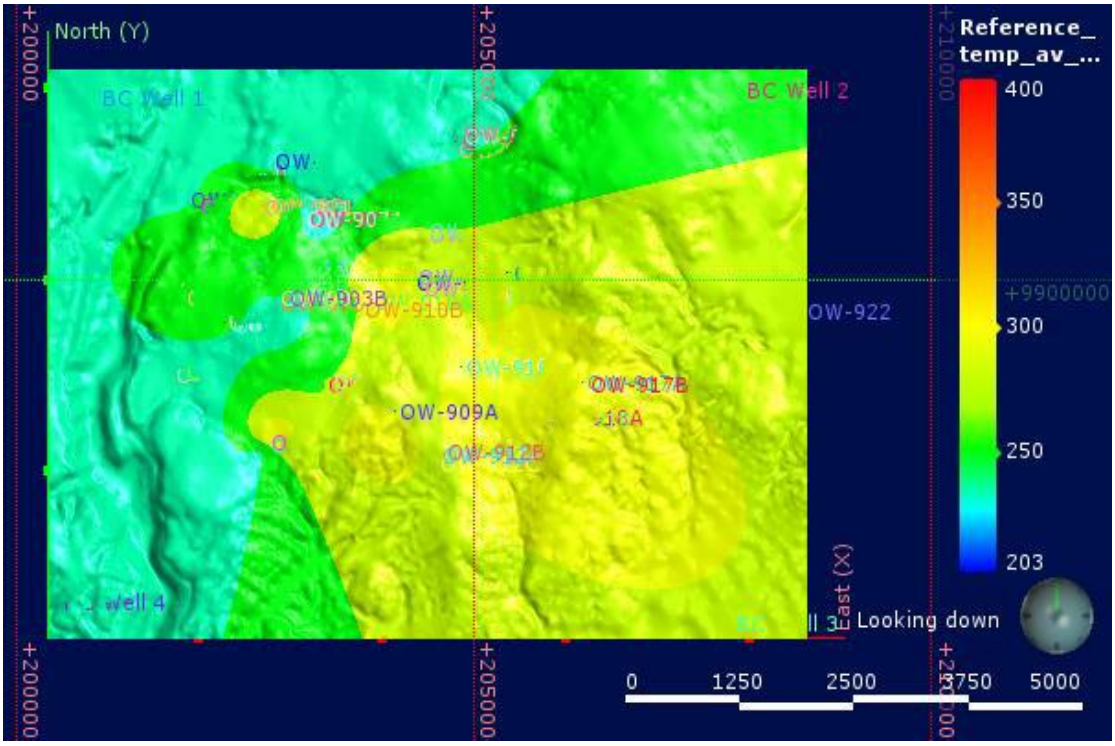


Figure 2.19: Na-K Geothermometer Temperature Average (Wamalwa et al., 2016)

2.7.2 Geophysical Techniques

Early exploration for reconnaissance mapping of the geothermal field was done by direct current (DC) resistivity method (Mariita, 2013). However, improvement in technology by introduction of state-of-the-art geophysical techniques such as MT made accurate mapping of subsurface structures achievable leading to development of accurate geothermal models (Figure 2.2

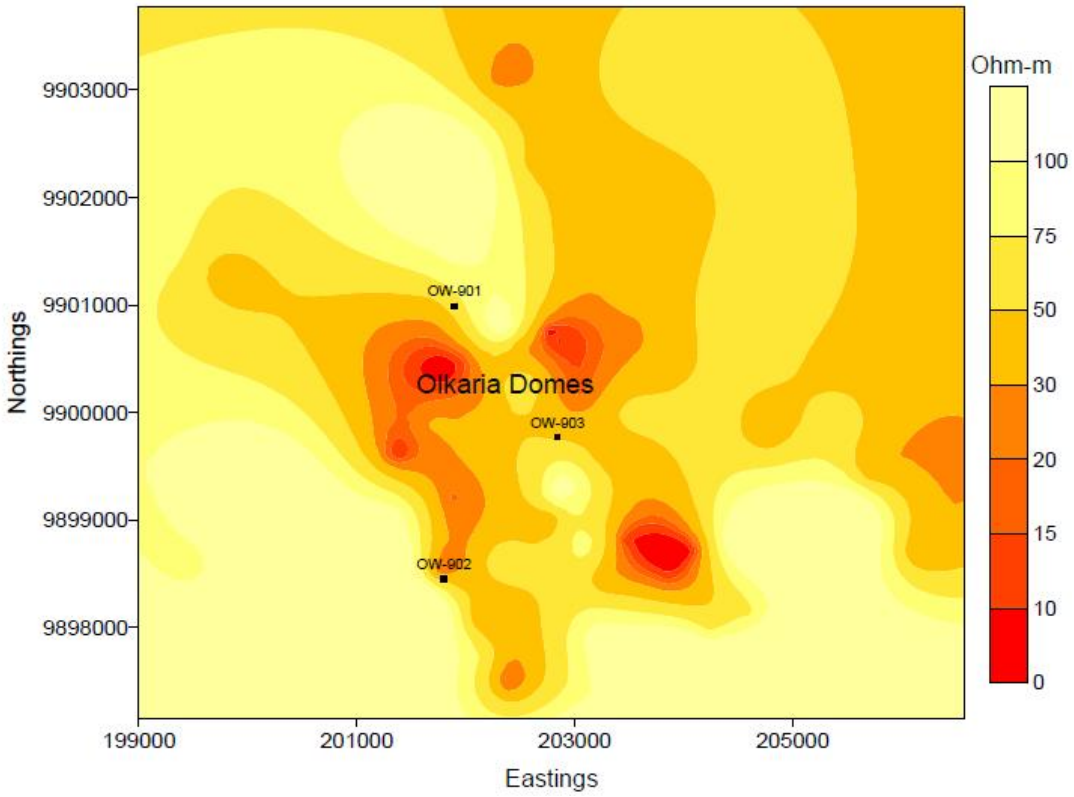


Figure 2.20: MT Resistivity Distribution at 1000 m b s l at Olkaria Domes Field (Mariita, 2013)

In the mid-1990s, TEM method that gives better resolution at depth was introduced. TEM and MT have been found to complement each other for better resolution.

2.7.3 Reservoir Analysis

The down-hole temperature and pressure profiles tests have been used in Olkaria domes to estimate permeability, transmissivity and storability for the purposes of determining individual well productivity (Mbithi, 2016). Feed zones for specific wells have been identified and interpreted through water loss zones during pumping tests and monitoring the loss circulation loss of fluid during drilling (Mbithi, 2016; Okoo, 2013).

2.8 The Research Gaps

Well drilling has been a continuous process and therefore new data which was not included in the previous temperature and geophysical resistivity studies has been obtained. The data for new wells were included in the research to improve on the existing information about subsurface structures.

Despite all the advancements that have made Kenya among the top ten leading producers of geothermal energy in the world and number one in Africa, geophysical data processing for different techniques is always done independently. Integration of multiple datasets is only done through manual visualization to generate a better model. Image processing technique is therefore needed to stack contours images from different data to generate one image to solve this.

Integration between the well completion tests such as temperature parameter and geophysical MT and TEM for correlation analysis and for the purposes of applying TEM/MT resistivity data as a tool for determining subsurface temperature is also desired. In data driven modelling, Machine Learning is applied to describe the system by use of data so that a model is developed to explain how the resulting physical processes behave based on knowledge, experience and the observed data (Ma, 2018). These models can be used to build up extremely nonlinear relationships (Herzog et al., 2018). The development of Machine Learning algorithms as an alternative tool to predict subsurface temperature from resistivity in geothermal reservoirs was inspired by the absence of geothermal reservoir research based on data-driven models in the literature.

CHAPTER THREE

MATERIALS AND METHODS

3.0 Data Collection Methods and Procedures

This research used the Electromagnetic (MT and TEM), pressure and temperature data that has been acquired by KenGen during geophysical exploration and well testing between 2004 to 2016.

3.1 Pressure Transient and Temperature Heat Up Data

The Kenya Electricity Generating Company has successfully drilled more than thirty deep wells in Olkaria domes field for the purpose of energy exploitation (Nyandigisi, 2020). For this study, 15 wells were randomly selected with the aim of getting a good representation of the entire field. These wells were: OW901, OW902B, OW903, OW905A, OW907A, OW908, OW910, OW911A, OW912, OW914B, OW915, OW917, OW918, OW919 and OW921. The data used for this study was collected by KenGen during well completion tests for each well at different times between 1998 to 2017.

Before any test was conducted during well testing, the maximum depth of the well was determined by running the logging tool down the well. This was also to ensure that the well was free from any obstructions (Rionomakal et al., 2018). To establish the depth where to station the logging tool, temperature profiles were obtained before step pumping was done. Step pumping was done with the first pumping rate of 1000 lpm. The rate was increased to 1300 lpm then to 1600 lpm and finally to 1900 lpm. Pumping was done for 4 hours for the first step and for 3 hours for the other steps. While still pumping at 1900 lpm, the pump was turned off. The pressure fall-off was then monitored for 8 hours. After fall-off a temperature profile run was conducted during well recovery. Grapher v13 software was used for estimation of the formation temperature and identification of permeable zones or feeder zones. Surfer 16.1 software was used for generation of contours

3.2 Resistivity Data Acquisition

3.2.1 MT Data Acquisition

The MT data were collected by KenGen between 2004 and 2016 by use of a 5-channel (MTU-5A) MT equipment. The equipment layout was as shown in Figure 2.11. MT data from each station was acquired for about 20 hrs. A total of 120 MT soundings covering an area of about 45 km^2 were used for interpretation in the Domes prospect.

3.2.1.1 MT Data Processing

The Synchronized Satellite Magntotelluric (SSMT2000) program, which is provided by Phoenix Geophysics in Canada (Phoenix Geophysics, 2005), was used to process the data in time-series form from the MT. through Fourier transformation, the data was changed to frequency spectrum. Noisy data points were eliminated by the MT editor. The resulting time-series data were decomposed through Fourier transformation to the frequency domain. From the Fourier transform band, the robust processing method was used to compute average cross-powers and auto powers, which were then edited to get rid of noisy data by the MT editor program graphically. The data was then saved in Electronic Data Interchange (EDI) files and later used as TEMTD input for joint inversion with the TEM.

3.2.2 Transient Electromagnetic (TEM) Data Acquisition

The TEM equipment used in this survey is from Zonge. A 200 m × 200 m transmitter wire loop was used as shown in Figure 3.1. The TEM data were collected by (KenGen) between 2004 and 2016.

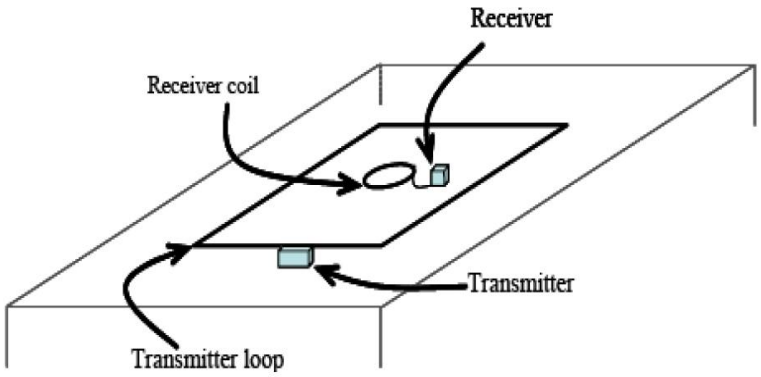


Figure 3.1: Transient Electromagnetic Equipment Setup (Flóvenz et al., 2012)

A total of 79 Central Loop TEM soundings were carried out and were distributed

3.2.2.1 TEM Data Processing

The raw TEM data was processed using the program TemxZ for Zonge where the data collected were averaged and late time apparent resistivity computed as a function of time after current turn-off. The program contains graphical interface that offers editing or omitting outliers of the noisy raw data possibilities.

3.2.3 The MT Static Shift Analysis

Static shift is a phenomenon in MT method caused by local resistivity inhomogeneities which disturb the electrical field. The static shifts can be a big problem in volcanic environments where resistivity variations close to surface are often extreme (Marwan et al., 2019). It is possible to solve the problem of static shift in MT method through joint inversion with TEM because TEM measurements at late time have no such distortion since they do not measure electrical field. The MT static shift was corrected by joint inversion with the TEM data, and then extracted from the jointly inverted models (Sirma, 2019).

3.2.4 1-D Joint Inversion of TEM and MT Soundings

The One-dimensional joint inversion was performed simultaneously for both TEM and MT data by fitting one inversion on both data sets to obtain one model according to Lichoro (2010). This was achieved by use of an algorithm which determined the appropriate shift factor to be used to constrain the MT data to fit the TEM response. Both the MT and TEM data collected on approximately the same location were brought together in a joint inversion where TEM 1-D inversion was used for static shift correction on MT data (Christensen, 2022). The TEMTD program was used to invert MT apparent resistivity and phase derived from the rotationally invariant determinant of the MT tensor elements. In the joint inversion, one additional parameter was also inverted for, namely a static shift multiplier needed to fit both the TEM and MT data with the response of the same model, (Rosenkjær, 2011).

3.3 Image Stacking

Image processing technique for stacking multidimensional contours image data sets into one image was needed to aid in identifying areas with high temperature and high conductivity in one image for perceptions and better visualization of the reservoir. The images were treated as bidimensional arrays.

To achieve this, Python Imaging Library (PIL) an open-source library with many functions for image processing was used. The model in Figure 3.2 was build using Python programming language and Anaconda framework. The open-source web-based application Jupyter Notebook for coding and visualization was used. PIL package was installed in the anaconda framework. The images from resistivity and temperature data at different depth locations were then loaded into the model. Image pre-processing was done by resizing the input images to have equal dimensions. The images were then merged by using Pivotal Focus algorithm.

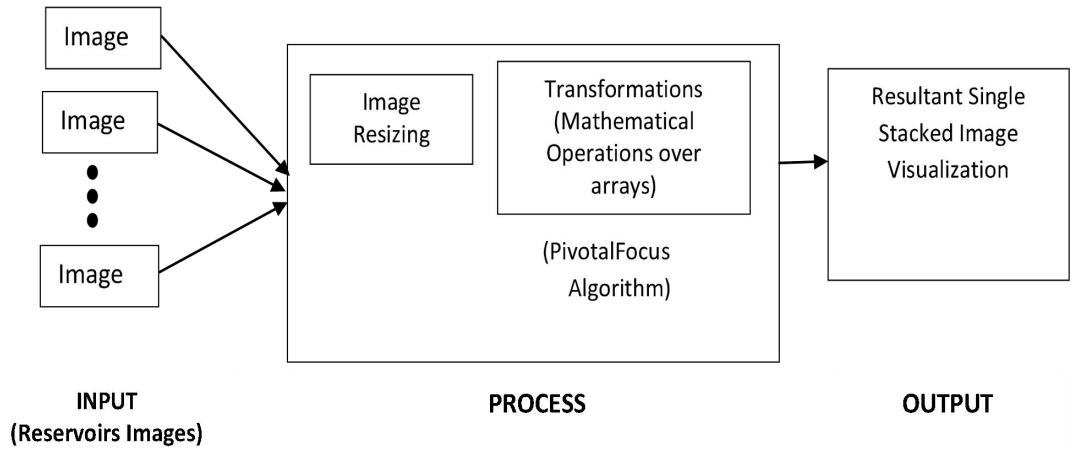


Figure 3.2: Image Stacking Pipeline

3.4 Implementation of Machine Learning Prediction Algorithms

The objective was to obtain independent input values from the geophysical resistivity data and use them to predict the subsurface temperature variables, implying that the computer would learn from the input data and predict the output data through Machine Learning algorithms.

To achieve this, regression analysis was employed based on the nature of the input and the output data. Looking into the data at hand, the desired output which would be in terms of temperature values meant that we knew what the output would be, hence, choosing supervised as opposed to unsupervised learning. The output temperature data was also continuous values and not discrete hence ruling out classification method of analysis. Having established regression as the tool for analysis, Linear Regression, Decision Tree Regression, Adaptive Booster Regression, Support Vector Regression and Random Forest Regression models were attempted in order to obtain the best model.

Resistivity and temperature dataset were originally provided in excel worksheet. The data was then merged into one Comma Separated Values (CSV) for modelling. Data

from thirteen wells was used for training the model while two wells data observations were set aside for the purposes of model verification during deployment evaluation. This was in the ration of 80% to 20% for training and testing respectively according to Chaki (2015). The training dataset was made up of 297 observations and two features; Resistivity and Temperature while the testing dataset comprised of 72 observations comprising of Resistivity and Temperature. The testing data was important as it was used to evaluate how the model was to handle new data that was outside the training data.

Data-Driven Discovery Predictive Model (3DPM) approach shown in Figure 3.3 using Machine Learning computational approach was used in this research. This approach was necessitated by the fact that cross all scientific disciplines, the governing equations have been of fundamental importance and the governing equations' traditional derivation has been based on underlying first principles. Limitations to obtain theoretical or even empirical but universal relationships between temperature and resistivity in geothermal systems are because of the very high variability range of chemical and physical conditions within the Earth. Therefore, the recourse to first-principles derivations could be untenable.

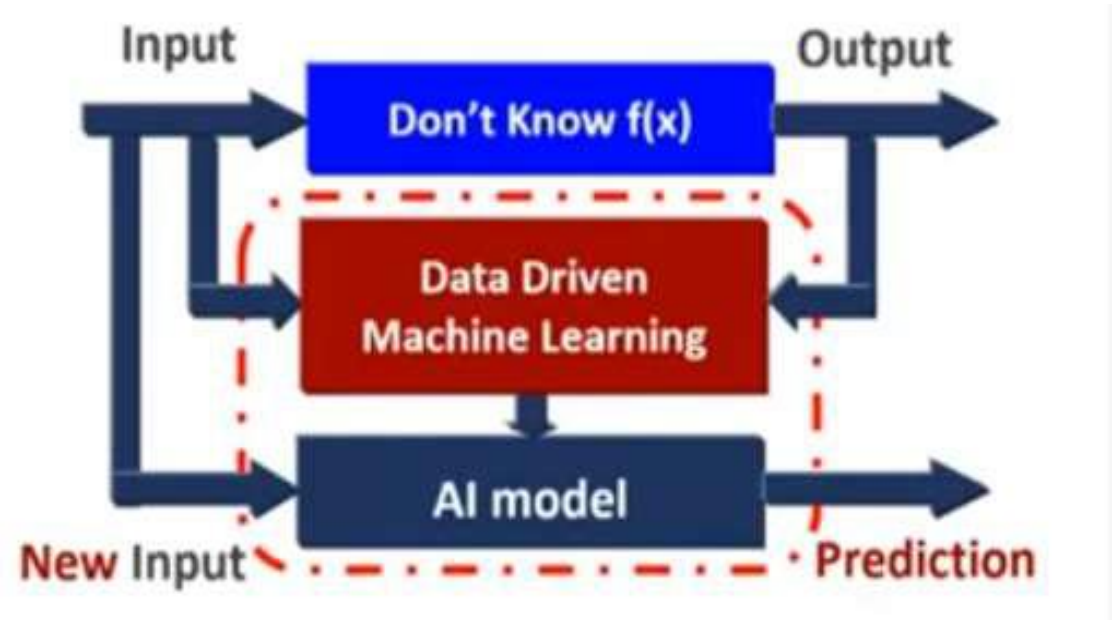


Figure 3.3: Data Driven Discovery Predictive Model (Abrahart et al., 2008)

A detailed Machine Learning procedure adopted is described in the flow chart of Figure 3.4. The model was build using Python programming language and Anaconda framework. NumPy an open-source Python library was used in working with arrays. The open-source web-based application Jupyter Notebook for coding was used. Matplotlib a graph plotting library in python was used for graph visualization. Analysis, cleaning, exploring, and manipulating data into comma separated files was carried by Pandas, while SciPy, a scientific computation library was applied in optimization and statistical processing. Sklearn was used for predicting temperature attribute from resistivity. Resistivity was used as the input parameter while temperature as the output variable. Using the same set of data, Linear Regression, Decision Tree Regression, Adaptive Booster Regression, Support Vector Regression and Random Forest Regression Machine Learning models were developed to establish the best fitting model based on the performance inspectors.

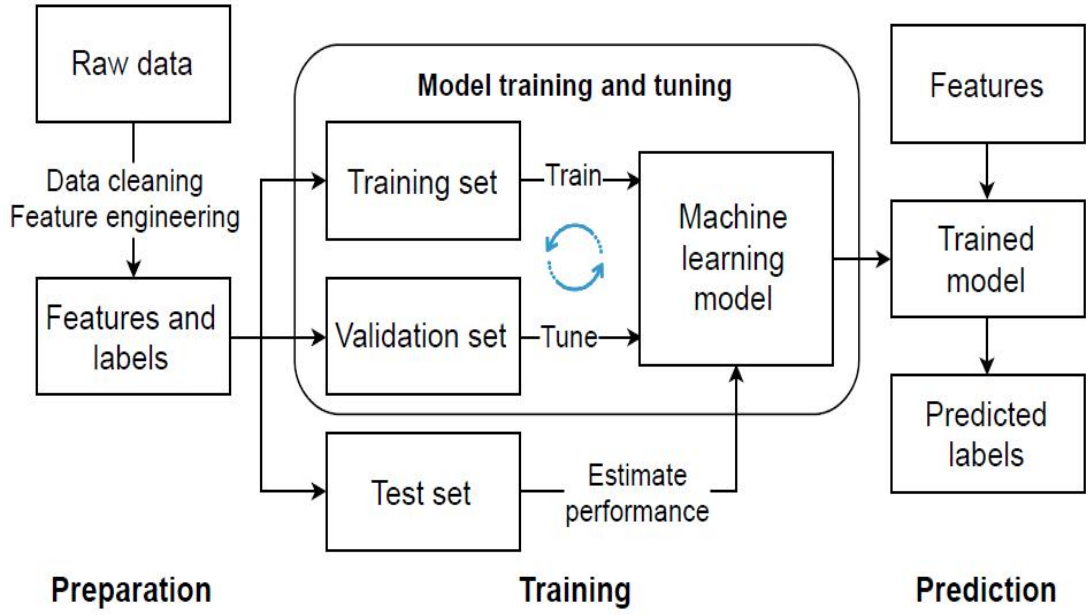


Figure 3.4: Machine Learning Prediction Approach (Steuer, 2018)

3.5 Model Performance Evaluation for Data-Driven Model

In this thesis four approaches were applied to analyze the performance of data-driven models, including cross-plot, line graphs, coefficient of determination (R^2) and Mean Absolute Error (MAE).

3.5.1 Descriptive Statistics

To observe the model predictions pattern and compare with the actual data patterns, graph of actual values and forecasting of testing data set was plotted.

3.5.1.1 Cross Plot

Cross plot is the most intuitive visualization method that was used to examine the performance of data-driven models. A schematic of a cross-plot is shown in Figure 3.5.

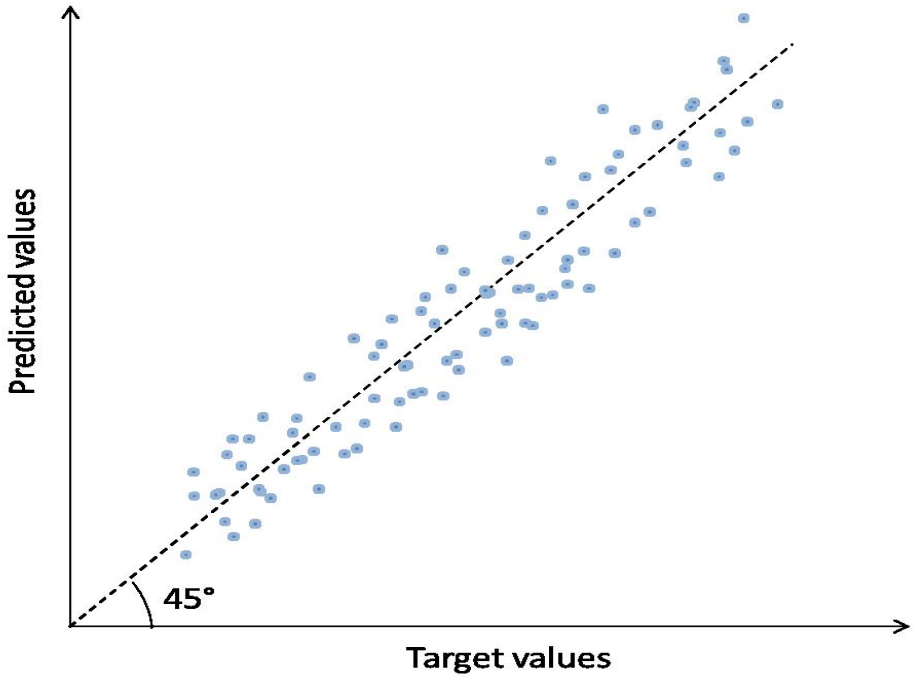


Figure 3.5: A Schematic of Cross Plot (Ma, 2018).

The x-axis as in Figure 3.5 represented the measured temperature, while the y-axis denoted the predicted temperature values from models. The location of the scatter points in the Figure depicted its corresponding prediction from the model (Das, 2018).

3.5.1.2 Line Graphs

Line graphs were used to display change of data points connected by straight line segments on the x and y-axes. They were important in determination of how temperature and resistivity values changed with depth. They were also used to display multiple line graphs to aid visualising the behaviours of actual and predicted temperature as functions of depth.

3.5.1.3 Coefficient of Determination (R^2)

Coefficient of Determination (R^2) is an important statistical measure that was used to show how well a model fitted the targets. According to Flynn (2019), R^2 is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(T_i - \hat{f}_i)^2}{\sum_{i=1}^{n}(T_i - \hat{T})^2} \quad 3.1$$

Where T_i was the observable variable or target (resistivity); $\hat{f}_i$ represented the predicted value from the model (temperature); $\hat{T}$ denotes the average of all targets; n is total number of samples; for a set of data samples, a larger value of R^2 indicates an increasing prediction precision from a regression. In this thesis, a perfect prediction means the value of R^2 is 1 according to Ma (2018).

3.5.1.4 Mean Absolute Error

Mean Absolute Error (MAE) was used as a measure of errors between measured and the predicted temperature values (Figure 3.6). MAE is calculated as (Kothapalli, 2021):

$$MAE = \frac{\sum_{i=1}^{n}|y_i - x_i|}{n} \quad 3.2$$

The smaller the MSE, the better the predictor fits the data

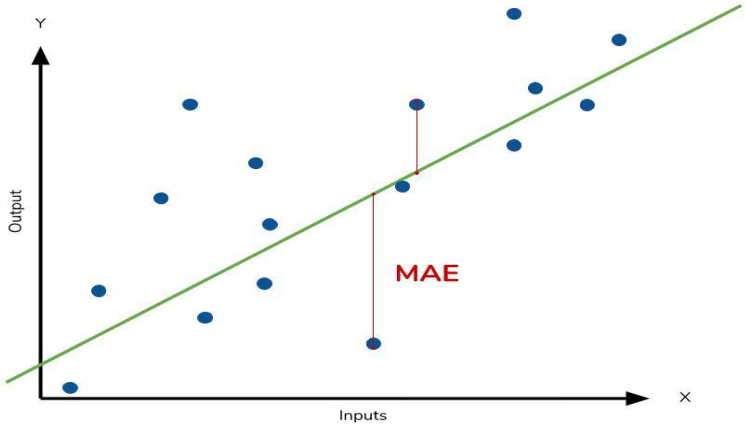


Figure 3.6: A Schematic of MAE (Willmott & Matsuura, 2005)

CHAPTER FOUR

RESULTS AND DISCUSSIONS

4.0 Introduction

This chapter discusses results with respect to the outlined objectives. Several contributions were presented in this research. First, the application of temperature heat up and pressure transient tests to evaluate reservoir sub surface structures. Secondly, the application of resistivity method by use of MT and TEM to image subsurface resistivity distribution in high-temperature Olkaria domes geothermal field. Thirdly, from the temperature and resistivity data, the application of data-driven models involving geophysical resistivity data to predict subsurface temperature by use of Regression algorithms of Machine Learning in Artificial intelligence was realized. Also image stacking by use of PivotalFocus Algorithm of machine vision of Artificial Intelligence was performed on temperature and resistivity images at particular depths to generate one image that was a representation of resistivity and temperature parameters at that particular depth is novel. This aided in identifying areas with high temperature and high conductivity in one image. This led to locating the best possible drilling sites based on the merged images of resistivity and temperature at different depths. In doing so, each of the outlined objectives was achieved.

4.1 Temperature Warm Up and Pressure Transient Tests

4.1.1 Temperature Recovery Profiles

Based on temperature recovery profiles, permeability zones were identified. Two major permeable structures were identified. The first was in the range of 900-1300 m a s l and the second permeable structure was at a depth of 250-0 m a s l in the wells as shown in Figures 4.1, 4.2 and 4.3 for wells OW905A, OW903 and OW919 respectively with the rest of the wells in appendix 1. At these depths, wells experienced adverting hotter zones in temperature recovery profiles indicating that hotter fluids from the surroundings entered the wells faster at these points during well recovery relative to other depths in

the wells. The two hotter structures appearing at same depths in most of the wells suggest that permeability in the Domes is dominated by two horizontal structures.

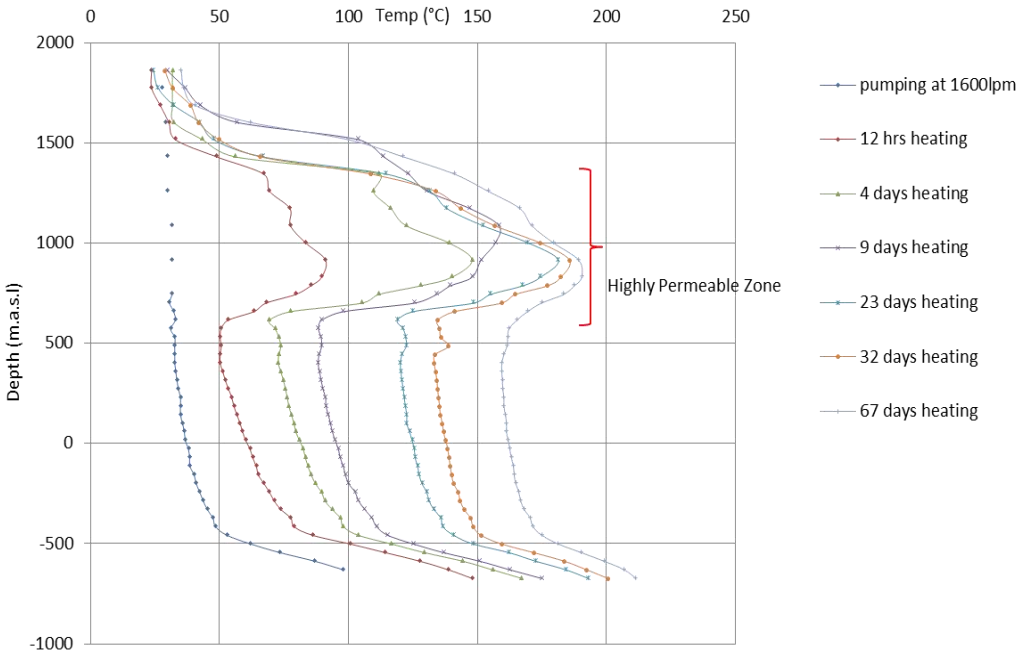


Figure 4.1: OW905A Temperature Profiles from the Heat up Test

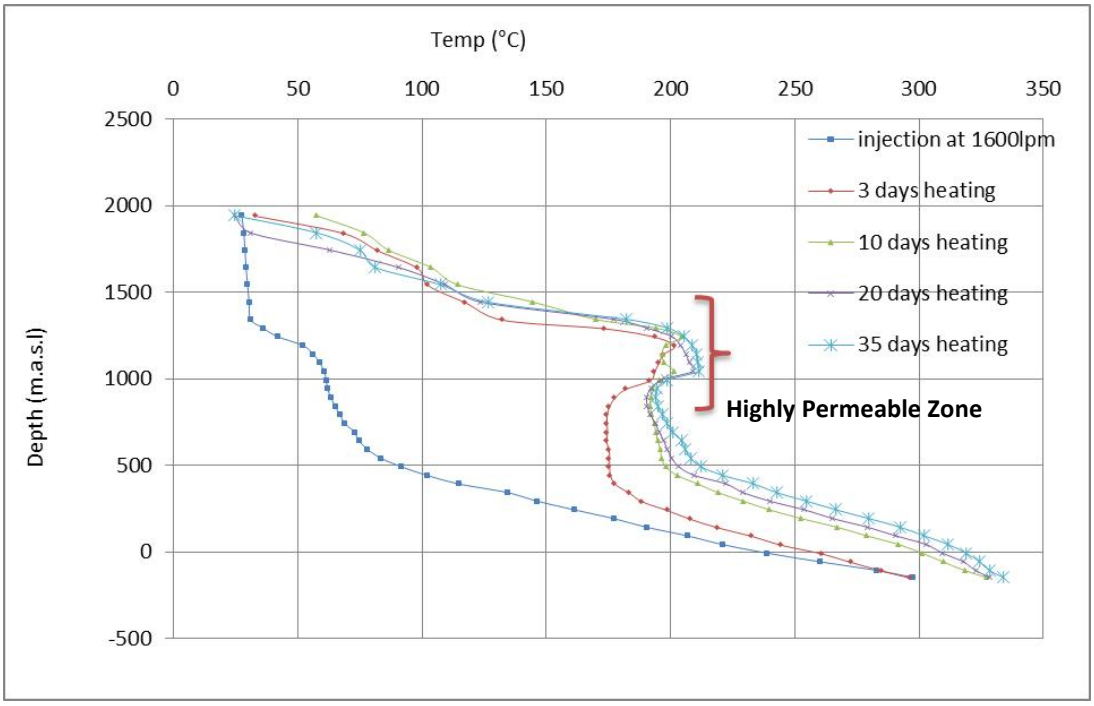


Figure 4.2: OW903 Temperature Profiles from the Heat up Test

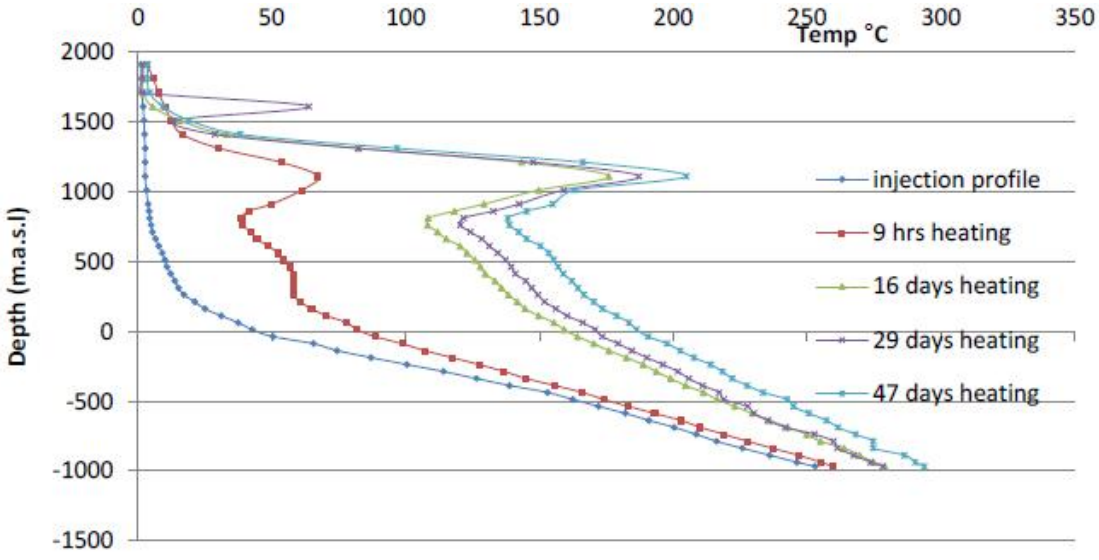


Figure 4.3: OW919 Temperature Profiles from Heat up Test

Three recharge zones were identified from the four lateral temperature contours at depths of 1000 m a s l, 500 m a s l, 0 m a s l and 200 m b s l as shown in Figures 4.4, 4.5, 4.6 and 4.7. These recharge zones were identified as areas with the lowest temperatures within the field. The low temperatures could have been occasioned by an inflow of cold fluids into the reservoir. These zones were in the NW, SW and SE directions.

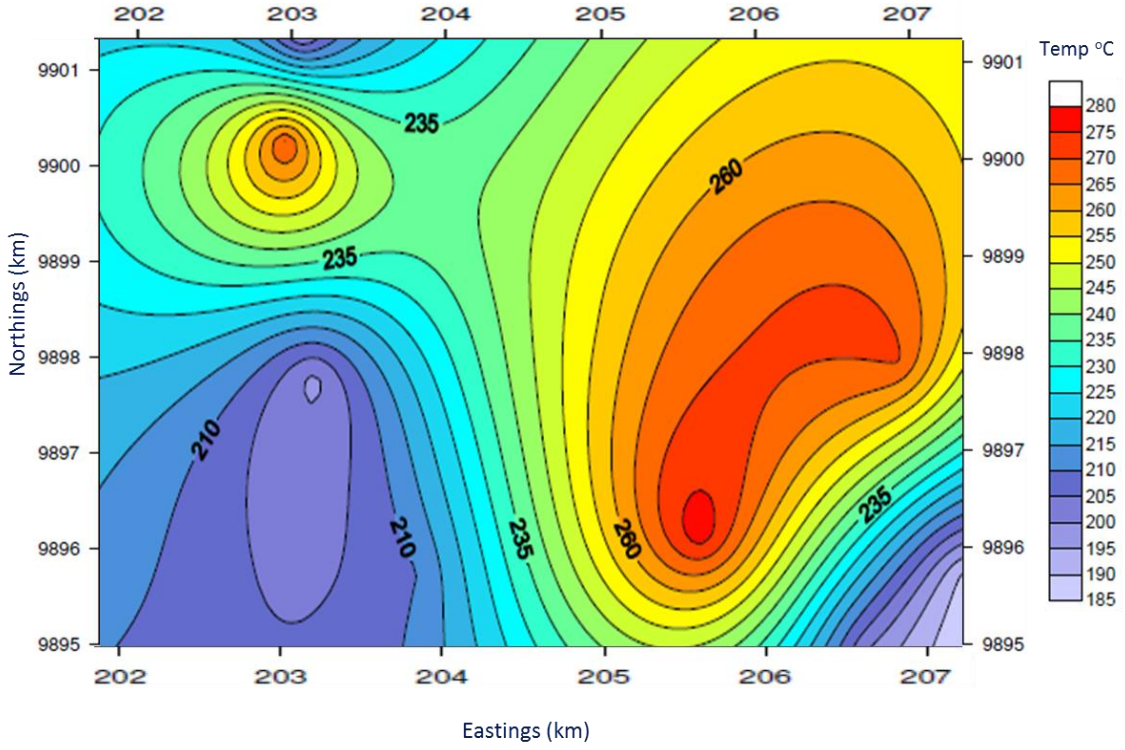


Figure 4.4: Lateral Temperature Contours at 1000 m a s l

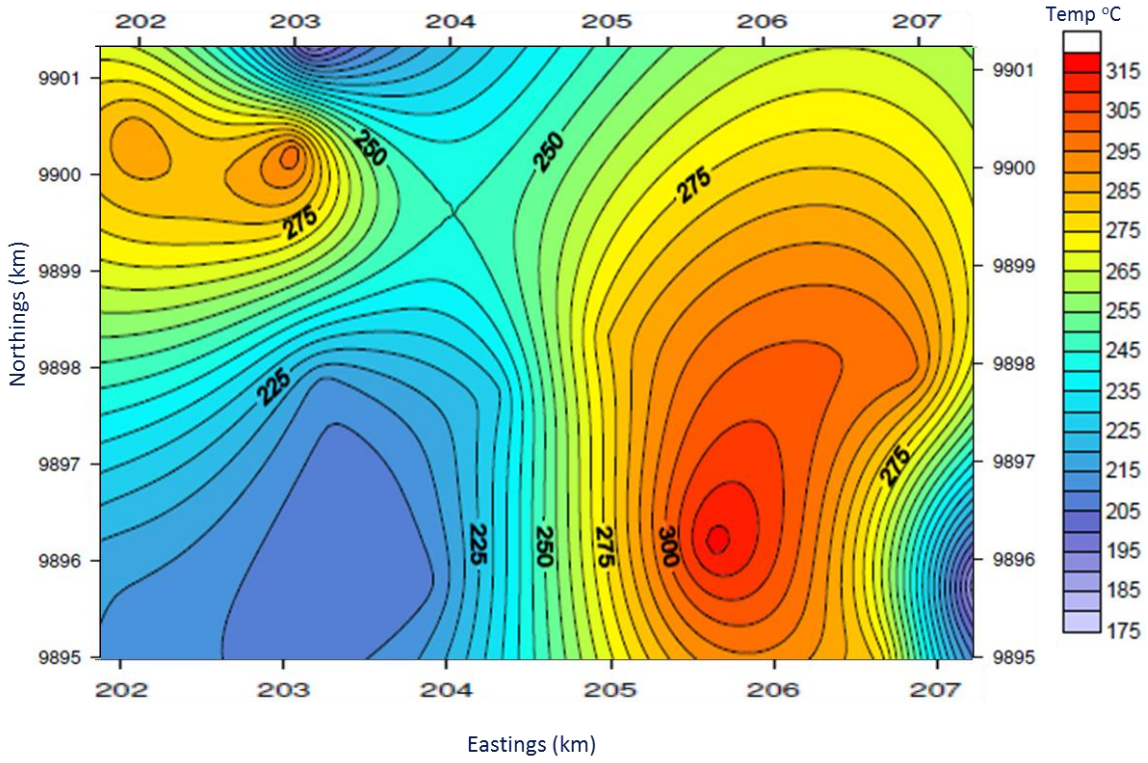


Figure 4.5: Lateral Temperature Contours at 500 m a s l

Also from Figures 4.4, 4.5, 4.6 and 4.7, two major heat sources were inferred. These two heat sources were identified as areas with the highest temperatures within the field as a result of being closer to the hot magmatic intrusions. The first heat source is located on the Northwest side of the field while the second heat source is on the SouthEast side of the study area. The two regions are characterized by high temperature isotherms. These regions exhibit temperatures more than 315 ^{0}C suggesting magma chambers beneath the areas. These heat structures could be associated with magmatic intrusions.

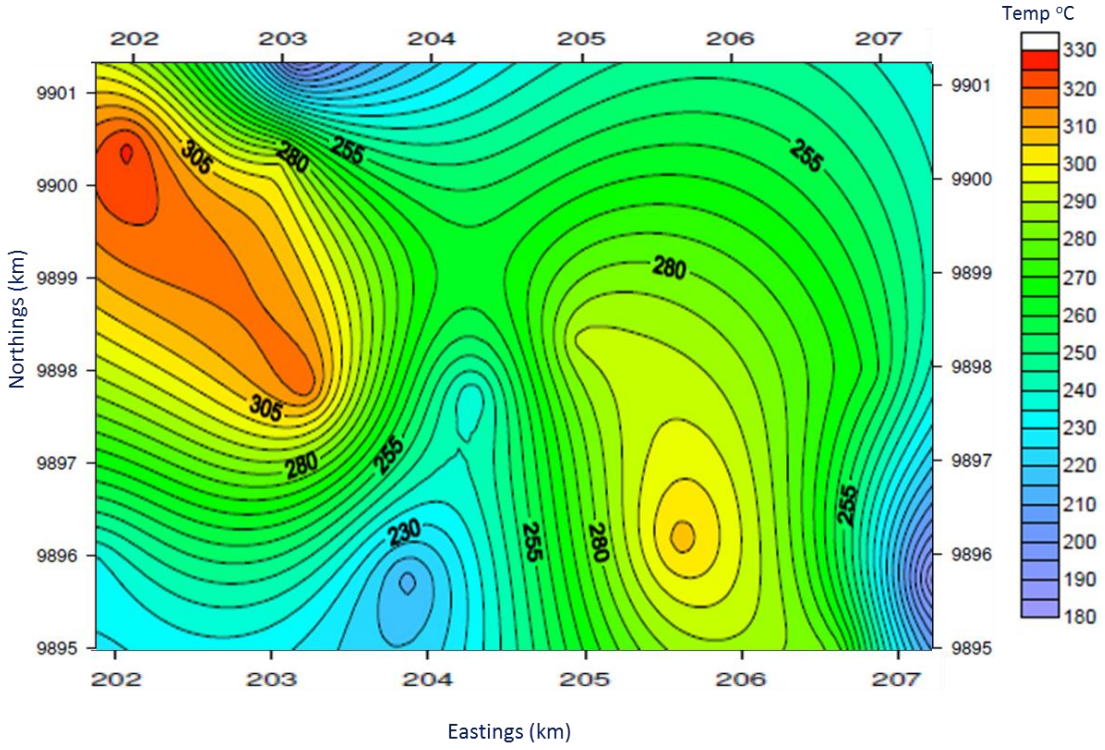


Figure 4.6: Lateral Temperature Contours at 0 m a s l

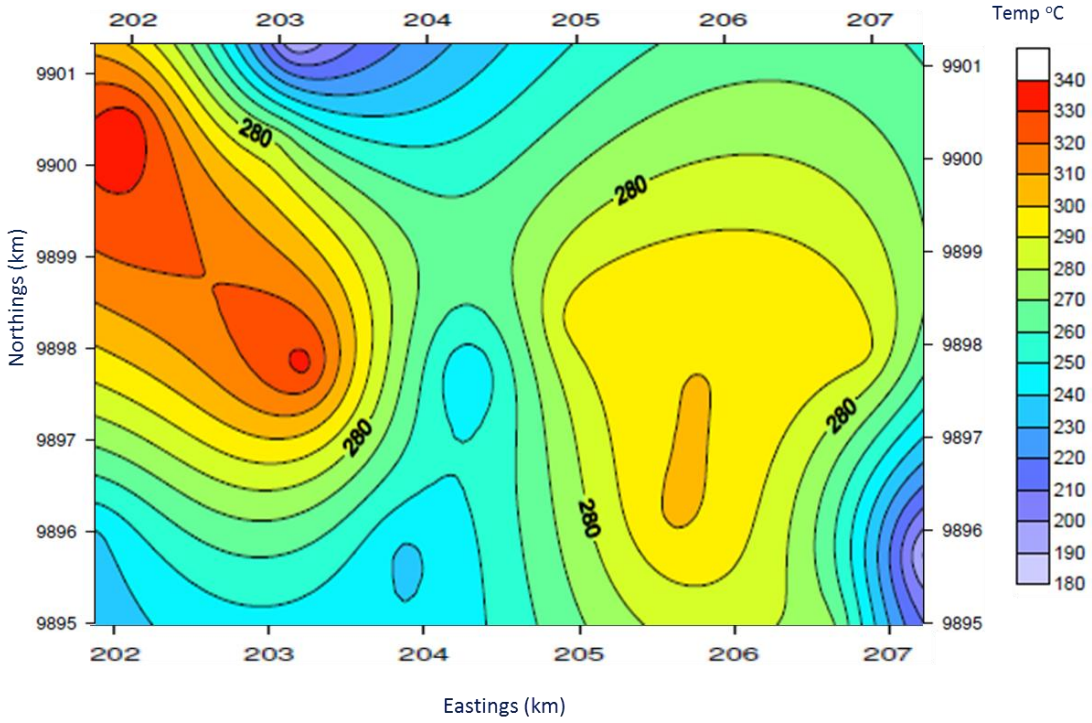


Figure 4.7: Lateral Temperature Contours at 200 m b s l

The two heat sources are separated by a low temperature region that runs in the NE-SW direction that coincides with the geological fault structure as shown in geological map of Figure 4.8.

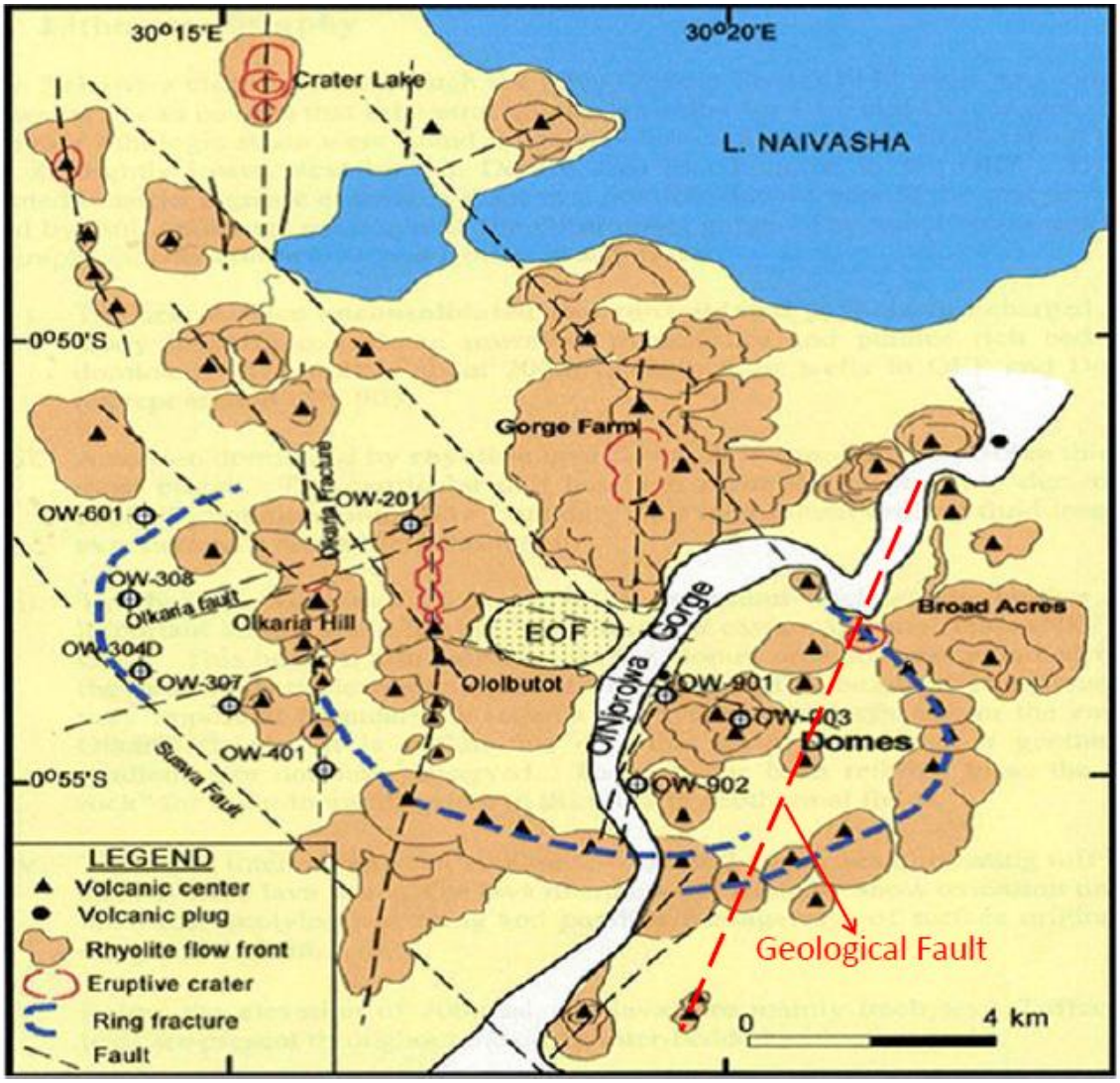


Figure 4.8: Geological Map of the Olkaria Volcanic Complex

Due to the low temperature nature of the NE-SW fault structure indicating cold fluids inflow, the fault could be controlling the fluid flow into the reservoir and therefore considered as the recharge to the reservoir.

4.1.2 Analysis of Nature of the Fault from Borehole Geology

Rock alteration entails transformation of the mineralogy of rocks as a result of the changes to the existing conditions such as temperature, chemical composition, tectonic setting and the interaction period that the rock is subjected to (Haldar, 2020). Temperature is the major determinant in the hydrothermal alteration as most of the chemical reactions do occur at elevated temperatures (Lagat, 2004). Also at elevated temperatures, minerals have been found to become thermodynamically stable.

Some of the favorable conditions for hydrothermal alterations are high temperature and high permeability to allow rock and fluid interactions. Fluids carry metals in solution, either from a nearby igneous source or from leaching out of some nearby rocks. This causes hydrothermal alteration of rocks by passing hot water fluids through the rocks and changing their composition by adding or removing or redistributing components. The primary minerals become unstable and forced to undergo chemical reactions with the hydrothermal fluids altering them to secondary minerals replacing the primary minerals (Nyandigisi & Katana, 2016). These secondary minerals tend to form at specific and stable temperatures thus making them suitable for mapping temperatures of a geothermal system and reservoir (Nyandigisi & Katana, 2016).

The alteration mineralogy and rock appearance were therefore used to investigate the nature of the fault in relation to the two magma chambers outlined from the temperature contours. To achieve this, alteration mineralogy and rock appearances in six wells that included OW901, OW902, OW905A, OW910, OW914B and OW917 were selected with the main material being the geological logging data obtained from rock cuttings during well drilling. Depth of the first appearance of Illites and Quartz minerals as well as Tuff and Rhyolites Rocks for each well was used with the surface elevation depth to find the depth above sea level of alteration. This depth and coordinates of the well were used to generate contour maps.

4.1.2.1 First Appearance of Quartz and Illite Minerals in Wells

Distribution of Quartz in Figure 4.9 that was generated from geological well logging data shows that in the Eastern region, Quartz was encountered at shallower depths of about 1600 m a s l compared to the western side where the first appearance of Quartz was as low as 1300 m a s l.

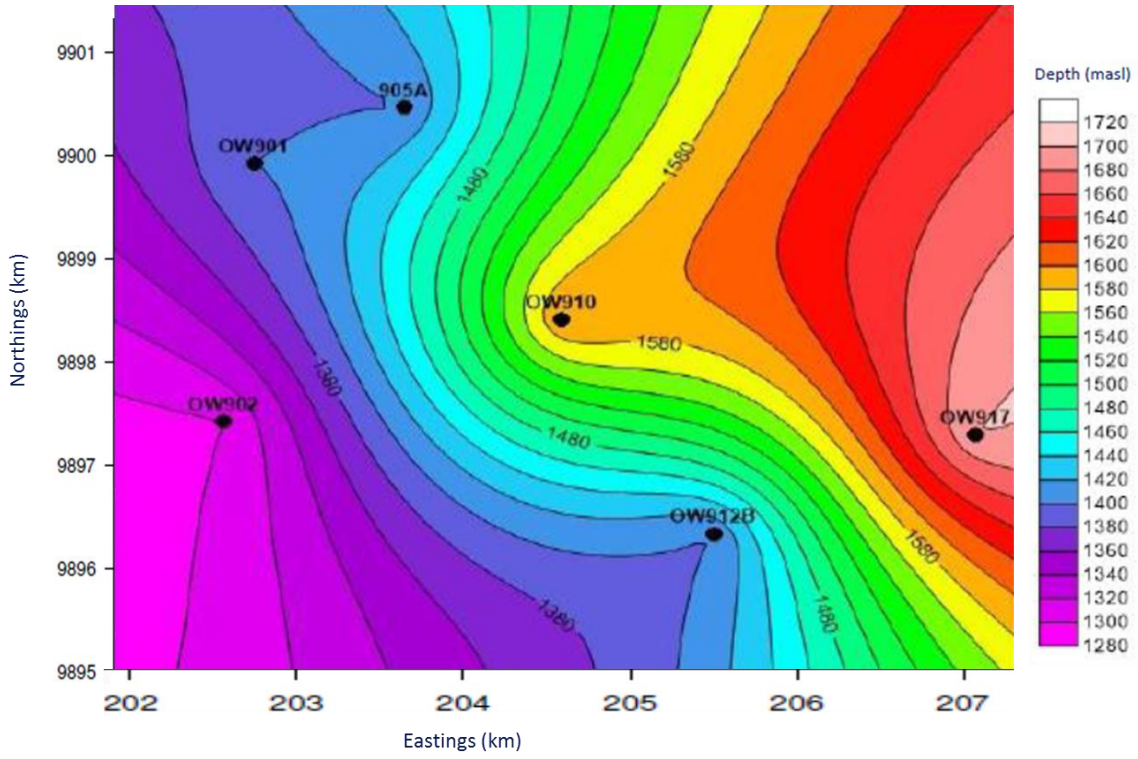


Figure 4.9: First Appearance of Quartz in Well Under Study

Illites appeared shallower on the SouthEastern side with the shallowest being at the depth of 1750 m a s l while the deepest was in the northwest side at 1150 m a s l as shown in Figure 4.10.

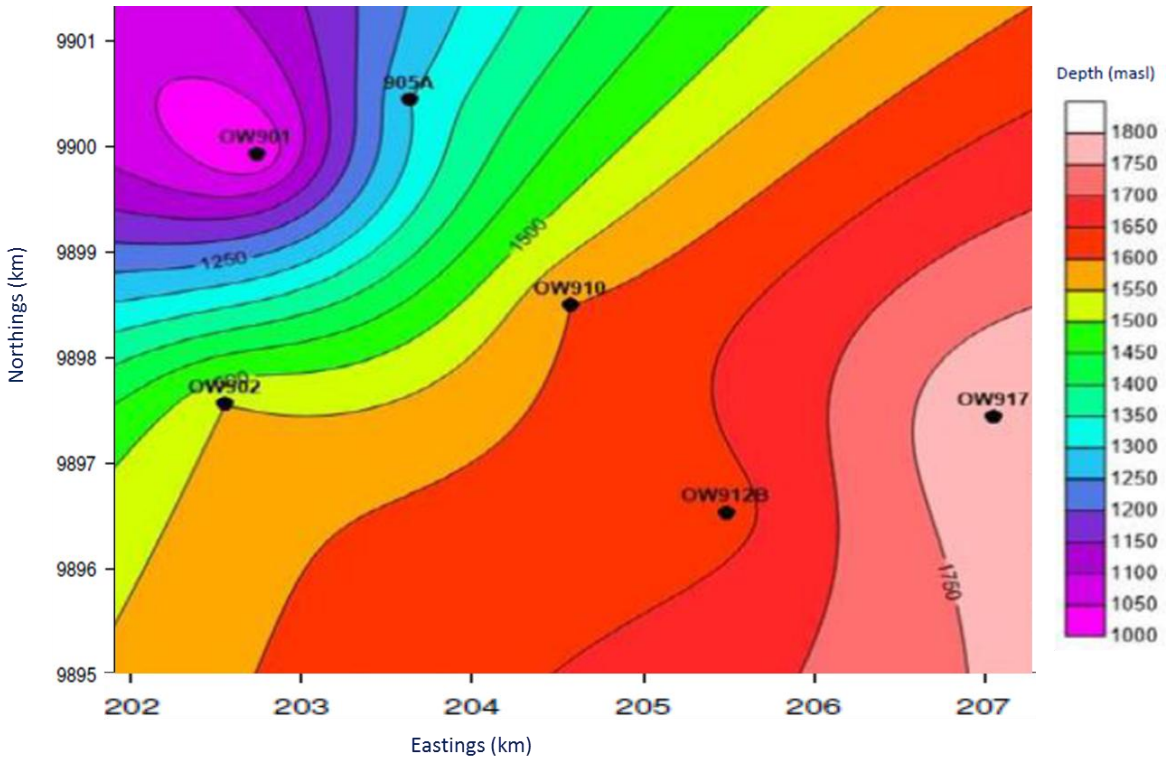


Figure 4.10: First Appearance of Illites in Wells under Study

The Illites and Quartz contour distribution indicate that minerals first appear closer to the surface on the Eastern region as compared to the Western region of the field hence high temperature zone characterized by the existence of high-temperature mineral assemblages in the eastern region. The presence of Quartz and Illites in a well indicate temperatures of over 180 ^{o}C and 200 ^{o}C respectively as shown in Figure 4.11 according to Fulignati (2020). Also, first appearance of Illites and Quartz in wells under study suggests that the two heat sources separated by the SW-NE fault are at different depths indicating that the heat source on the SouthEast side is closer to the surface compared to the Northwest side. Therefore SW-NE oriented fault could be as a result of more upthrow to the east as compared to the west leading to a normal fault.

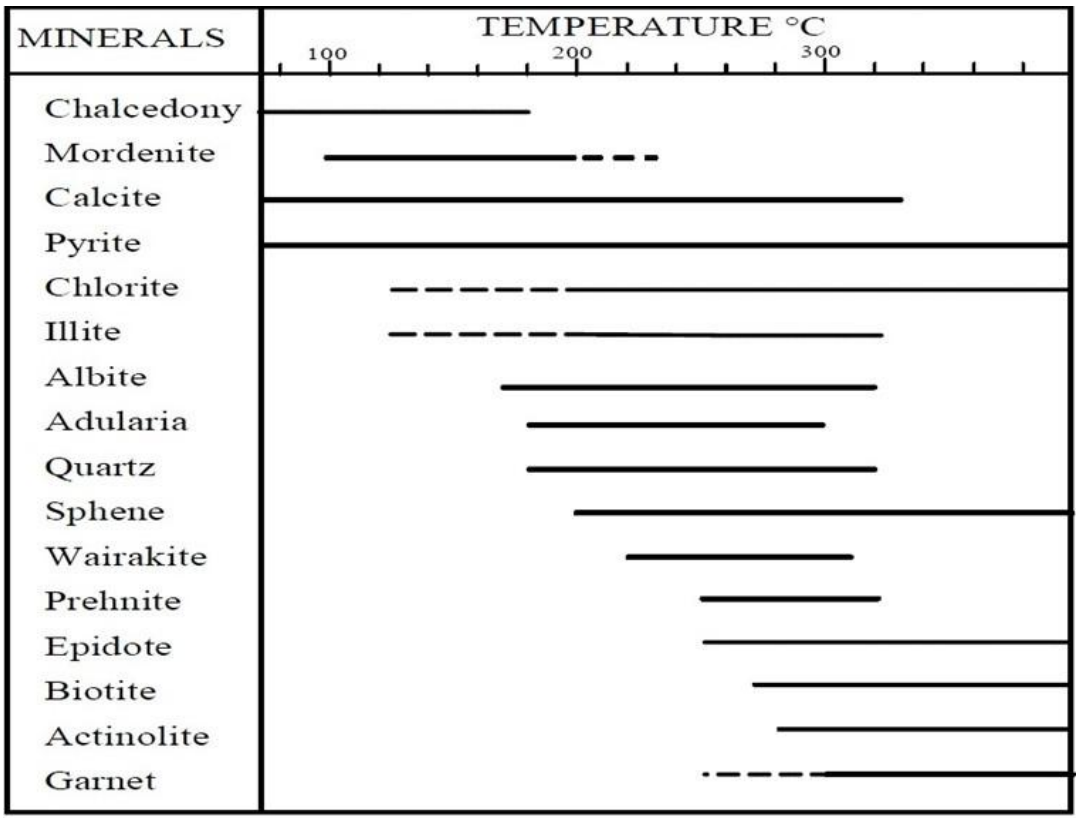


Figure 4.11: Hydrothermal Alteration Minerals and Their Temperature Stability Range (Fulignati, 2020)

This distribution trend of Illites and Quartz minerals correspond well with the SW-NE structure.

4.1.2.2 First Appearance of Tuff and Rhyolites Rocks in Wells

Distribution of tuff in Figures 4.12 and 4.13 that were generated from geological well logging data show that in the Eastern region, it was encountered at shallower depths of about 1900 m a s l compared to the western side where the first appearance of tuff was as low as 1340 m a s l.

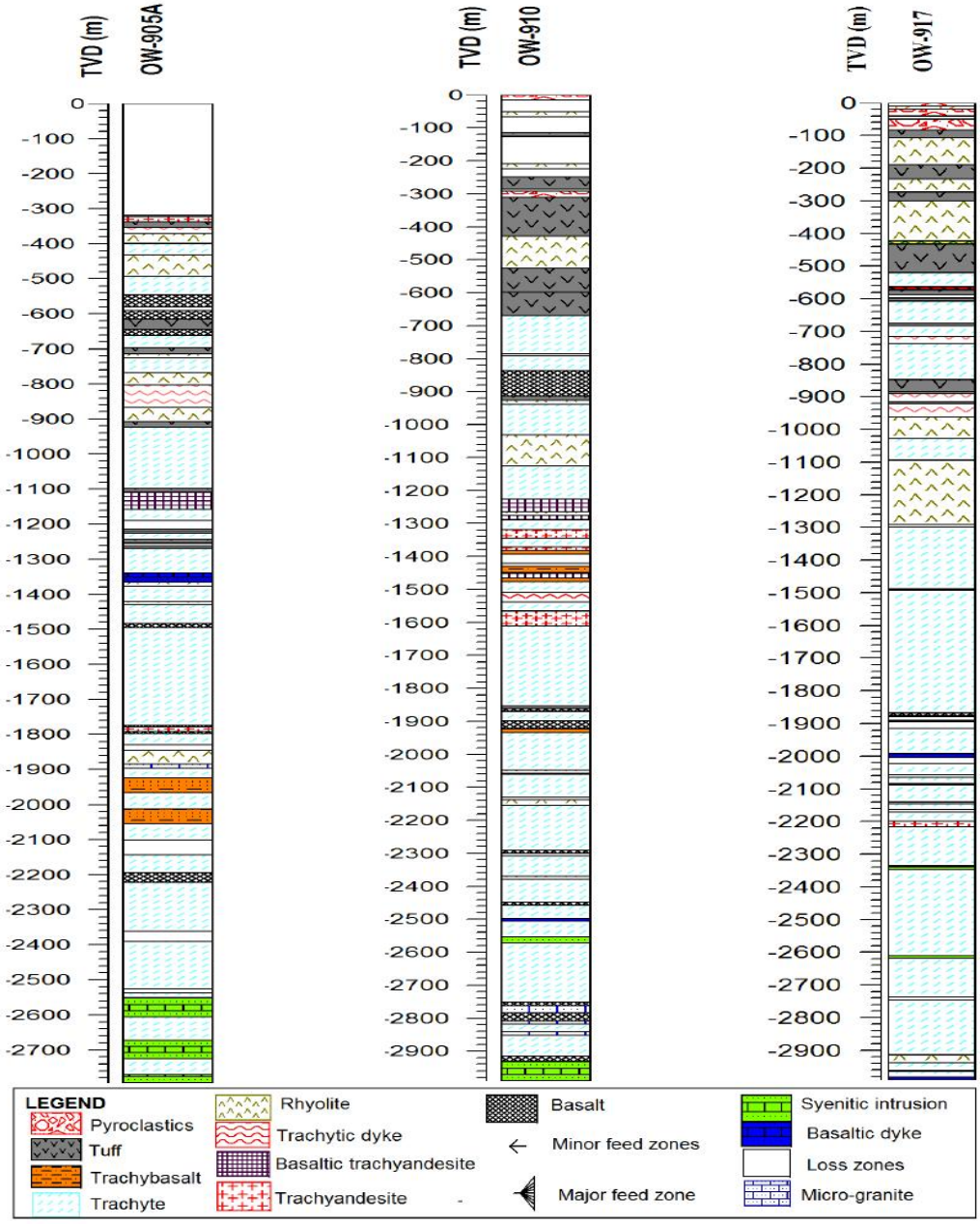


Figure 4.12: Rock Stratigraphy in Three Wells under Study

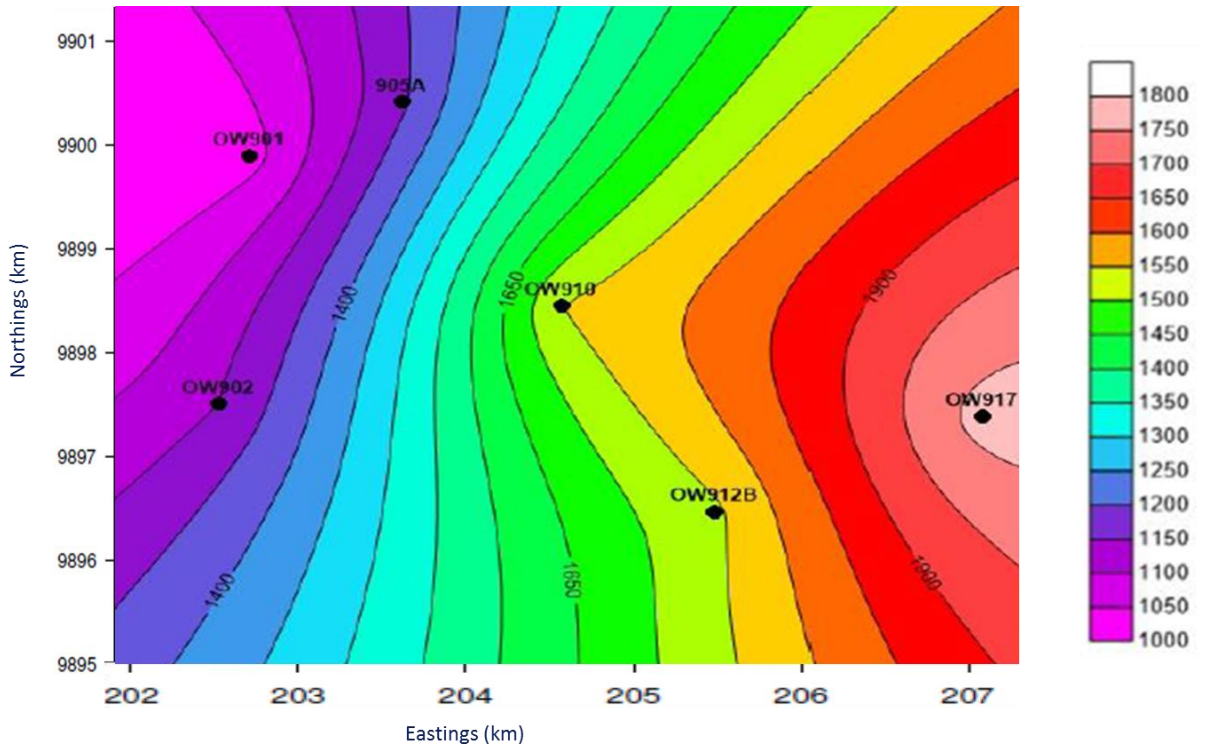


Figure 4.13: First Appearance of Tuff in Wells under Study

Rhyolites appeared shallower on the south-eastern side with the shallowest being in well OW917 while in the northwest side the shallowest in well OW905A m a s l as shown in Figure 4.14.

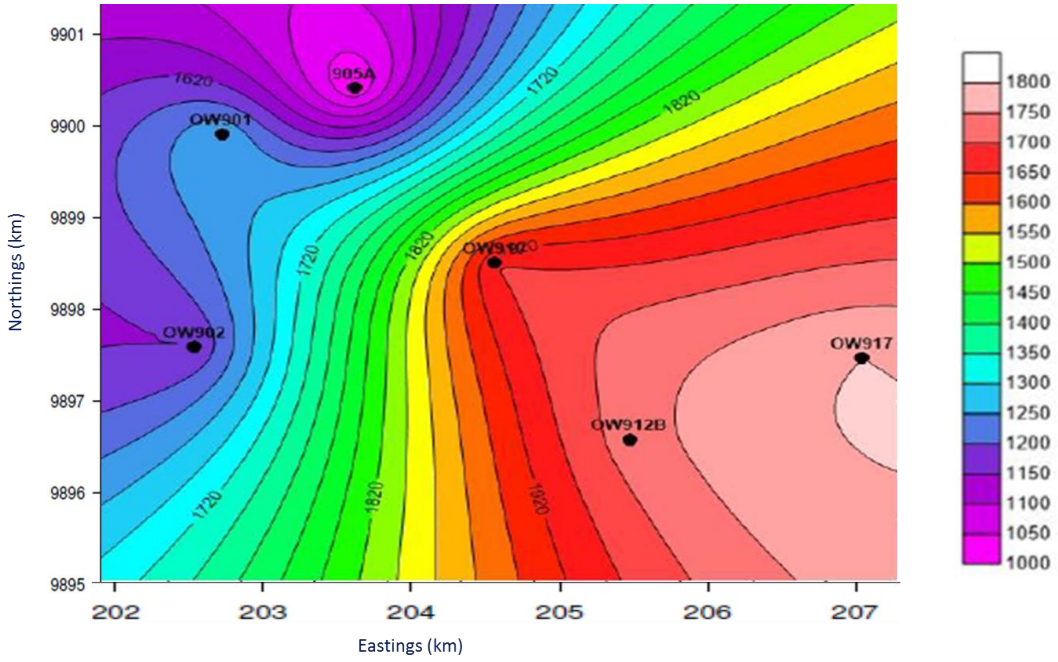


Figure 4.14: First Appearance of Rhyolites in Wells under Study

The Rhyolites and Tuff contour distribution indicates that these rocks first appear shallower on the SouthEast region as compared to the northwest region of the field. This indicates that there could have been a differential upward movement of magma chambers with the SouthEastern side moving more closely to the surface. This could also be the reason why Illites and Quartz minerals appeared shallower on the eastern region as compared to the western region.

4.1.3: Reservoir Heat Transfer Modes from Temperature Recovery Profiles

Considering individual well warm up profiles, it is observed that OW901, OW902B, OW903, OW905A, OW910, OW911A and OW919 are dominated by conductive heat transfer mode. This is indicated by gradual increase in temperature with increase in depth, with only convective heat transfer being observed at permeable zone at depth of 900-1300 m a s l where adverting hotter fluid is experienced as indicated in Figures 4.15 and 4.16 for wells OW905A and 902B respectively and appendix 1 for the rest of the wells.

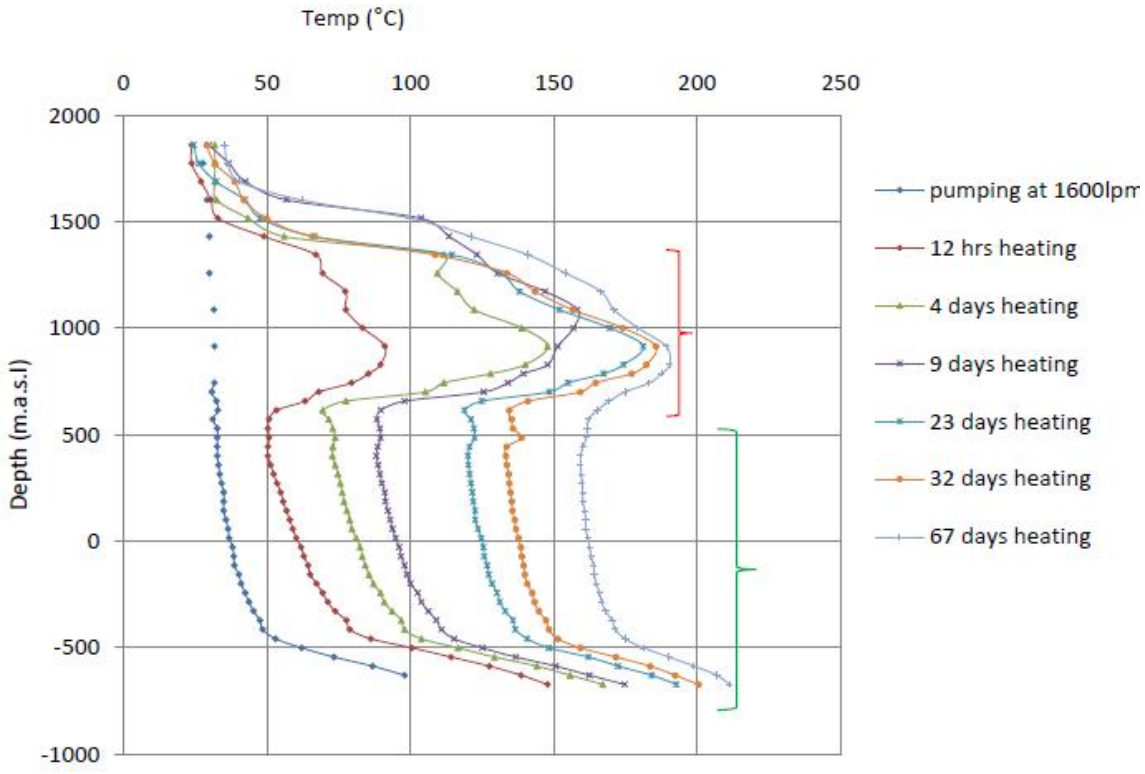


Figure 4.15: OW905A Heat Transfer Modes

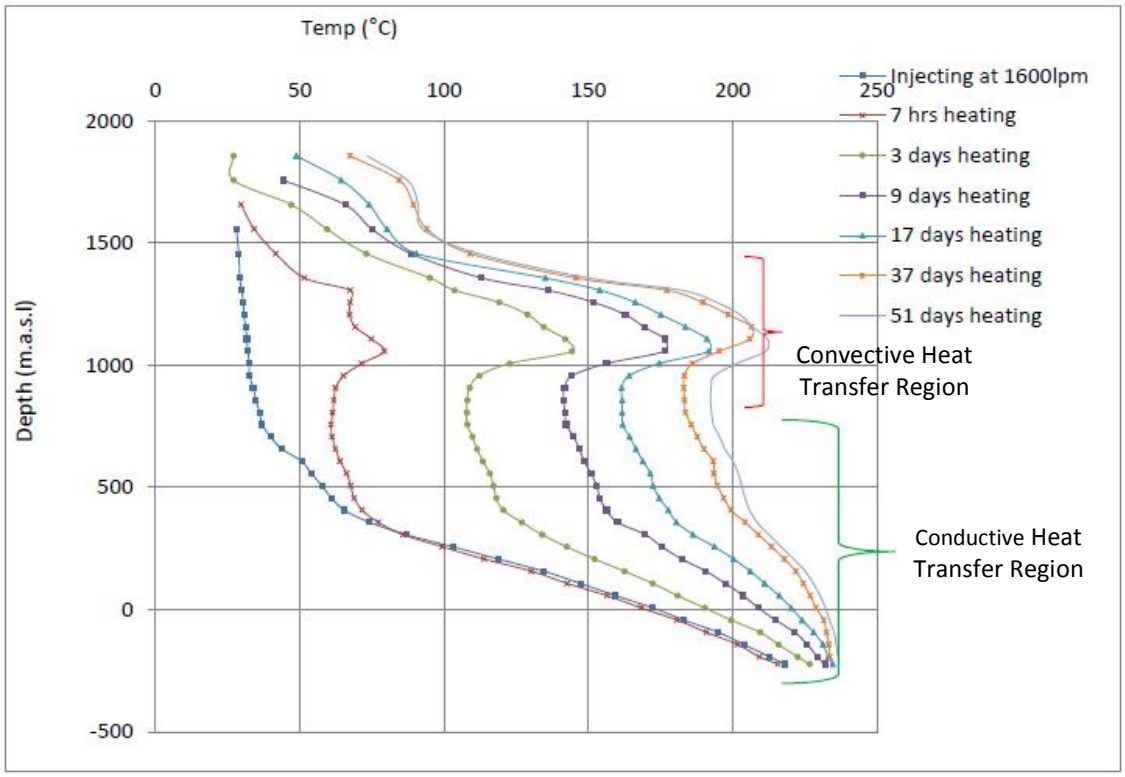


Figure 4.16: OW902B Heat Transfer Modes

Convective heat transfer mode dominates in wells OW908, OW912, OW914B, OW9015, OW917, OW918 and OW921. This is indicated by adverting hotter fluid dominating in most parts of these wells as indicated in Figures 4.17 and 4.18 for wells 912 and OW914B respectively and appendix 1 for the rest of the wells.

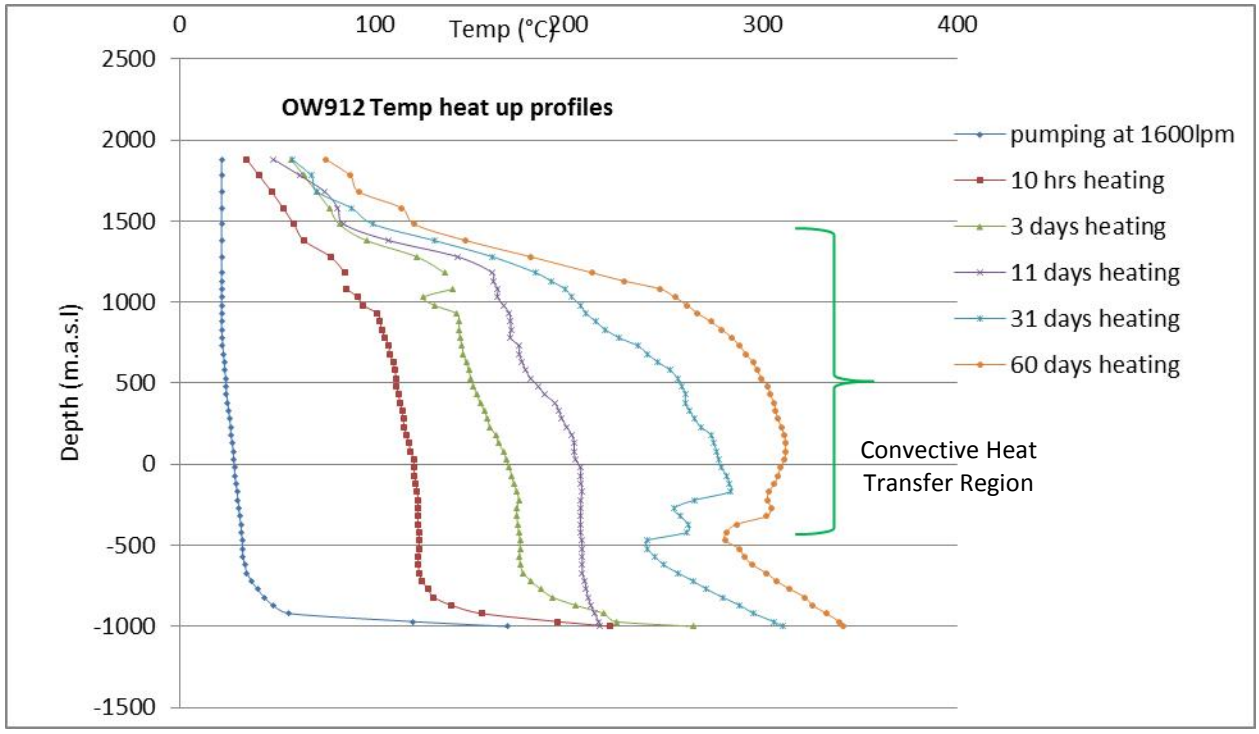


Figure 4.17: OW912 Heat Transfer Modes

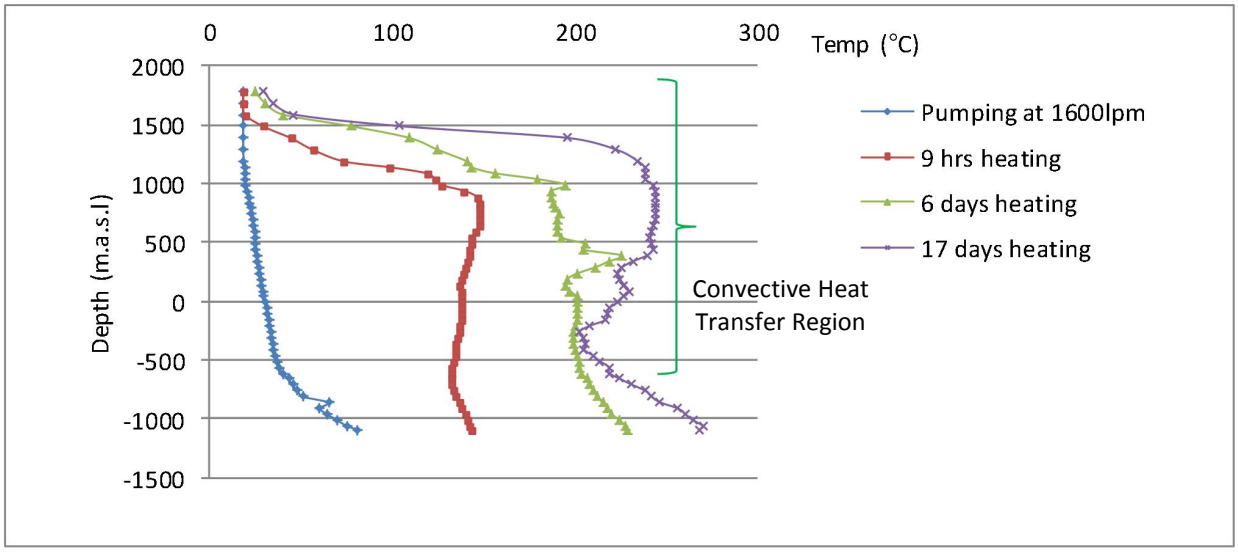


Figure 4.18: OW914B Heat Transfer Mode

All the wells in which convection heat transfer dominated were found to be located in the SouthEast while the wells with conductive heat transfer were found to be located in the northwest of the Olkaria Domes geothermal field as indicted in Figure 4.19. The blue rectangular boxes indicate conductive wells while black triangles indicate convective wells. This implying that the SouthEast region could be more permeable as compared to the northwest region since convection heat transfer happens only through the fluids which can only take place in permeable and porous medium.

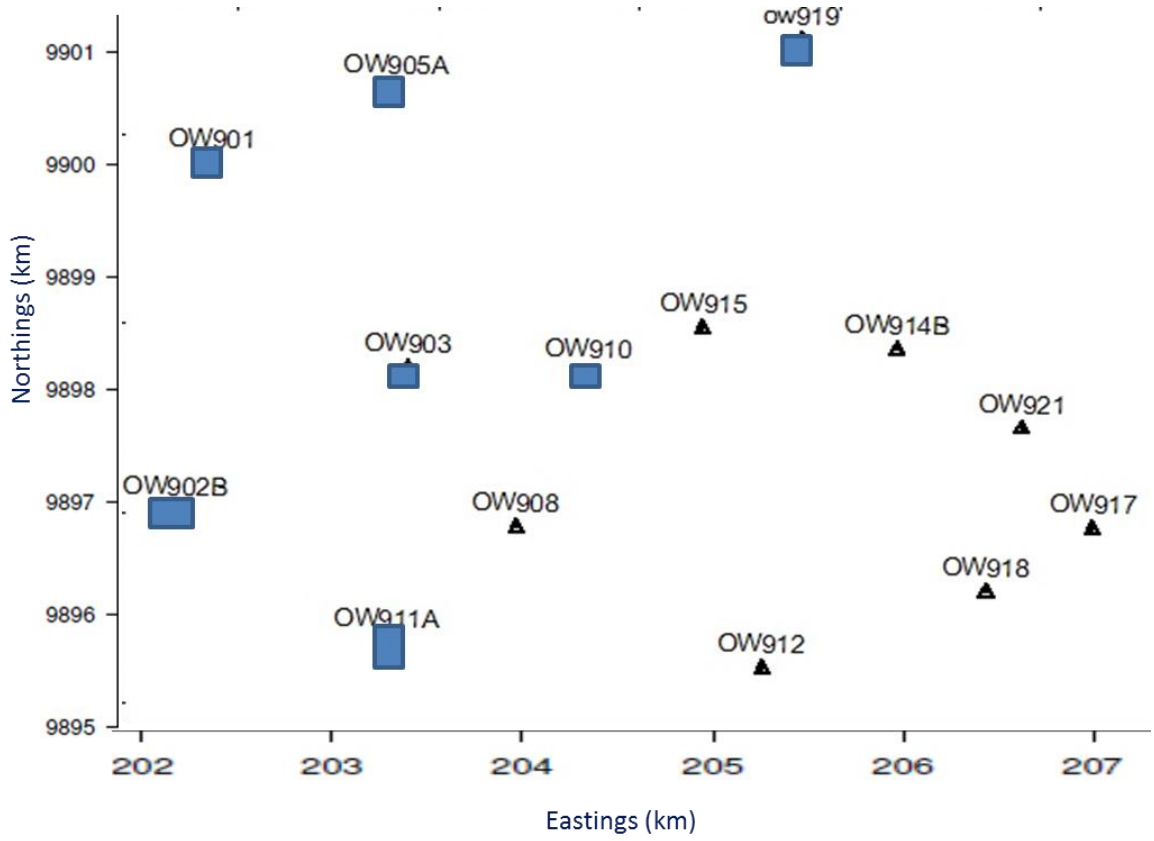


Figure 4.19: Convective and Conductive Heat Transfer Wells

4.1.4 Temperature Cross-Section

The vertical temperature cross-section model from west to east direction is shown in Figure 4.20. An up dome of temperature contours on the western side of the study area between eastings 202 and 203 indicates an upflow zone. The other upflow zone is found between eastings 204 to 207 Eastings. The map also indicates lateral recharge zone at a depth between 1500 m a s l and 1000 m a s l but extending deep into the subsurface at the fault line up to 800 m below sea level. These upflow zones could be associated with magmatic intrusions while the lateral recharge zone is associated with the fault line.

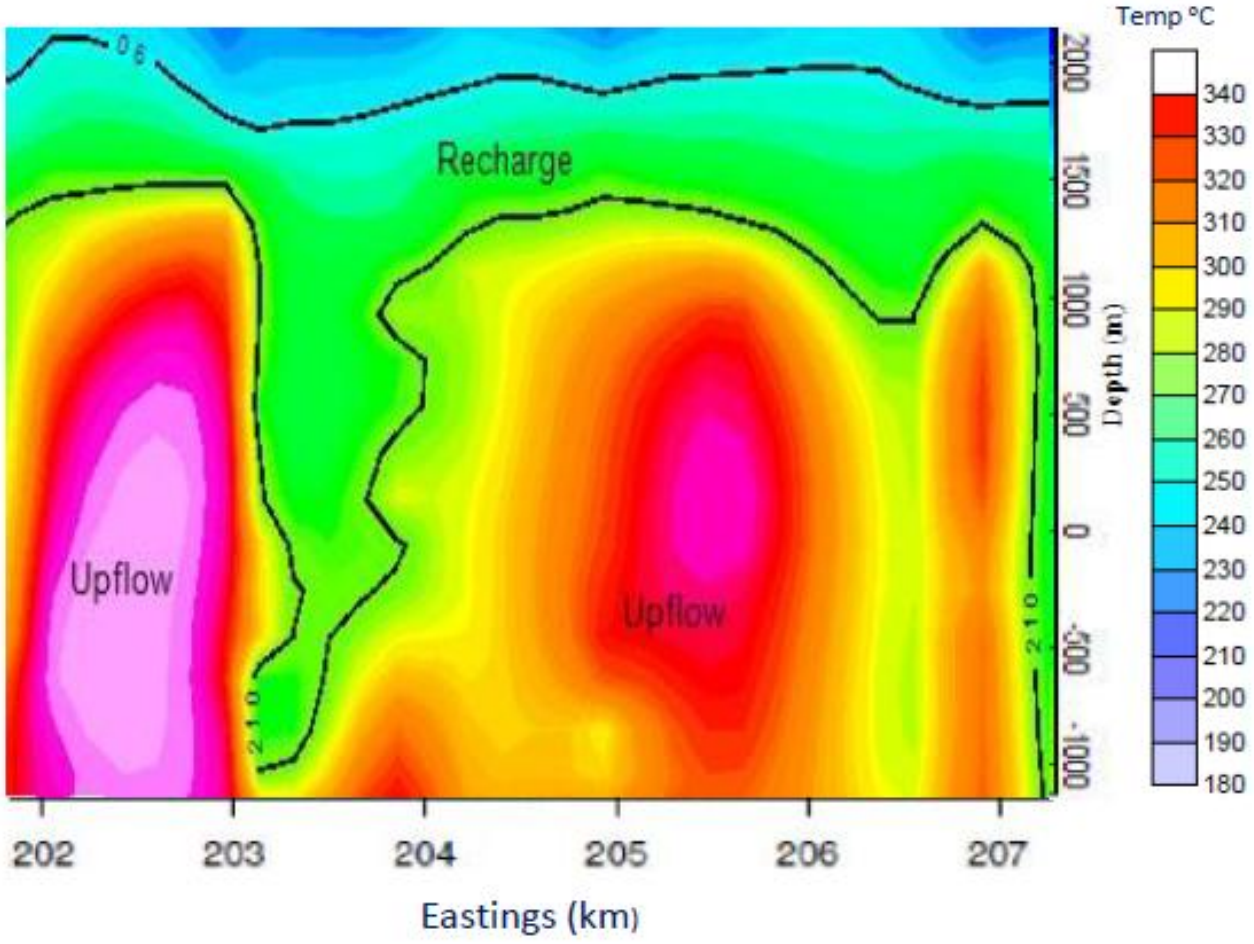


Figure 4.20: West - East Temperature Cross Section

4.1.5 Pressure Analysis from Step-Rate Injection Test

4.1.5.1 Injectivity Index

It is defined by the ratio of the change of injection flow rate to the change in reservoir pressure measured in the borehole, defined by equation 4.1.

$$\frac{\Delta Q}{\Delta P} = II \qquad 4.1$$

The last pressure values for each flow step of Figure 4.21 were plotted against the rates and best-fit lines drawn in Figure 4.22

.

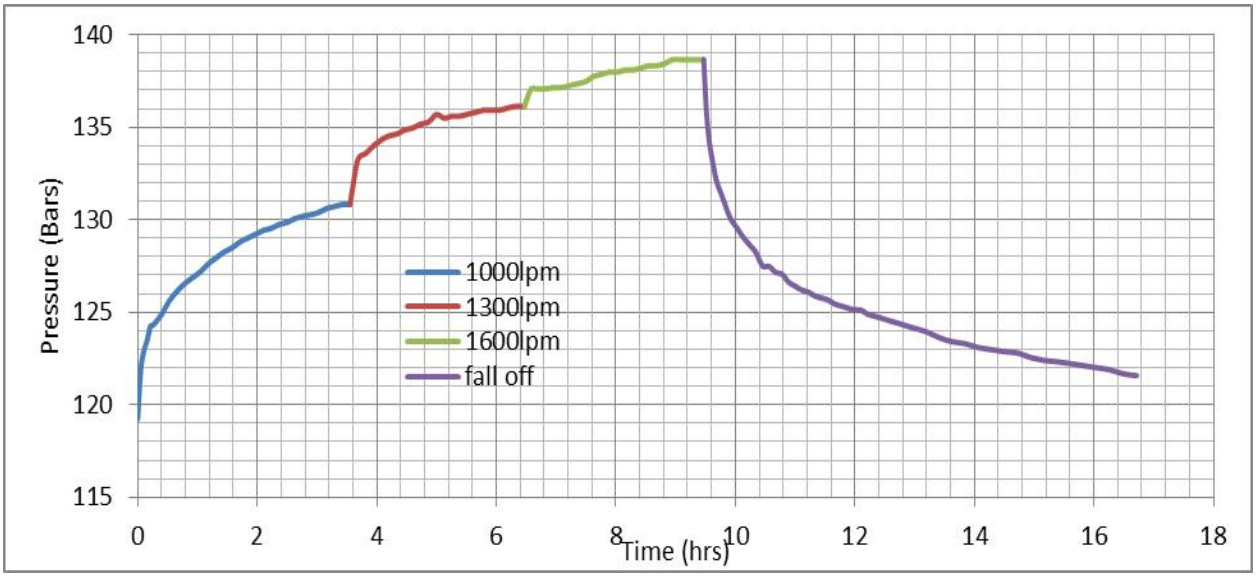


Figure 4.21: Pressure Buid Up and fall off Profile in OW901

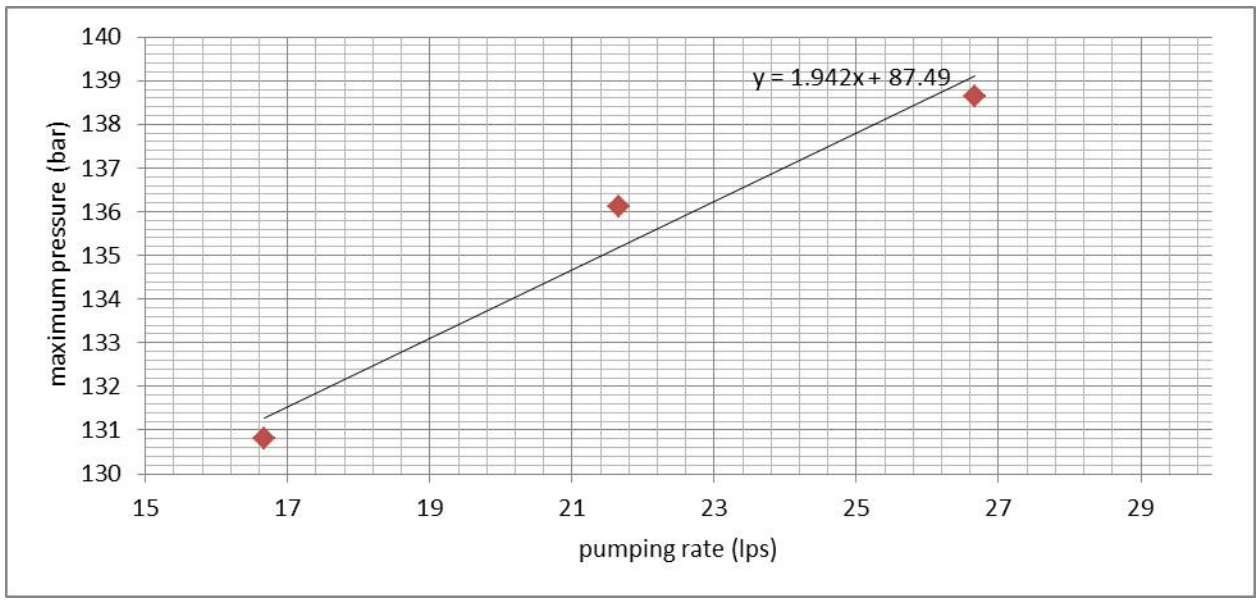


Figure 4.22: Injectivity Profile Plot for OW901

The inverse of the injectivity slope yielded injectivity of the OW901 well.

From the linear equation (injectivity curve)

$$Y = mx + C$$

Where *m* is the gradient

$$\text{Injectivity} = \frac{1}{m}$$

For OW901

$$Y = 1.942x + 87.5$$

Therefore,

m = 1.942

and

$$\text{Injectivity} = \frac{1}{m} = \frac{1}{1.942} = 0.51 \text{ lps/bar}$$

This was done to all the wells and the results are presented in table 4.1. For the rest of the wells under study, the pressure builds up and injectivity graphs are in appendix 2.

Table 4.1: Well Parameters

Well	Injectivity (lps/bar)	Transmissivity (m^3/pa.s)
OW901	0.51	8.28×10^{-9}
OW902B	2.94	7.82×10^{-9}
OW903	0.35	2.81×10^{-8}
OW910	3.13	1.17×10^{-8}
OW911A	1.79	5.63×10^{-9}
OW912	5.56	5.12×10^{-9}
OW914B	2.65	1.00×10^{-8}
OW915	2.78	9.38×10^{-9}
OW918	2.05	1.08×10^{-8}
OW921	3.13	1.00×10^{-8}

The results show the highest injectivity index 5.56Lps/bar in 0W912 while the lowest value of 0.35Lps/bar was recorded in 0W903. Also wells in the NE region of the reservoir have higher injectivity index while the SW wells have lower injectivity index. Contours for injectivity index were drawn as shown in Figure 4.23. From the results of this study, it can be said that the permeability of Olkaria Domes geothermal reservoir is not uniform; it varies for different parts of the reservoir. A well can be drilled into a high-temperature part of a reservoir but if the permeability is low, the well may not be productive. From Figure 4.23, injectivity index were generally higher in the Eastern region suggesting better the reservoir permeability and connectivity between the wells and the surrounding reservoir. This suggests that wells in the Eastern region of the reservoir could have intersected more permeable fractures.

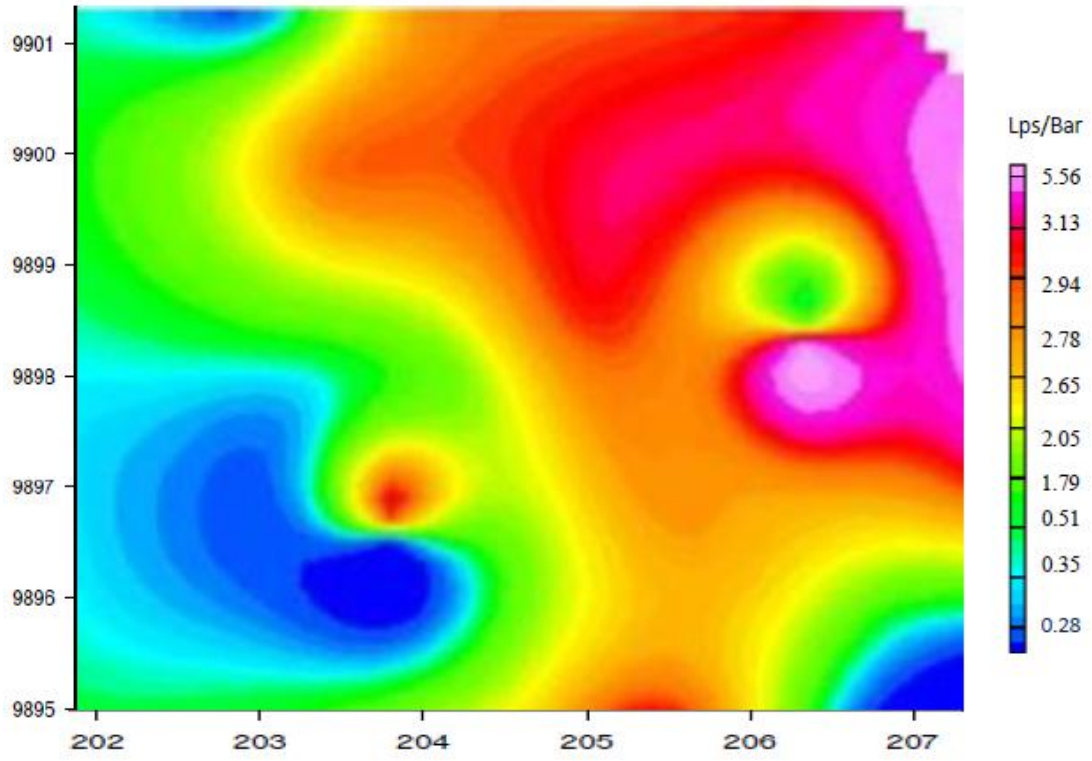


Figure 4.23: Wells Injectivity Distribution in Olkaria Domes Field

4.1.5.3 Semi-log Analysis: Transmissivity

A plot of the Theis solution for ΔPt vs. log t produced a semi-log straight line with a slope m per log cycle as shown in Figure 4.24 for well OW901 and appendix 3 for the rest of the wells. From equation 2.11

q = volumetric injection rate prior to fall-off

$$\frac{1900}{1000\times60} = 0.03167$$

m is the gradient of the semi-log straight line

For well OW901

m = 6.8 bars/log cycle

Hence Transmissivity is

$$\frac{2.0303 * 0.03167}{4\pi * 6.8}$$

T = $8.28\times10^{-09}m^3/pa.s$

This was done to all the wells and the results are presented in Table 4.1

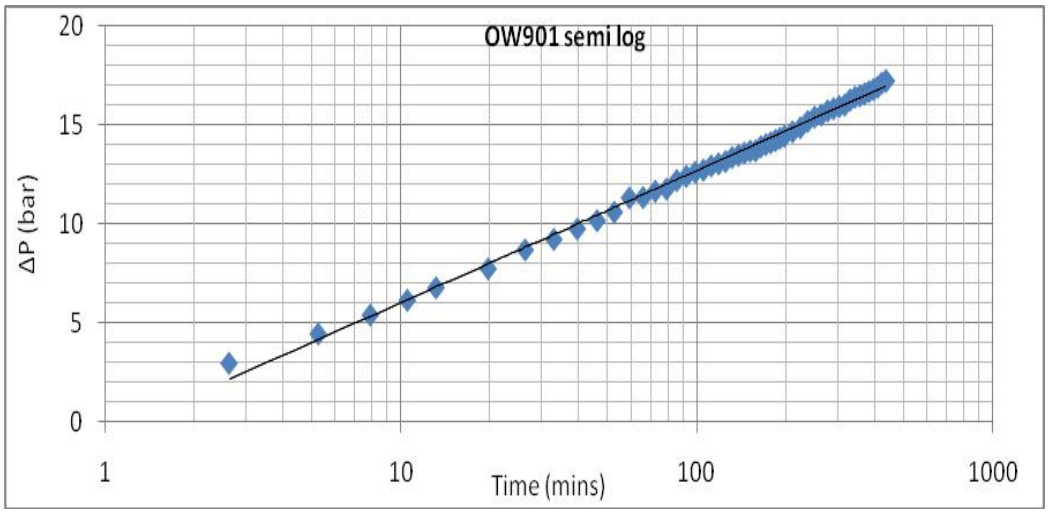


Figure 4.24: OW901 Semi-Log Graph

The results indicate that highest transmissivity values were 3.81×10^{-07} $m^3/Pa.s$ in 0W907A while well with lowest was 0W914B at 1.62×10^{-09} $m^3/Pa.s$. Wells in the eastern region of the reservoir have higher transmissivity while the NW and SE regions had the lowest transmissivity as compared to other regions in the study area.

These results could imply that wells in the in the eastern region of the reservoir could be more permeable as compared to the wells in the NW and SE.

4.2 MT and TEM Visualization

4.2.1 Resistivity Iso-Maps

Contour maps at different elevations from 1900 m a s l down to 2000 m b s l were constructed by TEMRESD program.

4.2.1.1 Resistivity Map at 1800-1700 m a s l Contours

This is shown in Figures 4.25 and appendix 7a. This covers a depth of 100 metres to 200 metres below the surface. The Figures show a dominant high resistivity region with resistivity ranging from 16 Ωm-1000 Ωm

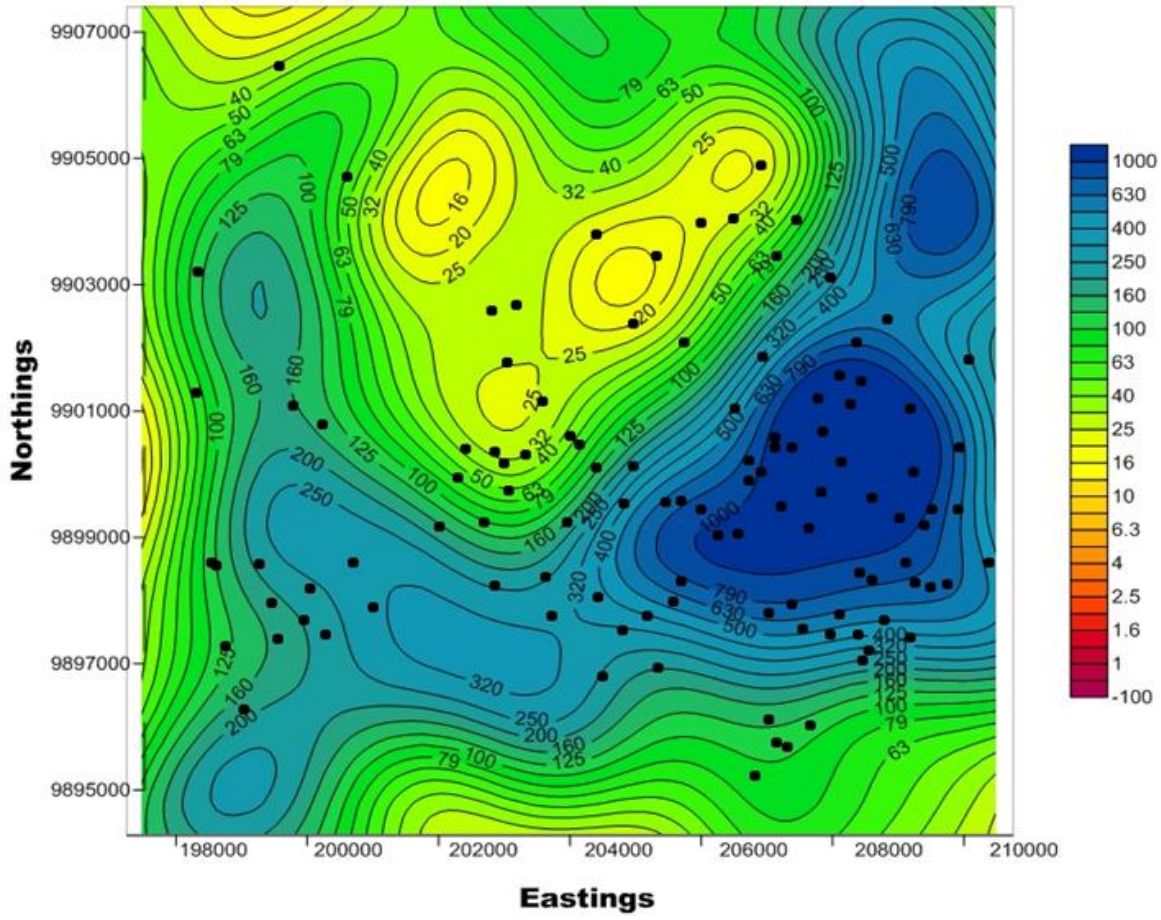


Figure 4.25: Resistivity Iso Map at 1800 m a s l

High resistivity in this region is an indication of un-altered subsurface formations near the earth's surface. The eastern region of the field exhibits the highest resistivity trending northwards and in the SW direction. This outlines the ring structure as indicated in the geological map in Figure 4.8. This could be as a result of resistive rhyolite rocks

that cover the top parts of the structure. The central part shows lower resistivity that shifts towards the west and northwest parts as we move down the surface.

In appendix 7a, high resistivity is concentrated to only two areas in the NE and central part of the East. The rest are low resistive region with the lowest being NW, W, S and SE. From the two Figures, resistivity decreases with depth

4.2.1.2 Resistivity Maps at 1600- 500 m a s l Contours

This is shown in Figures 4.26, 4.27and appendices 7b and 7c and covers a depth of 300 to 1200 metres below the surface. The Figures show a region of enhanced conductivity ranging from 1 Ωm-80 Ωm across the area especially in the western and eastern regions.

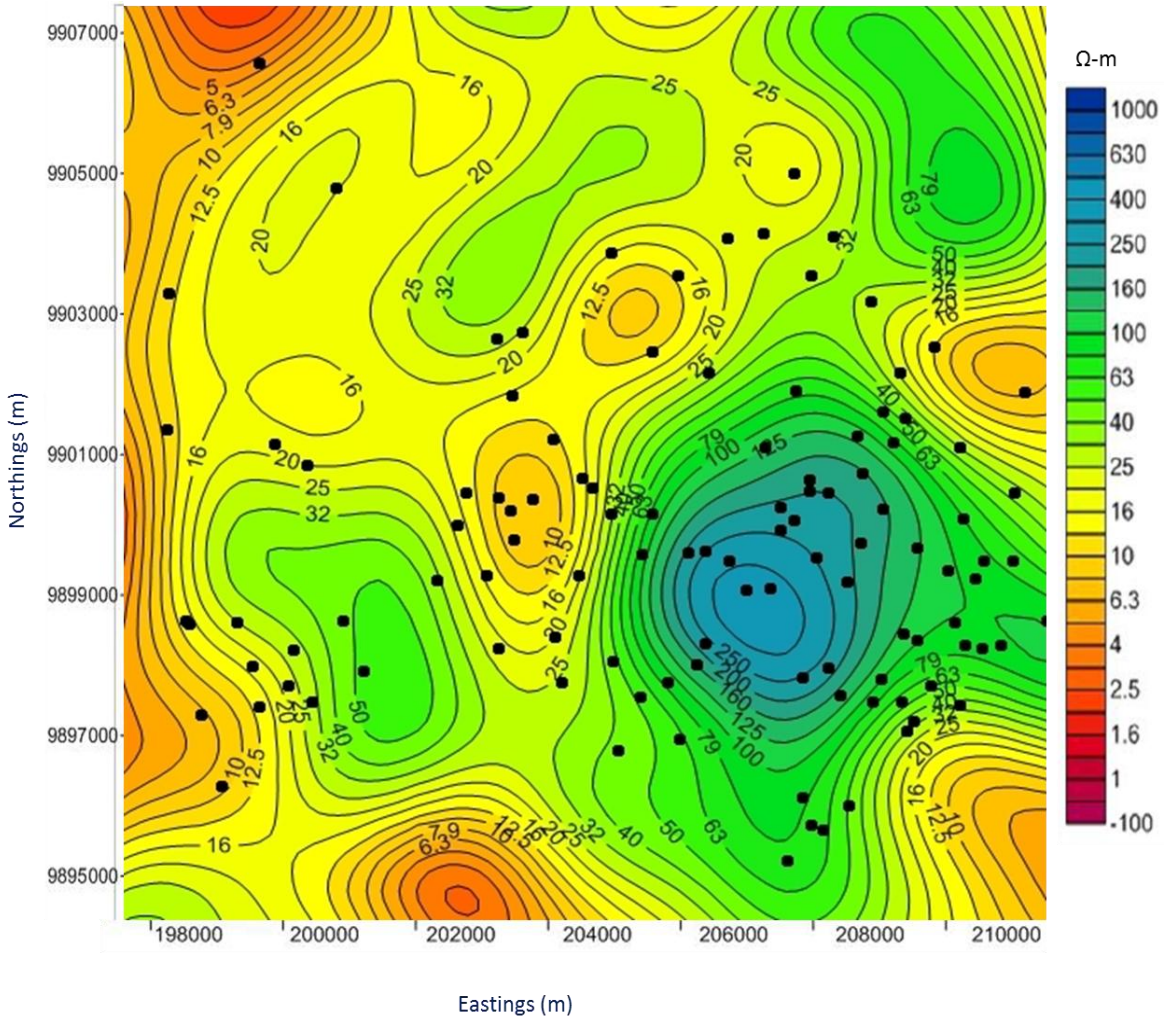


Figure 4.26: Resistivity Iso Map at 1600 m a s l

Relatively high resistive anomaly patches appear to emerge at the centre of the study area aligning in the NE - SW direction. This indicates a structural control that coincides with the geological fault structure in Figure 4.8. This high resistivity anomaly could be as a result of un-altered or presence of clay cap fills in the fault line.

High resistivity is observed in the SW region. This indicates patches of un-altered grounds at this depth. The western and eastern flanks have low resistivity also aligning

in the NE - SW direction. This is as a result of high permeability as shown in temperature recovery profiles (Figure 4.28) and low resistivity hydrothermal alteration minerals such as zeolites and smectites with their cations loosely bound are formed making the rocks conductive hence an indication of high hydrothermal fluid circulation exists at depth. The trend of resistivity decreasing with depth continues

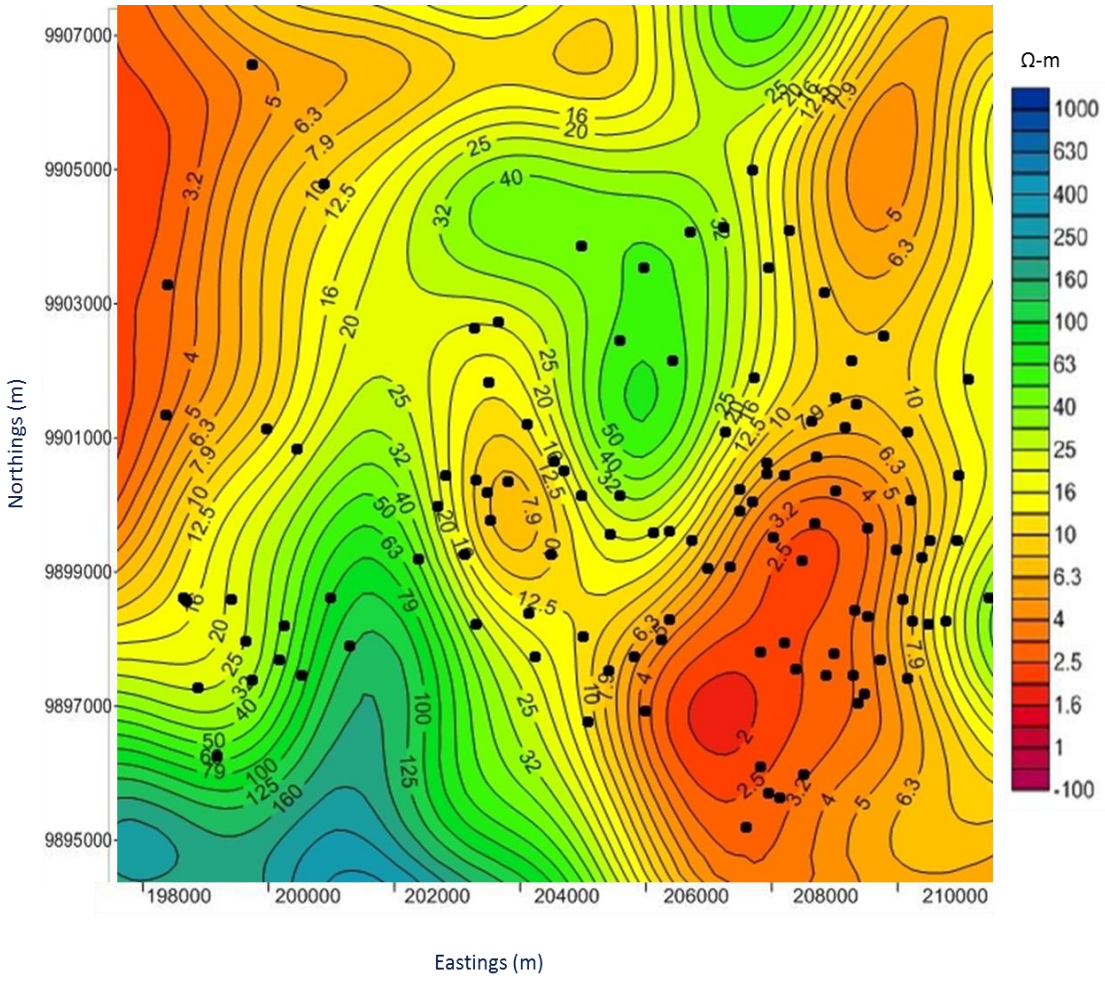


Figure 4.27: Resistivity Iso Map at 500 m a s l

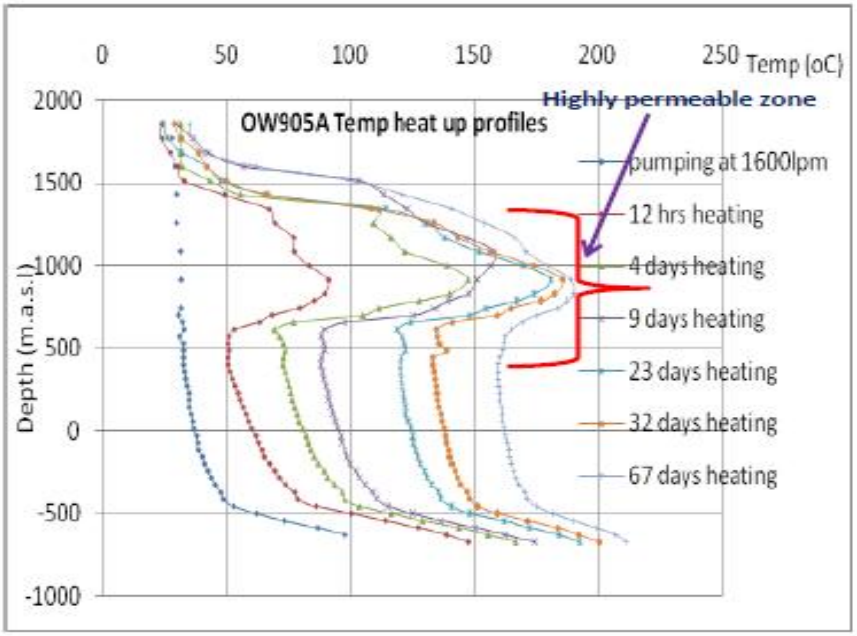


Figure 4.28: OW905A Temperature Profiles from the Heat Up Test

A low resistivity structure trending in the NW-SE direction appears from the depth of 1500 and 500 m a s l as shown in Figure 4.27 and appendix 47c cutting across the NE-SW trending fault. This could be a discontinuity in the fault line.

4.2.1.3 Resistivity Maps at 0- 2000 m b s l Contours

This is the region that covers 2200-4000 metres below the earth's surface. In this region, resistivity starts to increase with depth as shown in Figures 4.29 and 4.30 and appendices 7d and 7e. However western and eastern regions still show low resistivity as compared to other areas. The southern region is the most resistive.

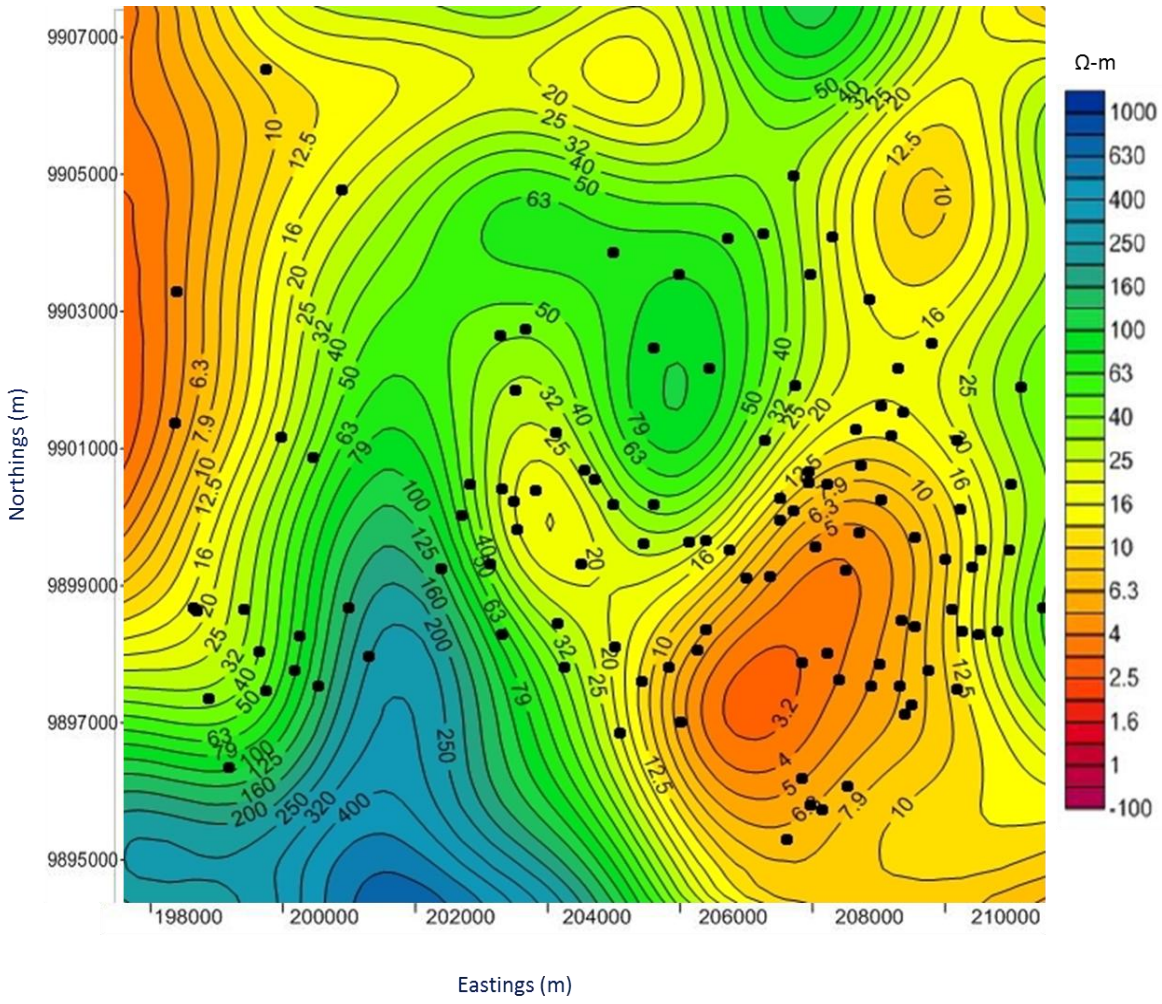


Figure 4.29: Resistivity Iso Map at 0 m a s l

The increase in resistivity with depth indicates the presence of high temperatures that has resulted into the formation of high temperature alteration minerals. This is because epidote and chlorite minerals with bounded ions in the crystal lattice dominate at higher temperatures exceeding 250 °C making the system more resistive.

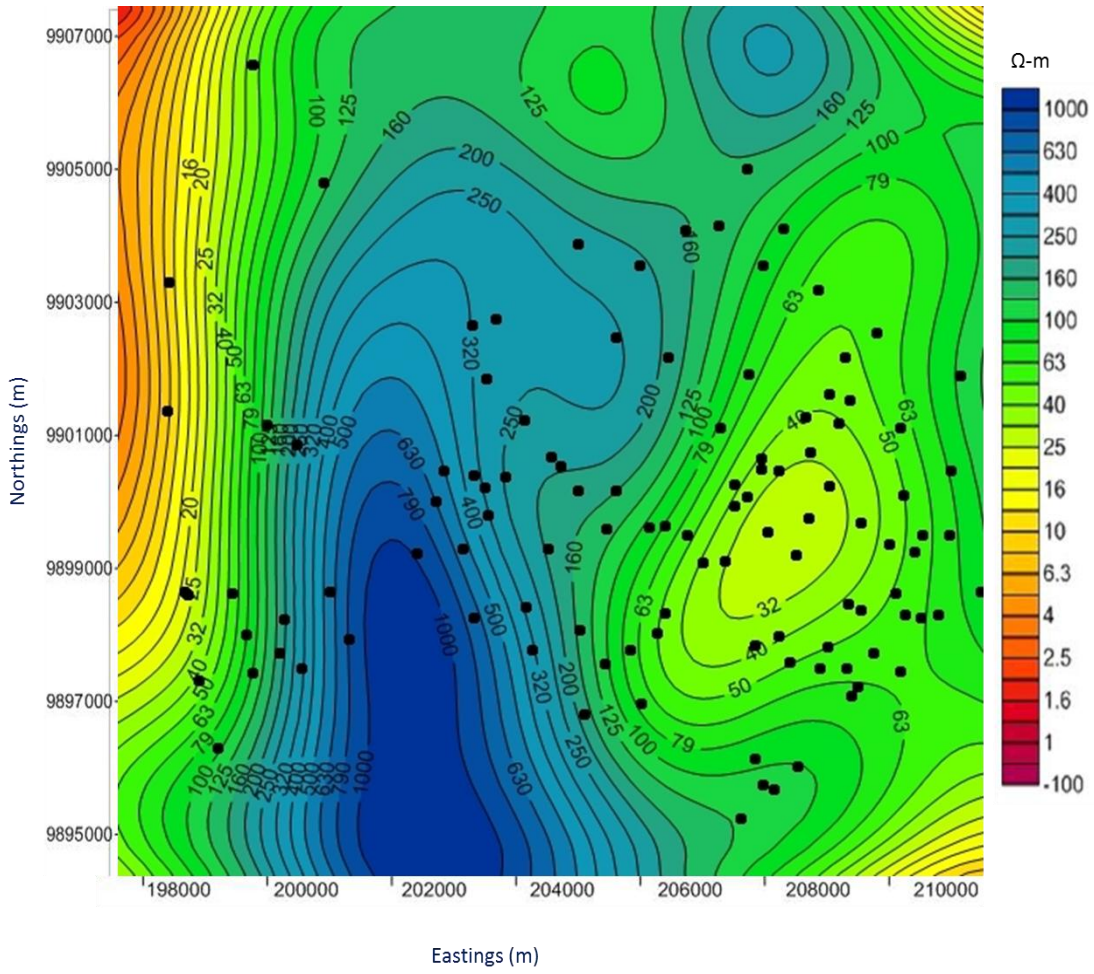


Figure 4.30: Resistivity Iso Map at 2000 m b s l

4.2.3 Resistivity Cross-Sections

Two cross-sections were generated by TEMCROSS program, the NW-SE and the SW-NE Cross-Sections as shown in Figure 4.31

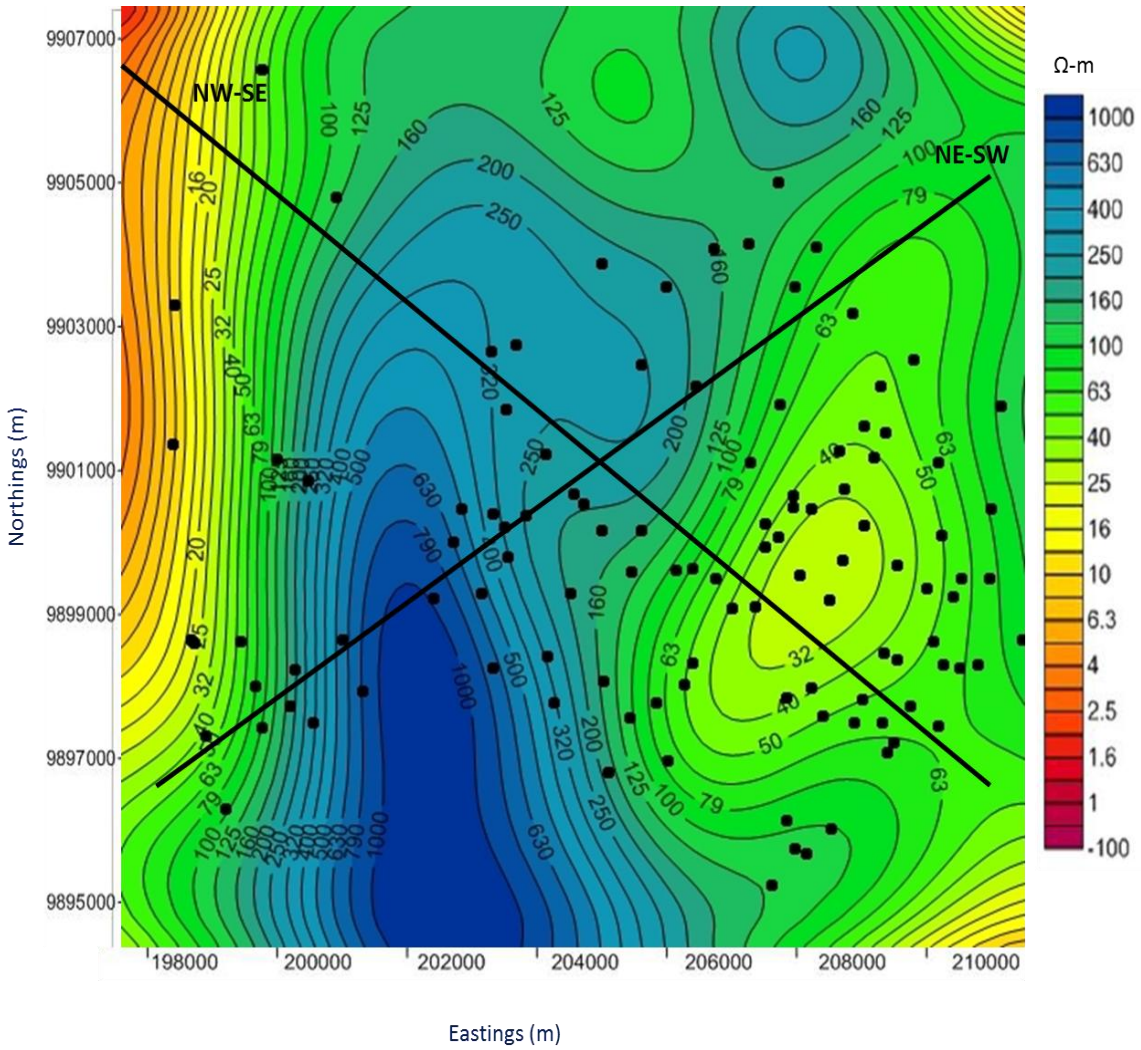


Figure 4.31: Cross Section Map.

4.2.3.1 NW-SE Resistivity Cross-Section

The NW-SE Resistivity Cross-Section depicts non-uniform resistivity structures as shown in Figure 4.32. This could be as a result of localized rock distribution in the area. At the surface, a narrow layer of high resistivity was realized. This could be as a result of unaltered rock formations. A higher resistive structure R1 on the surface indicates non uniform distribution of unaltered rhyolitic rocks lava flows and other eruptive materials such as ashy minerals. This could be an indication of poor water saturation at R1.

Another broader layer of low-resistivity elevated at the centre of the study area was observed. This could be attributed to the presence of low resistivity hydrothermal alteration minerals such as zeolites and smectites with their cations loosely bound that are formed making the rocks conductive hence an indication of high hydrothermal fluid circulation exists at depth. Conductivity zones C1 and C4 are indication of elevated permeability in these regions and therefore there could be a high hydrothermal fluid circulation. The resistivity discontinuity between C1 and C4 reflects a SW-NE trending fault line as shown in geological map of Figure 4.8 that could be as a result of un-altered or presence of clay cap fills in the fault line.

The third layer is dominated by relatively higher resistivity follows at depth. The high resistivity was attributed to the dominance of high temperature alteration minerals such as epidote and chlorite minerals with bounded ions in the crystal lattice. West and East sides of the reservoir are dominated by low resistive regions at depth indicating high permeability in these regions.

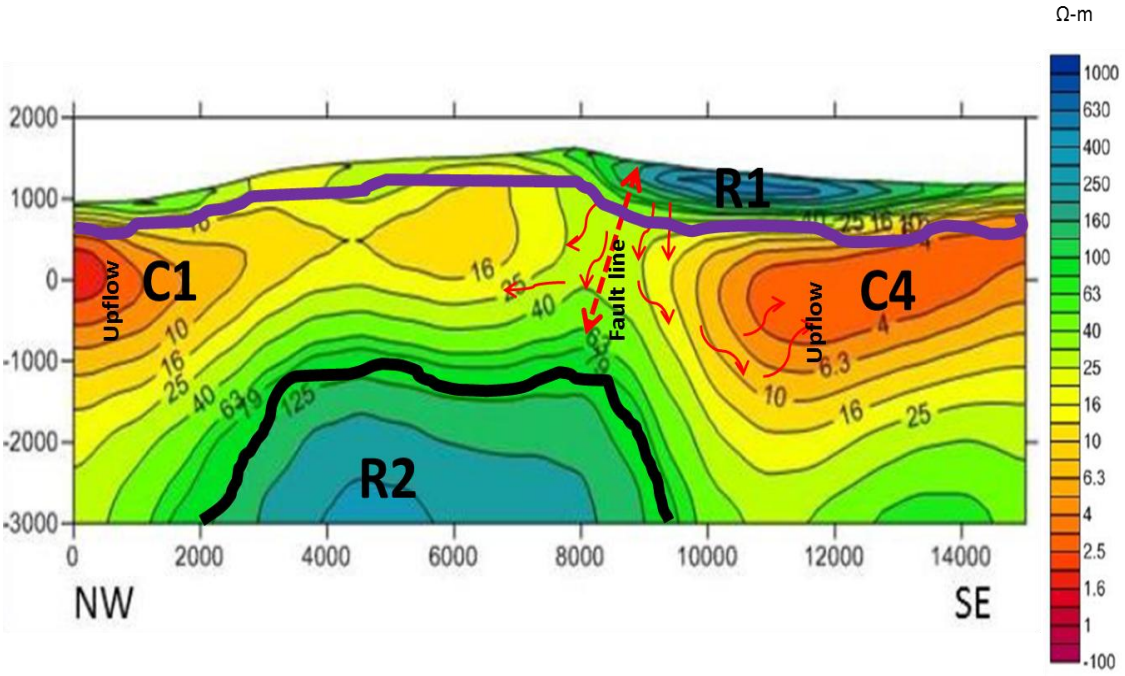


Figure 4.32: NW-SE Resistivity Cross Section

4.2.3.2 SW-NE Resistivity Cross-Section

Similarly, SW-NE Resistivity Cross-Section of Olkaria dome in Figure 4.33 revealed similar features as in Figure 4.32 with a non-uniform resistivity structure at depth and a narrow layer of high resistivity at the surface as a result of unaltered rock formations. It was also followed by another broader layer of low-resistivity layer with three main patches (west, central and east). Low resistivity within this region was attributed to the presence of low resistivity hydrothermal alteration minerals such as zeolites and smectites with their cations loosely bound are formed making the rocks conductive. A relatively higher-resistivity zone followed due to dominance of high temperature alteration minerals such as epidote and chlorite. West and East sides were dominated by low resistive at depth. High resistive body at the centre from the depth of 1250 m a s l downwards which is believed to be the heat source was observed.

In Figure 4.33, conductivity regions C1, C2 and C3 indicate elevated permeability in these regions and therefore there is a high hydrothermal fluid circulation. Resistivity discontinuity between C1 and C2 is an indication of a possible fault.

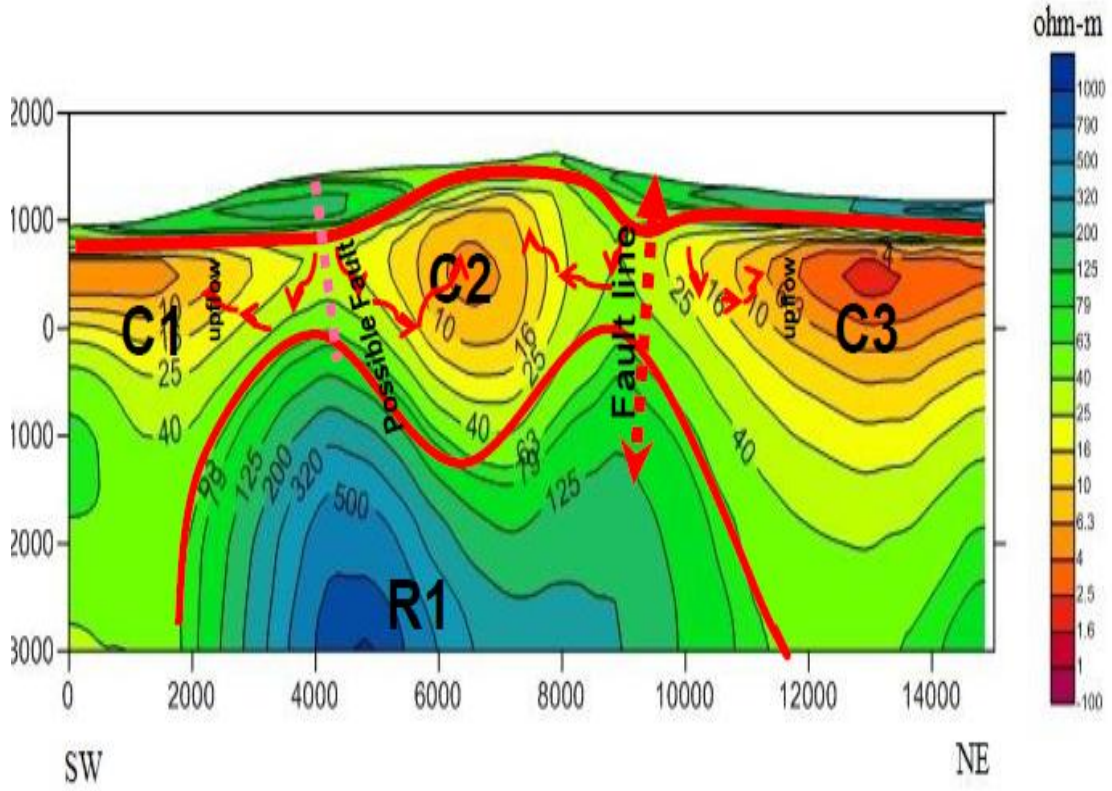


Figure 4.33: SW-NE Resistivity Cross Section

4.3 Predictive Modelling

The objective was to obtain independent input values from the geophysical resistivity data and use them to predict the subsurface temperature variables. The implication is that it would be useful to use the computer's ability to learn from data and predict the outcome, hence the use of machine learning. Regression analysis was used to help understand how the dependent variables would change their values on light of the variance of the independent variables. While there are many approaches through which this problem could have been tackled such as classification and clustering, they were not put into consideration based on the nature of the problem at hand. Considering that Machine Learning principal is to enable the machine to learn from the input data and be able to predict the output data, the nature of the input and desired output plays a great role in defining which paths to take in terms of algorithms and process. Looking into the

data at hand, the desired output which is to be in terms of temperature values means that the output is known hence, supervised learning was chosen. Also, the output temperature values were continuous values rather than discrete ruling out classification; given that the latter answers the question of as to which category a new observation may belong to. In other words, in classification, input labels are mapped out to output labels whereas in regression we are mapping input to continuous output.

Having established regression as the Machine Learning method to solve the problem at hand, the specific algorithm convenient for solving the problem also needed to be established. Linear regression was ruled out because the relationship between the dependent and independent variables is not linear. Therefore, other regression approaches comprising of polynomial Regression, Decision Tree Regression, Adaptive Booster Regression, Support Vector Regression and Random Forest Regression were applied.

4.3.1 Exploratory Data Analysis

Before constructing the model, the exploratory analysis of the dataset was performed with the aim of bringing out the summary of main characteristics of a dataset to understand what the data yields beyond the formal modeling or hypothesis testing task. The dataset was made up of 297 Observations and two features, resistivity and Temperature, of data type float. The statistical values for the two features are shown in Table 4.2.

Table 4.2: The Dataset Summary Information

	Resistivity (ohm-m)	Temperature (°C)
Count	297	297
Mean	56.26	215.49
Std	95.03	78.36
min	3	24.5
25%	11	186.6
50%	20	225.1
75%	51	269.1
max	630	364

The histogram in Figure 4.34 was used to show the nature of statistical data for resistivity. The histogram shows that the data was asymmetrical and skewed toward the left with a tail toward the right, Most of the resistivity data values lie between the resistivity values between 0 to 200 Ω-m and very few points lying above 200 Ω-m.

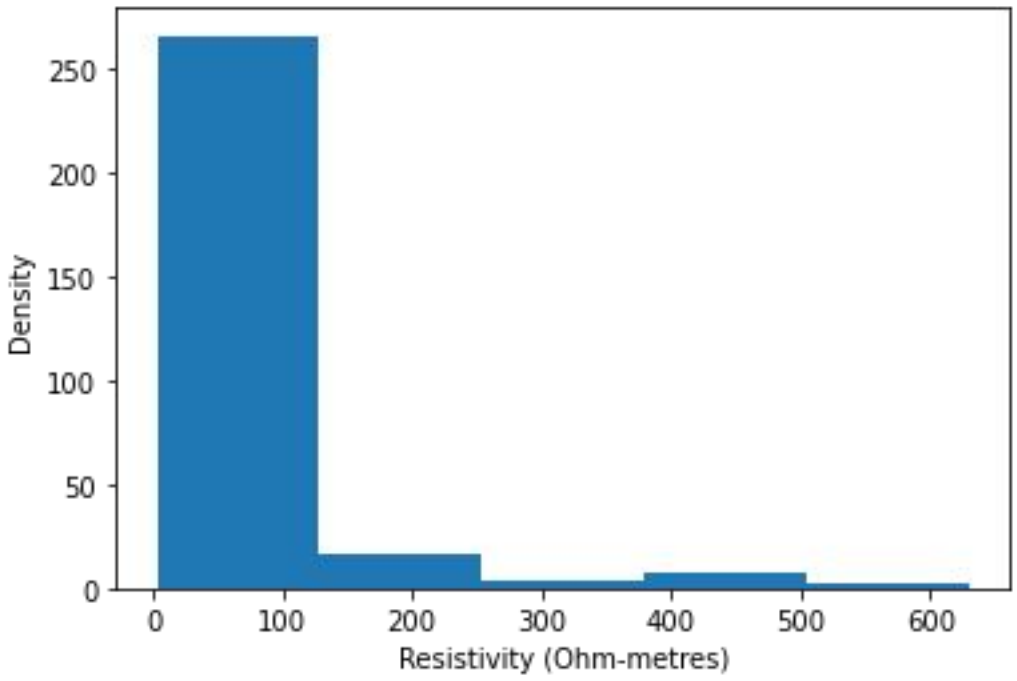


Figure 4.34: Resistivity Histogram

Temperature values were similarly plotted, and Figure 4.35 shows a right skewed graph. More than 60 percent of our temperature data points were between approximately 90 °C to 290 °C.

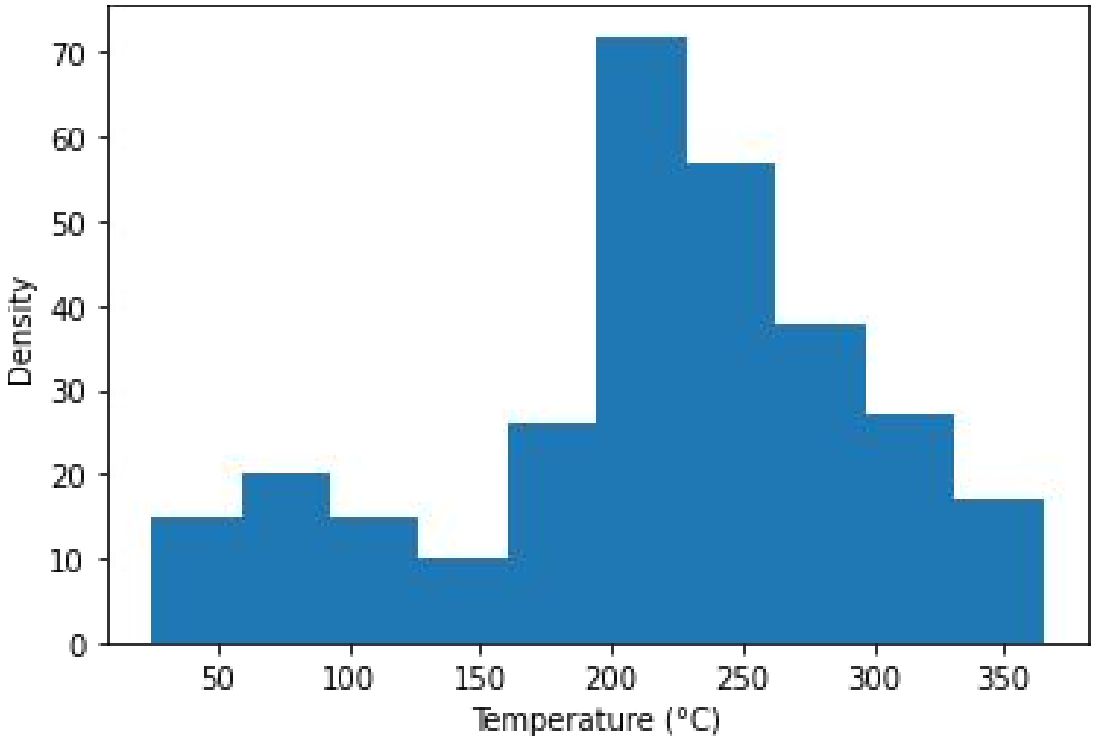


Figure 4.35: Temperature Histogram

As from the scatter plot in Figure 4.36, a complex non-linear relationship between the temperature and resistivity features is observed. The non-linear relationship in the data can be modelled by the powerful nonlinear regression through model training and testing.

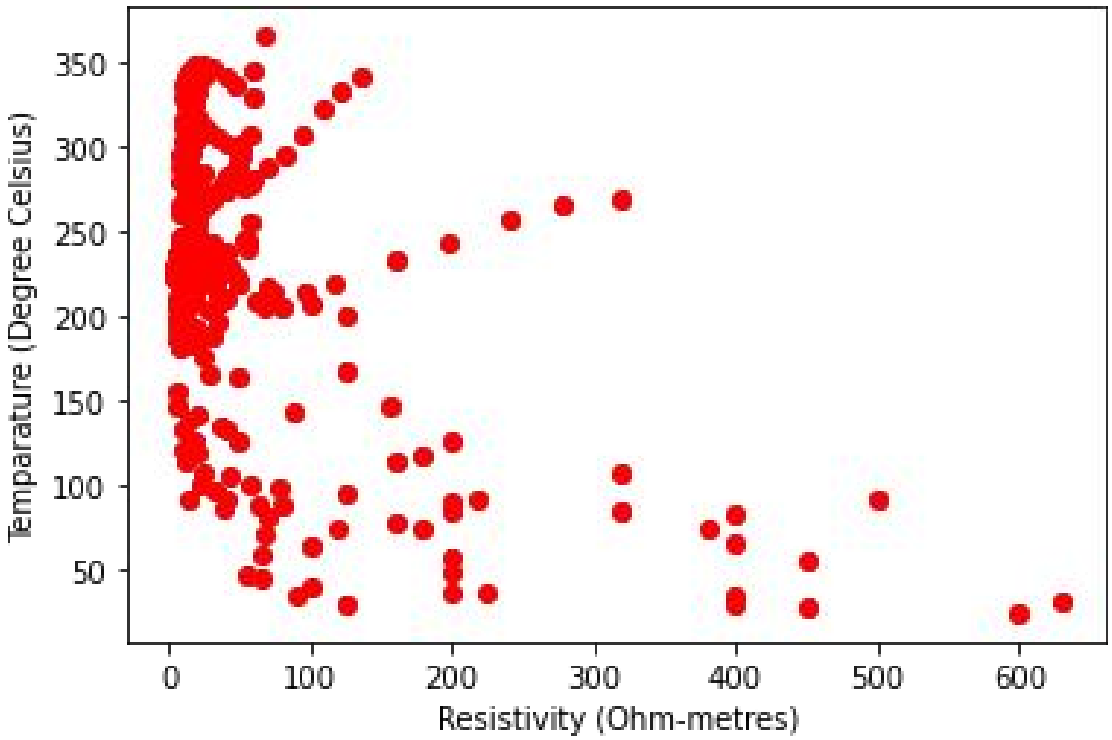


Figure 4.36: Temperature Vs Resistivity Scatter Plot

4.3.2 Nonlinear Regression Model Training and Testing

The model was trained and tested using the polynomial regression, Ensemble regression, SVR and Decision Tree Regression (DTR) and the results and the graphs are illustrated in the subsequent sections.

4.3.2.1 Polynomial Regression

Four polynomial regression models of orders 2, 3, 4 and 5 were developed to predict temperature

from resistivity as shown in the appendix 4a. The Mean Absolute Error (MAE) and Coefficient of Correlation (R^2) for the polynomial models are shown in Table 4.3.

Table 4.3: MAE and R^2 for Polynomial Models

Algorithm	2-Degree	3-Degree	4-Degree	5-Degree
MAE	50.6	51.18	50.47	49.77
R^2	0.5241	0.518	0.521	0.5338

From the table, R^2 an important measure of the strength of the relationship between the independent variables and the dependent variables ranged from the minimum of 0.518 to a maximum of 0.534. This means that between 51.8% to 53.4% of the variation between subsurface temperatures is accounted for by the surface resistivity. This is a moderate relationship between temperature and resistivity. MAE was observed to range from 49.77 to 51.18 indicating that between the predicted temperature value and the actual temperature value, the absolute value ranged from 49.77 to 51.18.

The estimated values that match the relevant field data should follow the line of Y = X if the data well fits the model. Therefore, the polynomial regression scatter plot for predicted temperature vs actual temperature should show a linear positive correlation. This was not the case as indicated in Figures 4.37, 4.38, 4.39 and 4.40 as a result of the complexity of the relationship between resistivity and temperature.

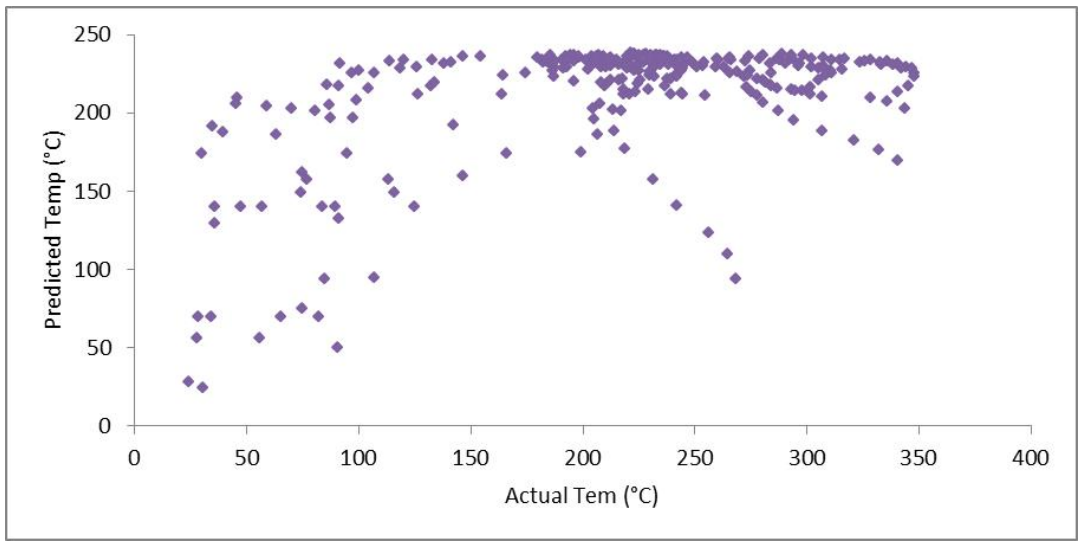


Figure 4.37: Predicted Vs Actual Temperature Degree 2 Polynomial Scatter Plot

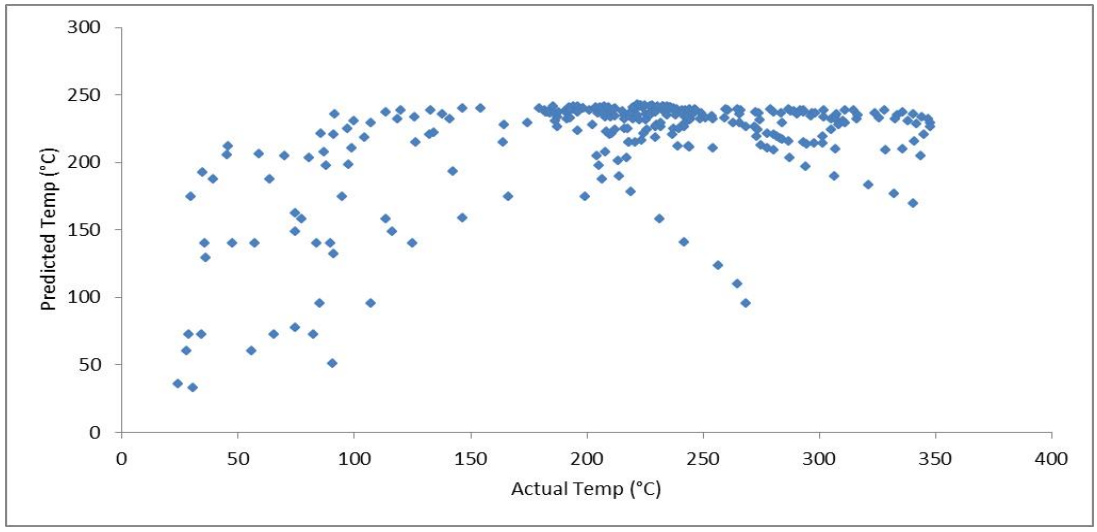


Figure 4.38: Predicted Vs Actual Temperature Degree 3 Polynomial Scatter Plot

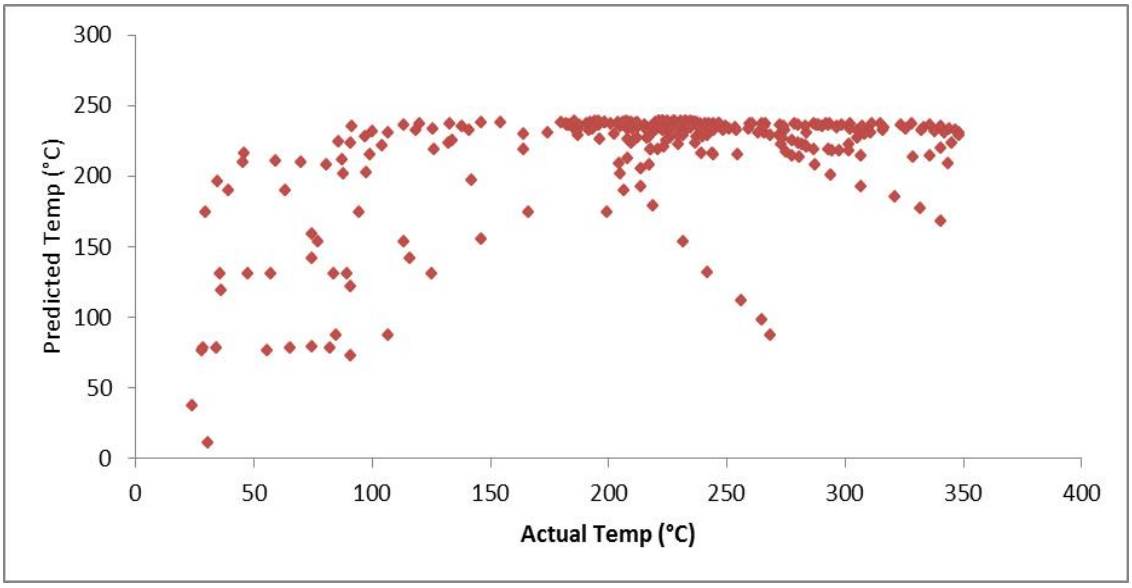


Figure 4.39: Predicted Vs Actual Degree 4 Polynomial Scatter Plot

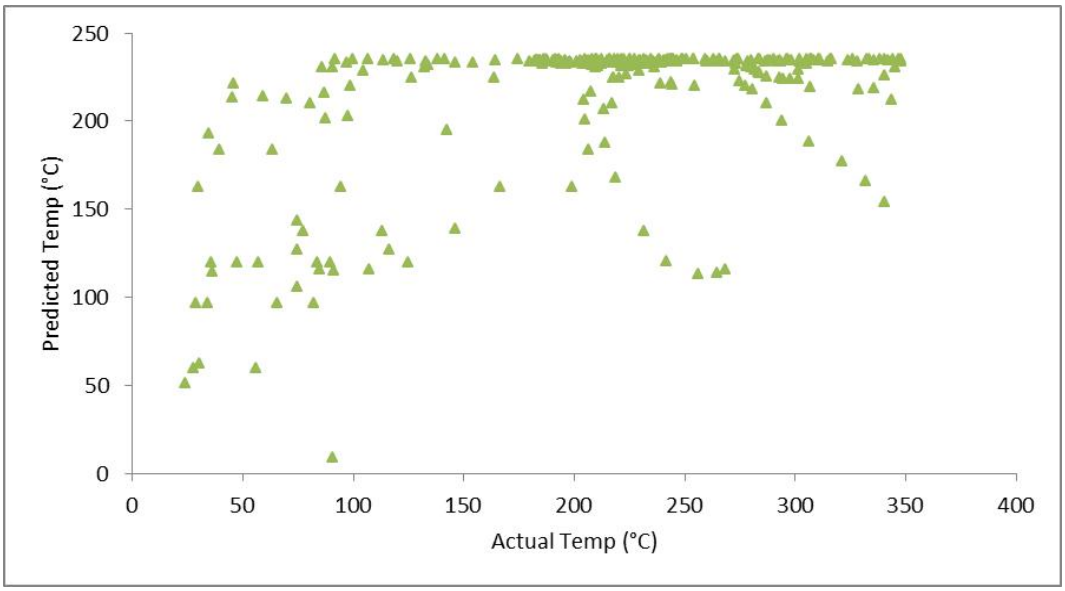


Figure 4.40: Predicted Vs Actual Temperature Degree 5 Polynomial Scatter Plot

The variability in the predicted vs actual temperature values in the polynomial scatter plots indicates that the data features were not learned by the polynomial model. This was

observed in Figure 4.41 and appendices 5 and 6 where resistivity values were observed to decrease with depth up to 1000 metres below the surface, the trend was then reversed where resistivity values started to increase with depth, but temperature continued increasing. At this point, the predicted temperatures started to decrease which was contrary to the measured temperature values. All the four polynomials could not learn from this and therefore the prediction was unreliable.

F

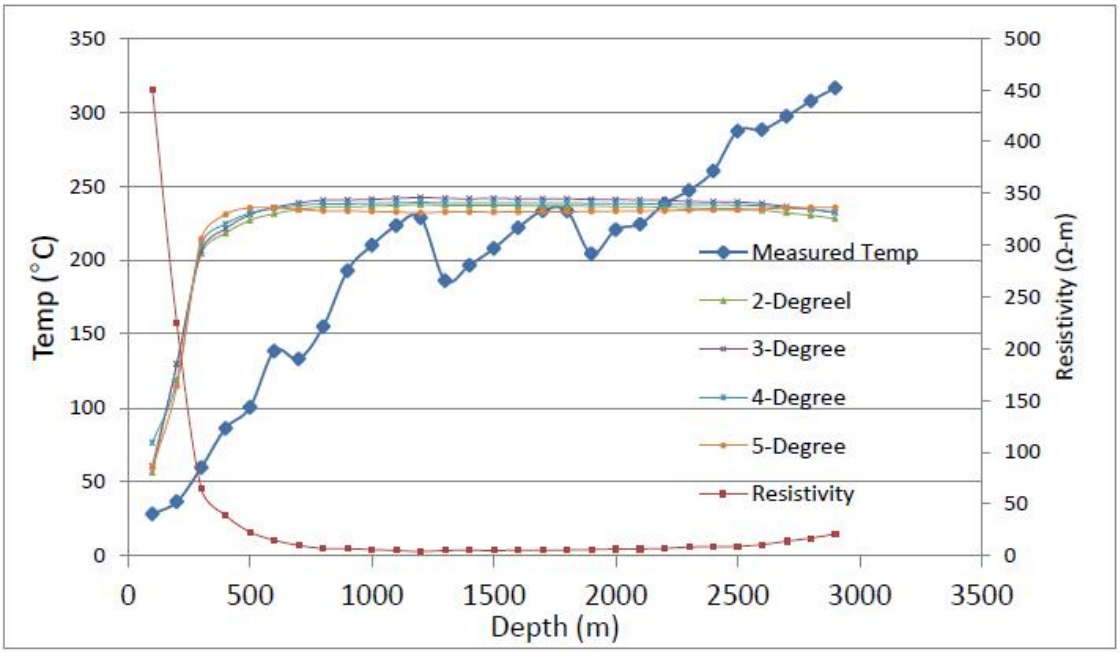


Figure 4.41: OW908 Actual and Polynomial Predicted Parameters

From Figure 4.41, It is observed that resistivity values decrease with depth up to 1000 metres below the surface and then the trend reverses increasing with depth, but temperature continue increasing.

4.3.2.2 Ensemble Models

The goal of ensemble methods is to combine the predictions of several base estimators built with a given learning algorithm in order to improve generalizability or robustness over a single estimator. Adaptive Boosting (AdaBoost) Regression and Random Forest ensemble regression models were considered for the purpose of this research work and the results, and the graphs are illustrated in the subsequent sections.

4.3.2.1.1 Adaptive Boosting (AdaBoost) Regression

For AdaBoost Regressor, the following codes in appendix 4b were used for implementing the algorithm. Mean Absolute Error and coefficient of correlation for the AdaBoost regressor are shown in table 4.4.

Table 4.4: AdaBoost MAE and R^2

Algorithm	AdaBoost
MAE	46.99
R^2	0.66

From table 4.4, R^2 was 0.66 implying that 66% of the variation between subsurface temperatures is accounted for by the surface resistivity which was an improvement as compared to polynomial models whose R^2 value ranged from 0.518 to a maximum of 0.534 an indication that AdaBoot model performed better than polynomial algorithms. Mean Absolute Error was 46.99 indicating that the absolute value difference between the predicted temperature value and the actual temperature is 46.99 which was also a better performance than polynomial models. The relationship between the predicted values and the actual temperature values of AdaBoost model is shown in scatter plot of Figure 4.42. From Figure 4.42, there is a better correlation between the predicted and the actual temperatures especially at temperatures above 300 ^oC as compared to lower temperatures where plots are nonlinear.

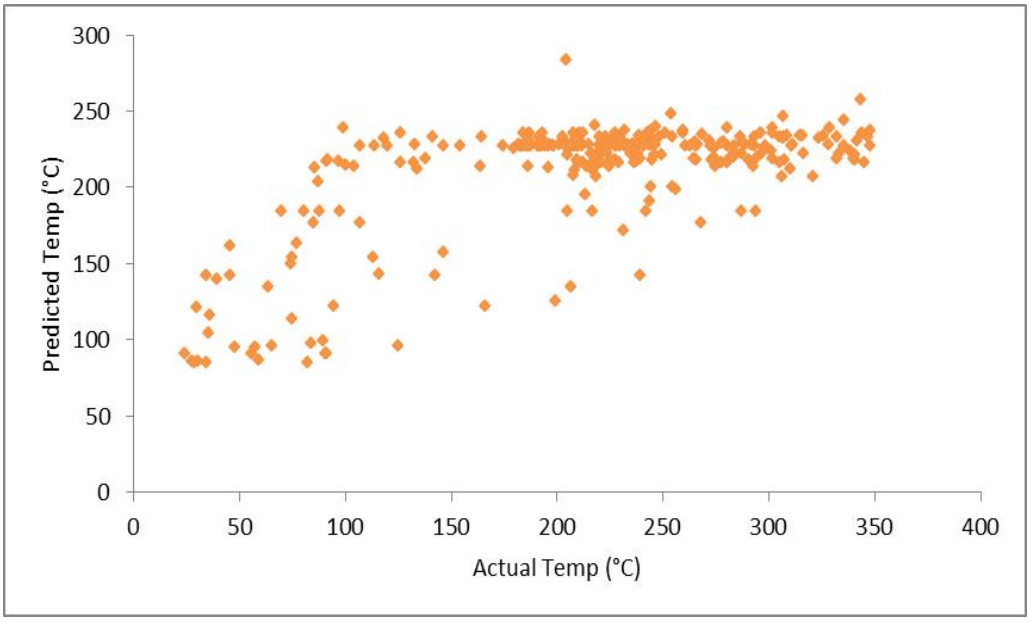


Figure 4.42: Predicted Vs Actual AdaBoost Scatter Plot

The variability in the predicted vs actual temperature values in the scatter plot indicates that the AdaBoost model was able to learn more features from the data as compared to polynomial regression where resistivity values were observed to decrease with depth of 1000 metres below the surface, the trend was reversed where resistivity values started to increase with depth, the AdaBoost predicted temperature showed a slight increase in temperature as opposed to temperature decrease experienced in polynomials, however the temperature increase did not match with the actual temperature increase since they were of lower values shown in Figure 4.43 and appendices 5 and 6.

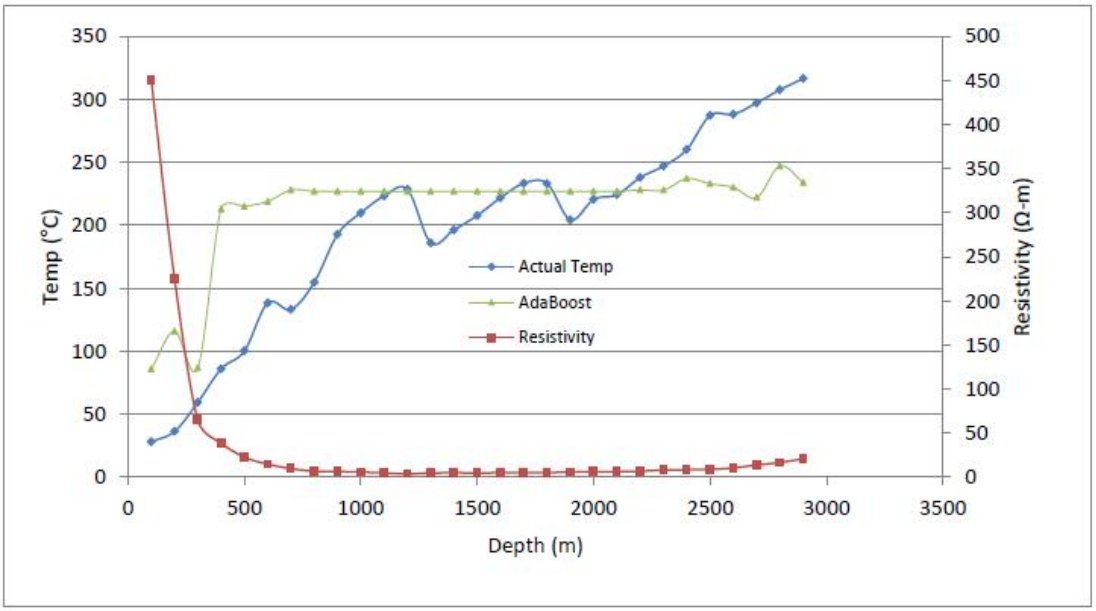


Figure 4.43: OW908 Actual and AdaBoost Predicted Parameters

4.3.2.1.2 Random Forest Regression

For Random Forest Regressor, the following codes in appendix 4c were used for implementing the algorithm. Mean Absolute Error and coefficient of correlation for the RFR model are shown in table 4.5.

Table 4.5: RFR MAE and R2

Algorithm	Random Forest
MAE	36.17
R^2	0.758

From the table, R^2 was 0.758 implying that 75.8% of the variation between subsurface temperatures is accounted for by the surface resistivity. This shows that RFR model performed better that AdaBoost and polynomial algorithms. Mean Absolute Error was

36.17 indicating that the absolute value of the difference between the predicted temperature value and the actual temperature is 36.17.

Figure 4.44 shows the relationship between the predicted values and the actual temperature values of RFR model. From the scatter plot, there is a better correlation between the predicted and the actual temperatures especially at temperatures above 200 ^{o}C as compared to lower temperatures where there are more scattered instead of being linear.

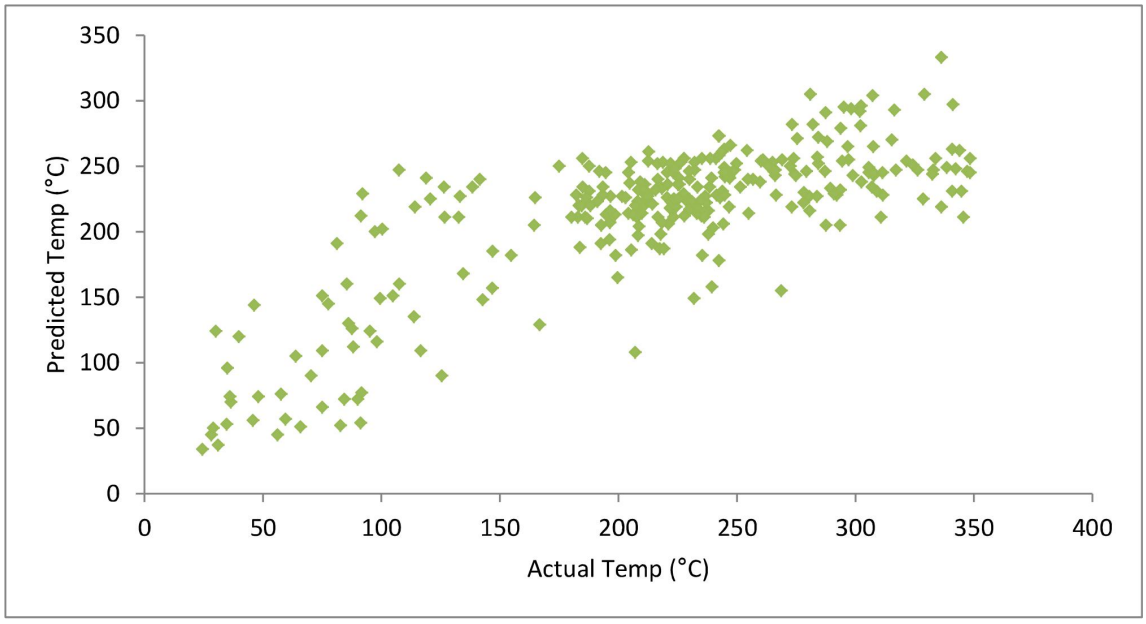


Figure 4.44: Predicted Vs Actual Temperature RFR Scatter Plot

The variability in the predicted vs actual temperature values in the scatter plot indicates that the RFR model was able to learn more features from the data where resistivity values were observed to decrease with depth of 1000 metres below the surface, the trend was reversed where resistivity values started to increase with depth, the RFR was able to cope up with the increase in temperature as opposed to polynomial and AdaBoost regression models as shown in Figure 4.45 and appendices 5 and 6.

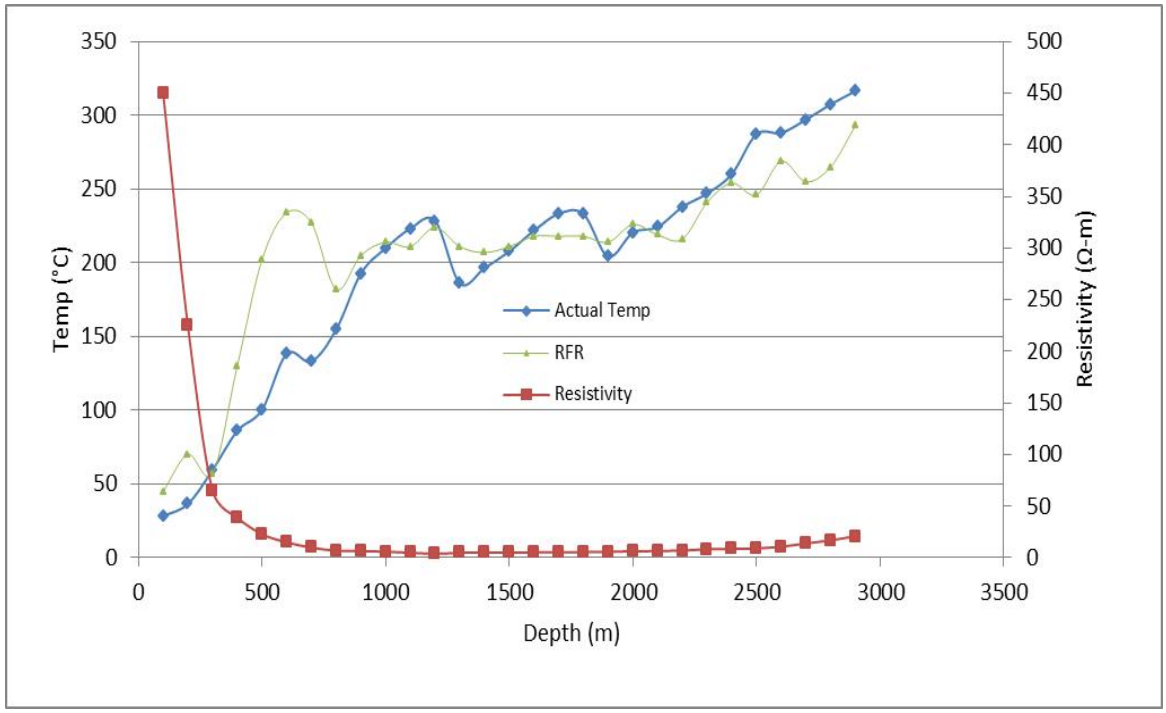


Figure 4.45: OW908 Actual and RFR Predicted Parameters

4.3.2.2 Support Vector Regression (SVR)

For SVR, the following codes in appendix 4d was used for implementing the algorithm. Mean Absolute Error and coefficient of correlation for the SVR model are shown in table 4.6.

Table 4.6: SVR MAE and R^2

Algorithm	SVR
MAE	34.40
R^2	0.789

From the table, R^2 was 0.789 implying that 78.9% of the variation between subsurface temperatures is accounted for by the surface resistivity which was an improvement as compared to polynomial and Ensemble models. SVR model also recorded Mean Absolute Error was 34.4 indicating that the absolute value of the difference between the

predicted temperature value and the actual temperature is 34.4 which was also a better performance compared to polynomial and ensemble models. The relationship between the predicted values and the actual temperature values of SVR model is shown in scatter plot of Figure 4.46. From the Figure, there is a better correlation between the predicted and the actual temperatures.

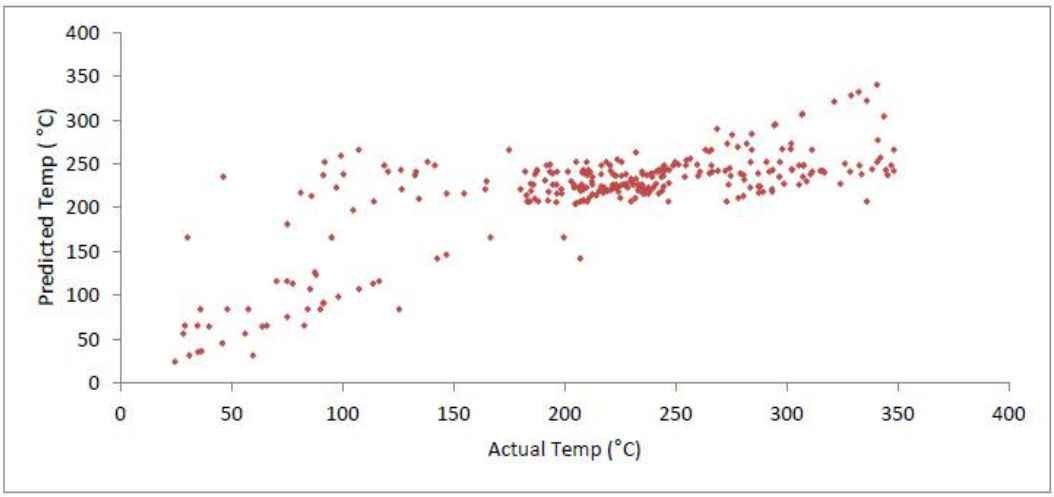


Figure 4.46: Predicted Vs Actual Temperature SVR Scatter Plot

The variability in the predicted vs actual temperature values in the scatter plot indicates that the SVR model was able to learn more features from the data where resistivity values were observed to decrease with depth of 1000 metres below the surface, the trend was reversed where resistivity values started to increase with depth, the SVR was able to cope up with the increase in temperature as shown in Figure 4.47 and appendices 5 and 6.

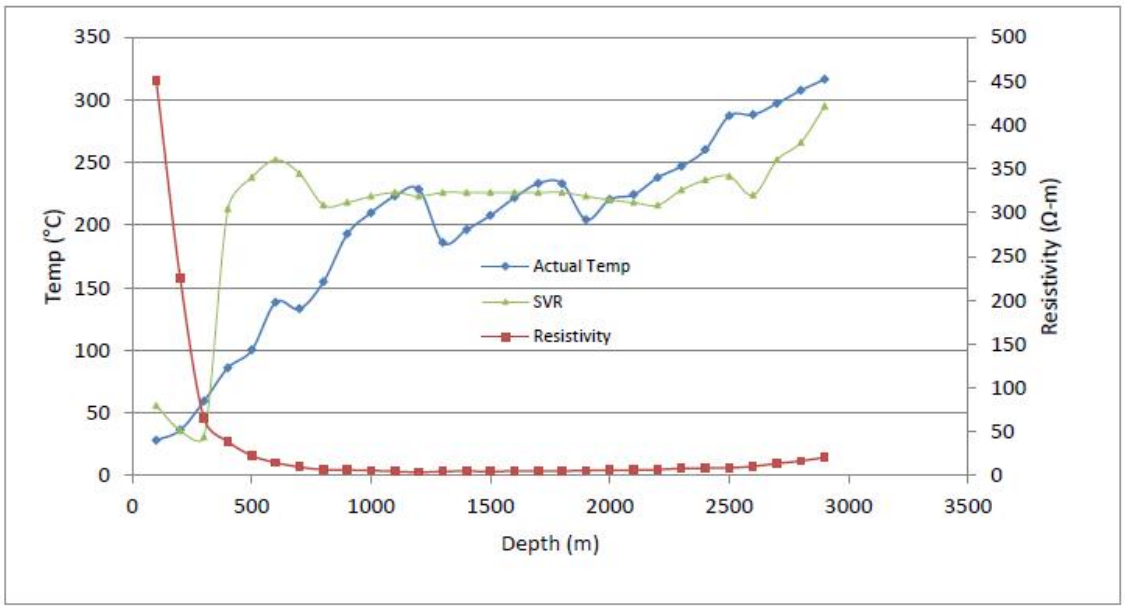


Figure 4.47: OW908 Actual and SVR Predicted Parameters

4.3.2.3 Decision Tree Regression (DTR)

For DTR, the following codes in appendix 4e were used for implementing the algorithm. Mean Absolute Error and coefficient of correlation for the DTR model are shown in table 4.7.

Table 4.7: DTR MAE and R^2

Algorithm	DTR
MAE	29.8
R^2	0.8147

From the table, R^2 was 0.8147 implying that 81.47% coefficient of correlation between the actual and the predicted temperatures. This was the highest value in all the algorithms used in this research. Mean Absolute Error was 29.8 which was also the

lowest value in all the algorithms used in this research. The relationship between the predicted and actual values is depicted in scatter plot of Figure 4.48.

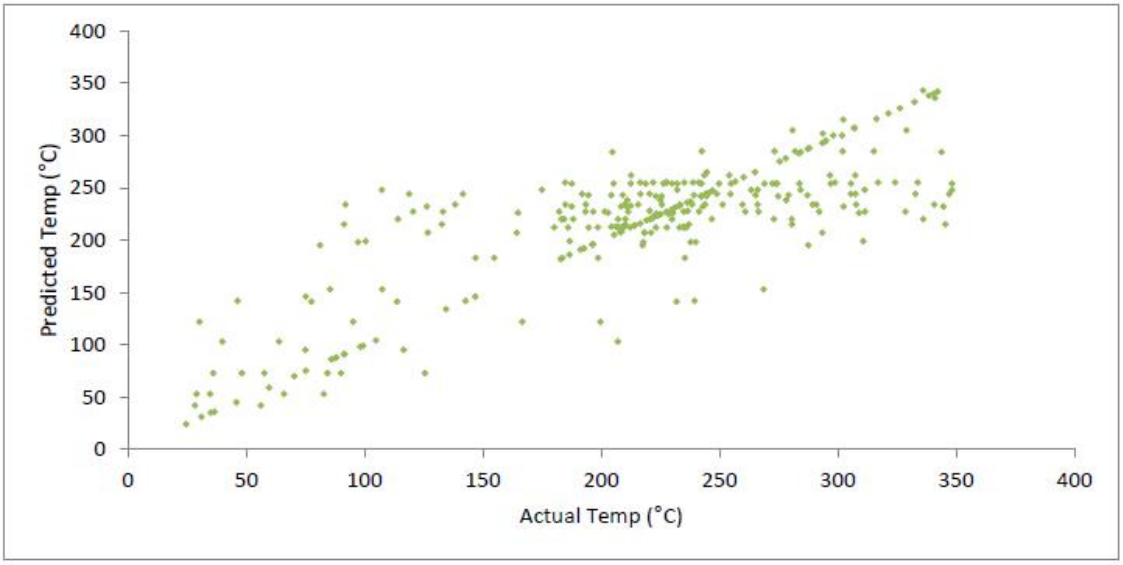


Figure 4.48: OW908 Actual and DTR Predicted Parameters Scatter Plot

As seen from the Figure 4.48, the data points representing the line Y = X was more visible indicating that the predicted data well fits the model. Also, the DTR model was able to learn the trend of decreasing resistivity with depth up to 1000 metres below surface and they started increasing in values after 1000 metres as shown in Figure 4.49 and appendices 5 and 6. From Figure 4.49, the predicted temperature graph matched well with the actual temperature graph. Training the model using the DTR algorithm approach provided superior outputs to other algorithms. This is based on the suggested statistical performance indexes.

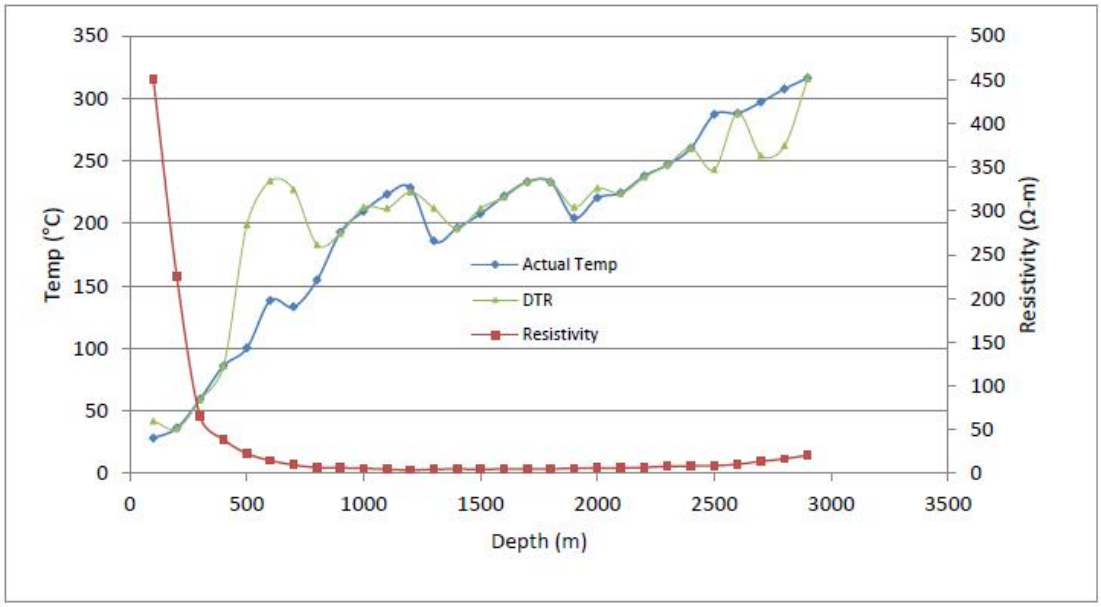


Figure 4.49: OW908 Actual and DTR Predicted Parameters

4.3.3 Testing Best Model with New Data

DTR algorithm having produced the best results as compared to the rest in the study was tested with new data and the R-squared was 0.835 with MAE at 21.7 as shown in Table 4.8.

Table 4.8: DTR New Data MAE and R^2

Algorithm	DTR Test Data
MAE	21.7
R^2	0.835

The Predictive performance of the new data is visualized in Figures 4.50 and 4.51.

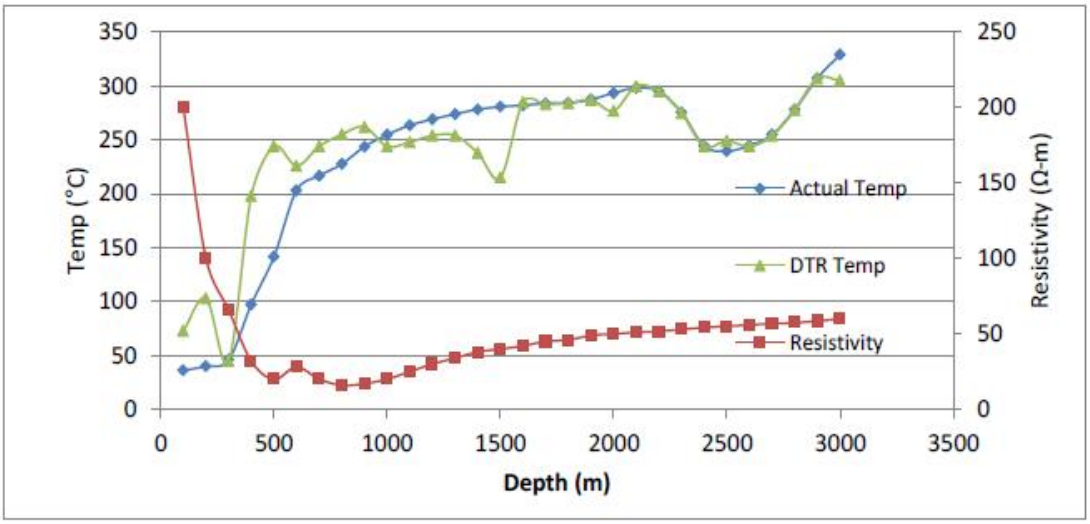


Figure 4.50: The Predictive Performance of the New Data, OW915

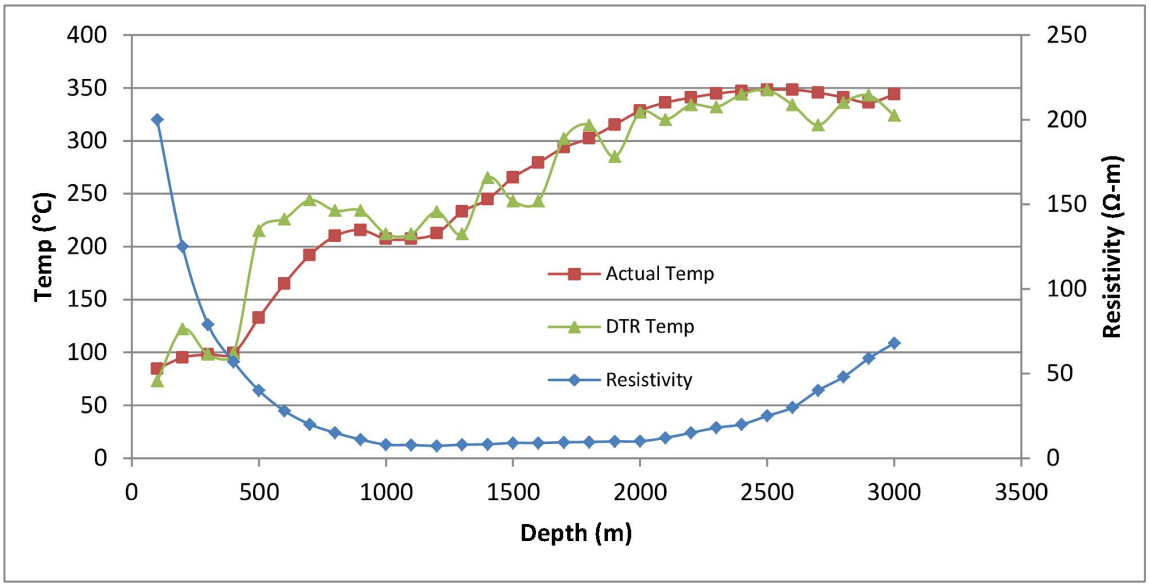


Figure 4.51: The Predictive Performance of the New Data, OW906A

DTR was then used to predict Temperature at a depth of 2000 m b s l that was beyond depth of the wells and the results are as shown in Figure 4.52

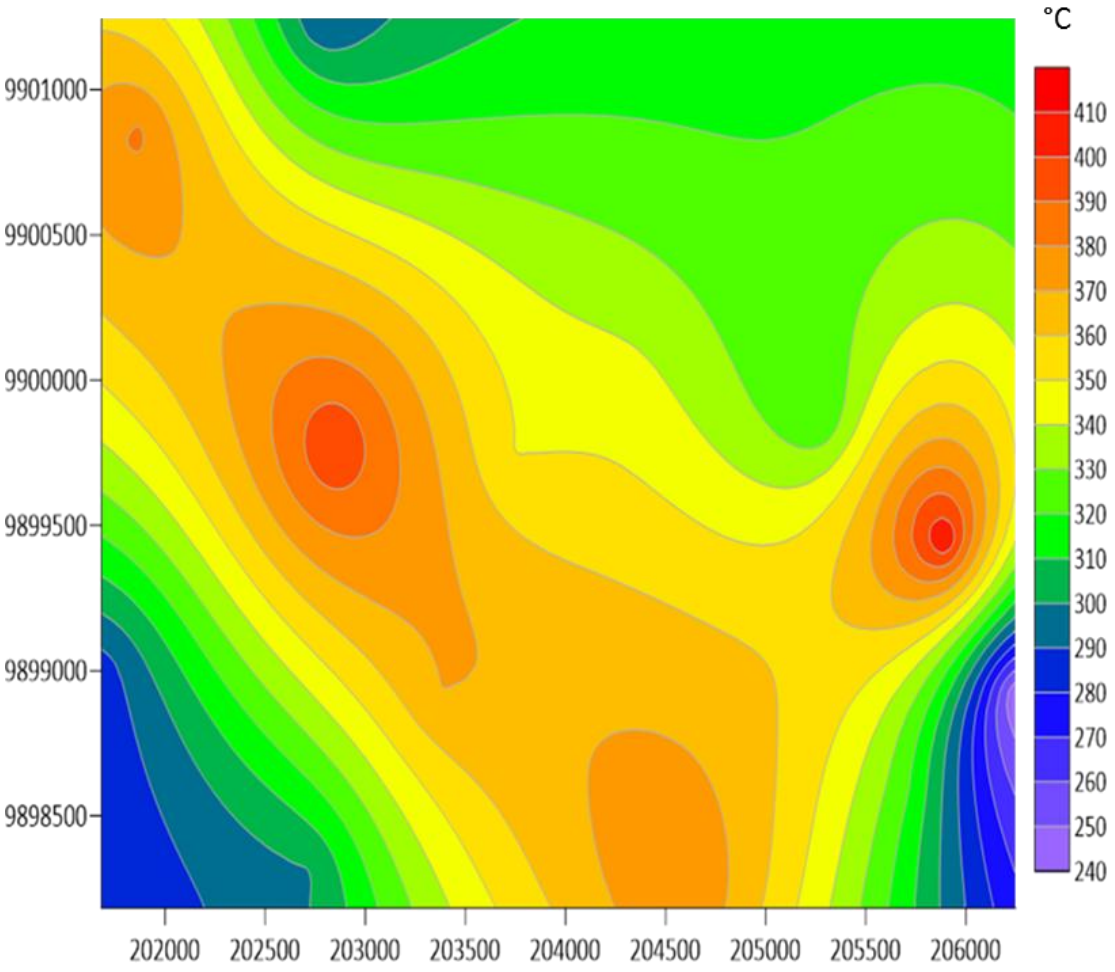


Figure 4.52: DTR Temperature Prediction at 2000 m b s l

From Figure 4.52, there is a high temperature anomaly aligning in the NW-SE direction indicating a possible deep NW-SE magmatic intrusion. This intrusion separates into two dykes that move further nearer to the surface that resulted in high temperature zones as shown in Figures 4.4 4.5, 4.6 and 4.7.

The DTR Algorithm could be implemented in high temperature hydrothermal fields with three main resistivity zones with the first zone being characterized by a narrow layer of higher resistivity near the surface that represents unaltered region followed by the second broader layer of high conductivity that is interpreted to be due to high conductive hydrothermally altered mineralogy such as zeolites and the third zone being a relatively higher-resistivity zone where resistivity is controlled by the formation of high temperature mineralogy at depth such as epidote.

4.4 Image Stacking

The Program Code in appendix 4f was used for stacking resistivity and Temperature images at 500 m a s l and 200 m b s l depths. Figures 4.53 and 4.54 were used as input images for resistivity at 500 m a s l and 200 m b s l depths respectively while Figures 4.55 and 4.56 were used as input images for temperature at 500 m a s l and 200 m b s l depths respectively.

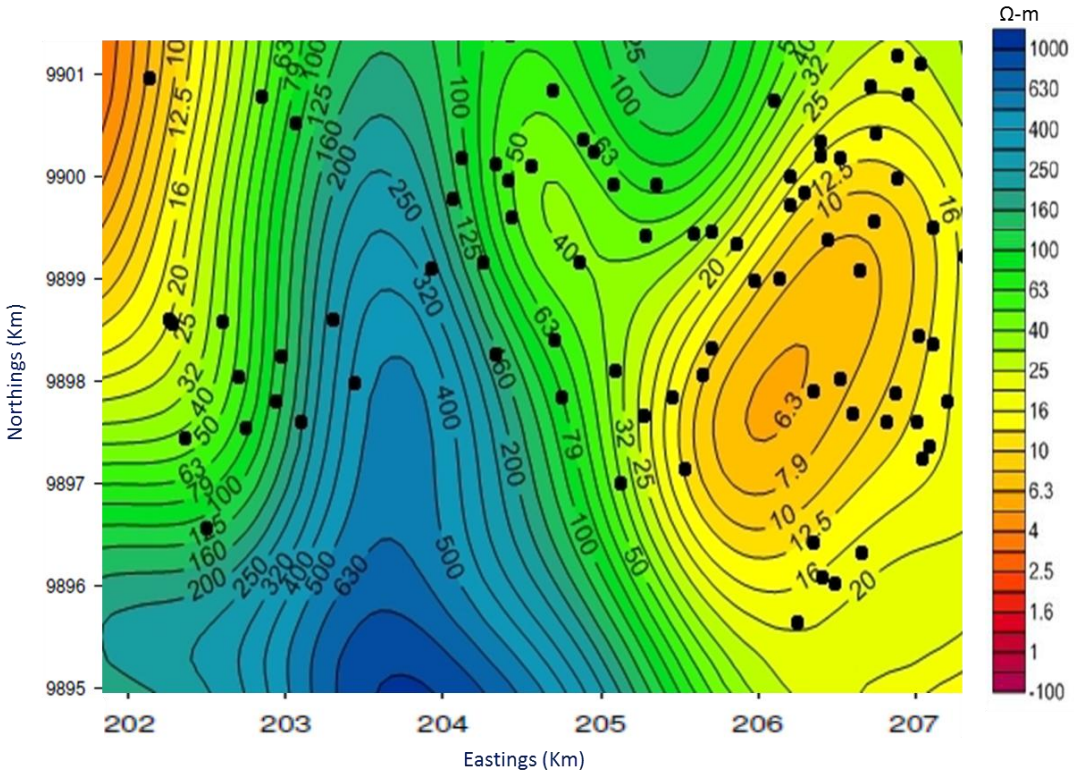


Figure 4.53: Input Images For Resistivity at 500 m a s l

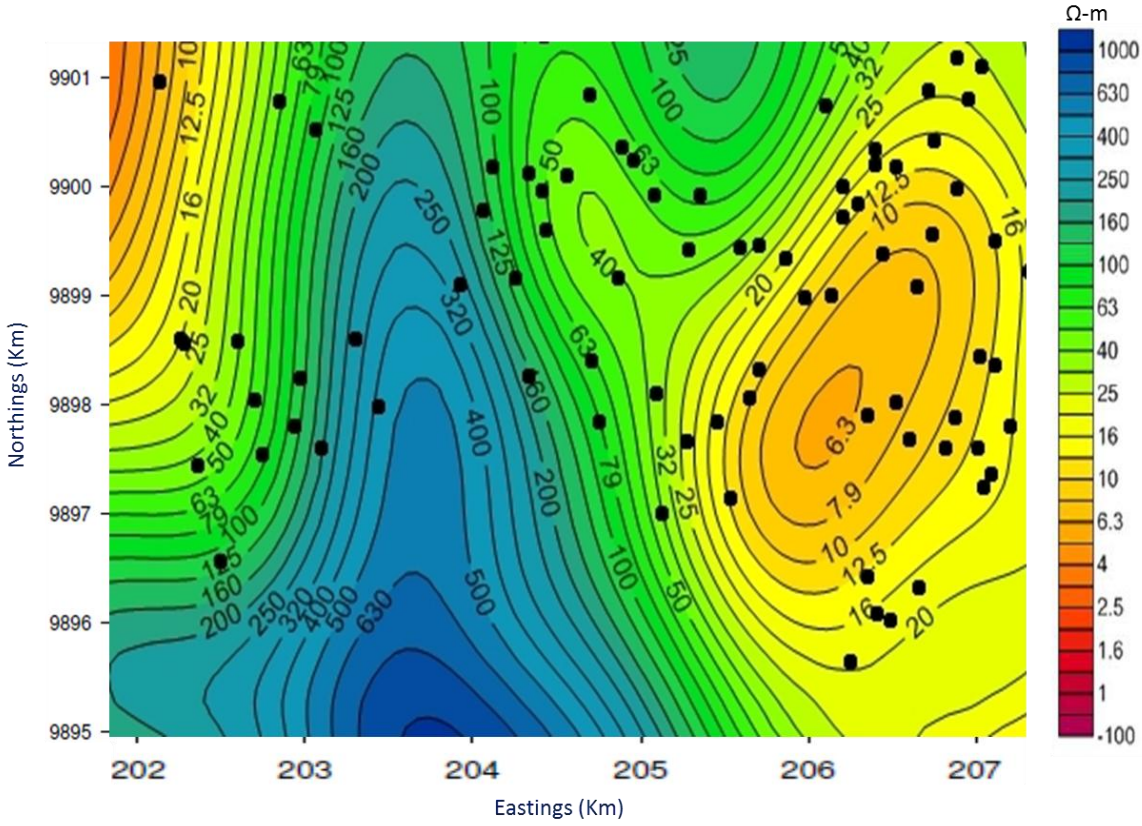


Figure 4.54: Input Image for Resistivity at 200 m b s l

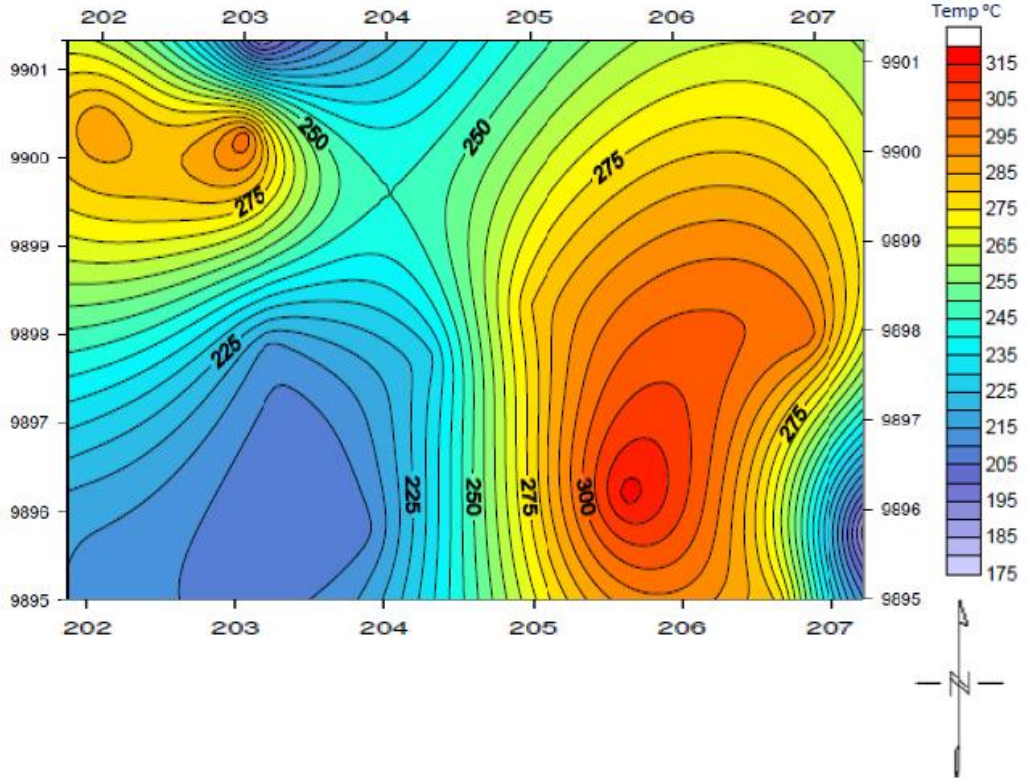


Figure 4.55: Input Image For Temperature at 500 m a s l

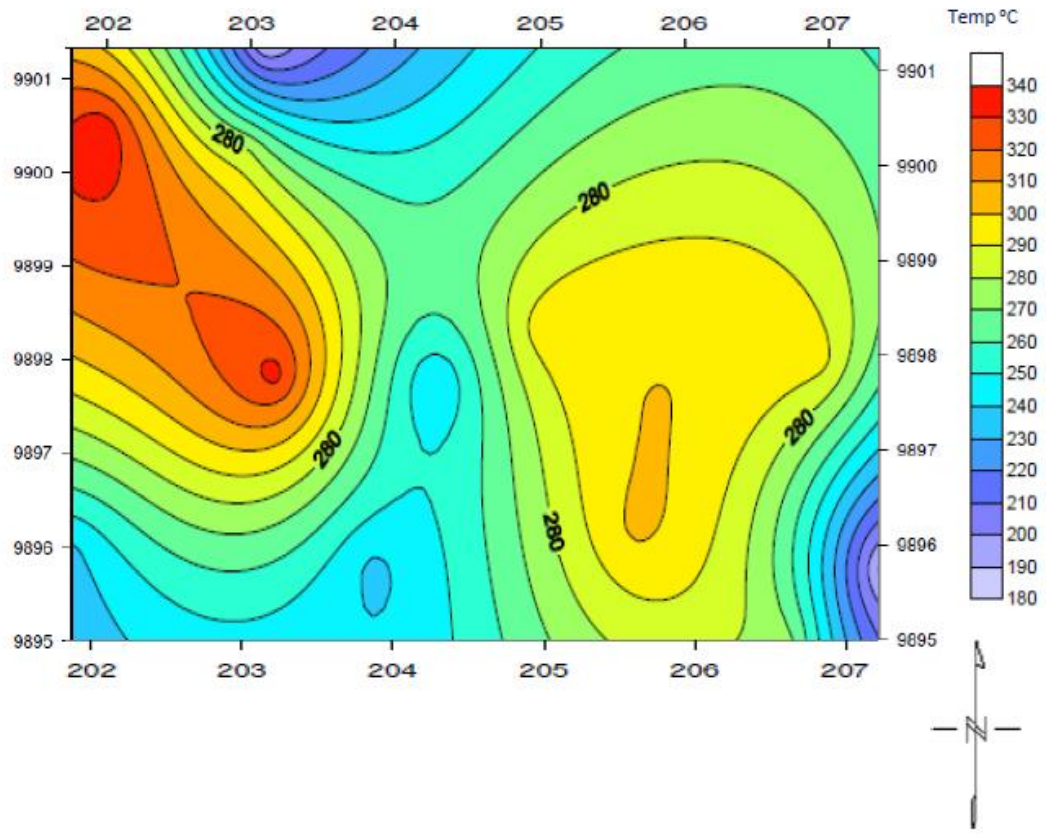


Figure 4.56: Input Image for Temperature at 200 m b s l

The output images indicate areas with low resistivity coincide with areas with high temperatures. The image stacking of temperature and resistivity narrowed down Olkaria domed geothermal field into four main regions of concern.

Regions A and B

The regions are labelled "A" and "B" as shown in Figure 4.57 for the depth of 500 m a s l and Figure 4.58 that show merged images at at 200 m.b.s.l. These are areas of high temperature and high conductivity structures within the field. Resistivity is less than 25 Ωm while Temperature is above 275 ^{o}C. This may be an indication that permeability and porosity are highest in these areas. More wells drilled in these regions may lead to increased production of steam.

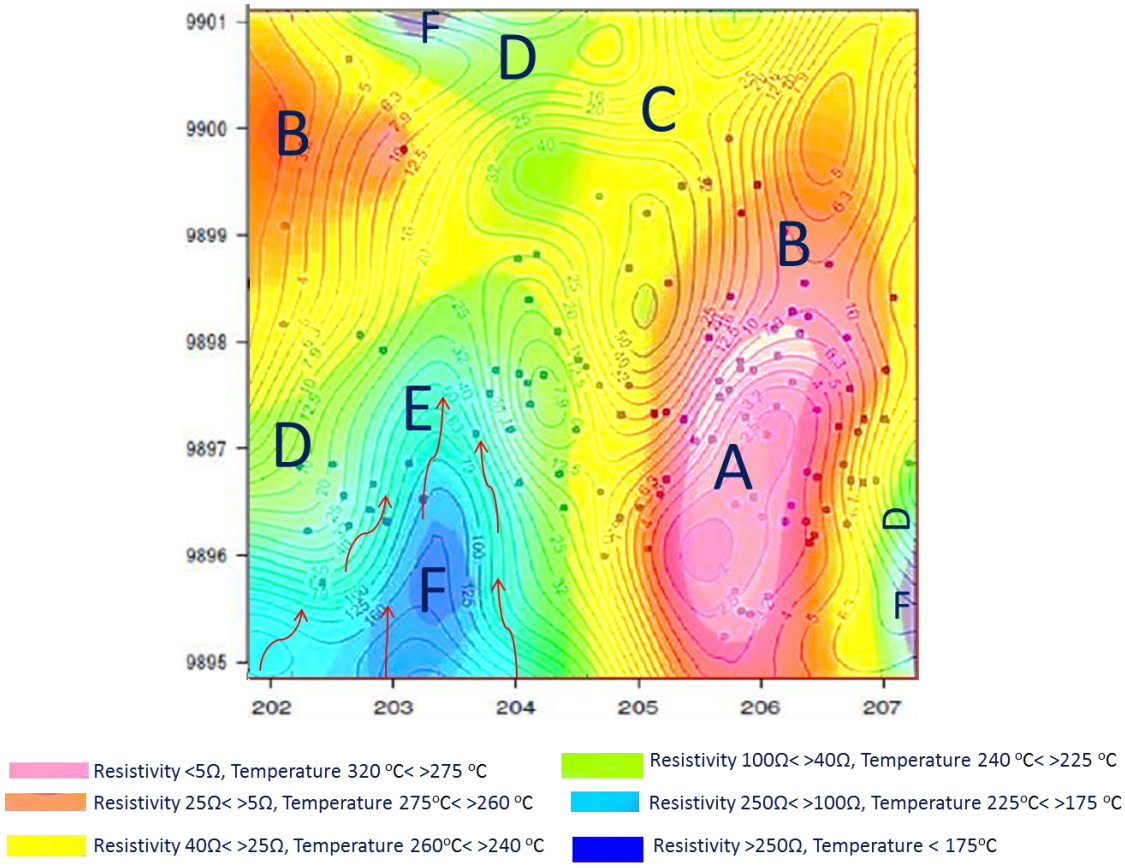


Figure 4.57: Resistivity and Temperature Image at 500 m a s l

Regions C and D

These are regions with lower conductivity and temperatures as compared to A and B with Resistivity ranging between between 25Ωm and 100Ωm while Temperatures range between 225 °C and 260 °C. This may suggest low permeability and therefore wells drilled in these regions may produce less steam.

Regions E and F

These are regions with lowest conductivity and temperatures with resistivity above 100 Ωm while temperature below 225 °C. These may be recharge zones where cold fluid is entering the reservoir. Well drilled in these regions may lead to unproductive wells. In cases of decrease in microgravity, reinjection wells may be drilled in these regions in

order to balance mass withdrawal. This may increase chances for proper economical exploitation of the reservoir.

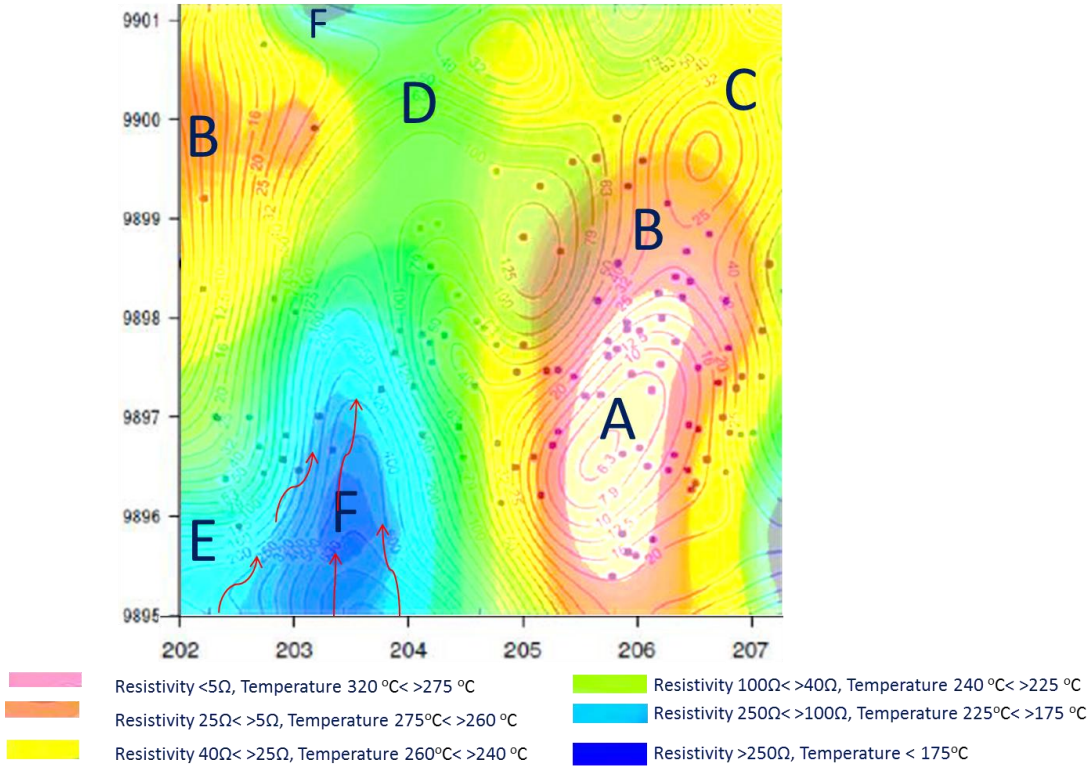


Figure 4.58: Resistivity and Temperature Image at 200 m b s l

4.4.1 Model Validation

The stacked image model was compared with the gravity model carried out within the same field. The gravity study established two main areas of high gravity anomaly which was interpreted as denser bodies as a result of magmatic intrusion from the subsurface. The first denser region was situated in the northwest side of the field. The second denser area was in the SouthEast region of the field as shown in Figure 4.59. These two denser bodies correlated with the regions labeled as ‘A’ and ‘B’ in Figures 4.57 and 4.58 that were categorized as high temperature and high conductivity areas.

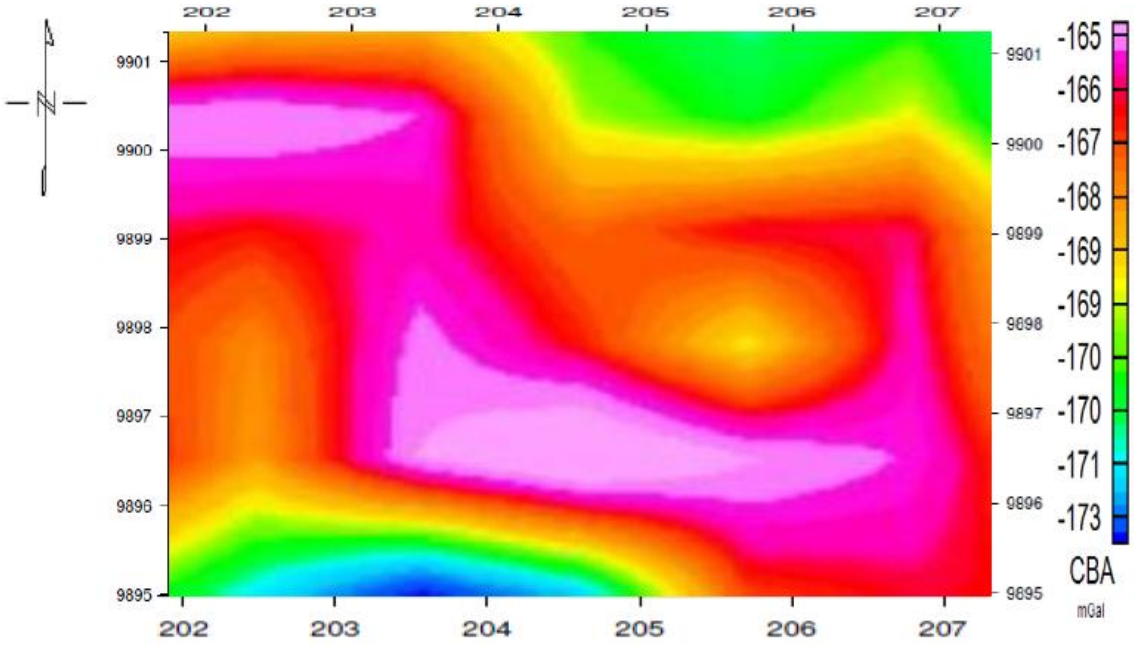


Figure 4.59: Gravity Contour Map of Olkaria Domes

High temperature and high conductivity may therefore be associated with the intrusions of denser magma chamber that had moved closer to the surface creating density contrast with the surrounding rocks. The magmatic intrusion aligns in the NW-SE direction that separates into two dykes as it approaches near surface. This validates the DTR temperature prediction obtained in Figure 4.52 where the high temperature anomaly aligned in the NW-SE direction and suggested a possible deep NW-SE magmatic intrusion.

Figure 4.59 confirmed that to achieve maximum productivity from wells and save well drilling costs, drilling may be concentrated in the areas marked ‘A’ and ‘B’. These are areas with denser rock structures as compared to the surroundings, a phenomenon that was associated with denser magmatic intrusions. The two intrusions resulted in raise in temperature, an important parameter for a good geothermal reservoir. Also, these two areas have the lowest resistivity indicating that permeability and porosity may be highest

in these areas. High permeability and porosity are important in geothermal reservoir as they allow ease of fluid flow into the well for steam extraction at the well separator.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.2 Conclusions

With high temperatures exceeding 315^{o}C being encountered in Olkaria Domes geothermal field, it can be categorized as a high temperature field. From the temperature recovery profiles, wells in the SouthEast exhibited convection heat transfer as the dominant mode of heat transfer while the wells with conductive heat transfer are located in the northwest of the Olkaria Domes geothermal field. This indicates that the SouthEast region is more permeable since convection heat transfer happens only through the fluids which can only take place in permeable and porous rocks as compared to the northwest region.

Ilites and Quartz minerals were encountered at shallower depths of in the eastern side as opposed to the western side. On the Eastern region Quartz was encountered at shallower depths of about 1600 m a s l compared to the western side where the first appearance of Quartz was as low as 1300 m a s l. Illites appeared on the SouthEastern side at the depth of 1750 m a s l while the deepest was in the northwest side at 1150 m a s l. Therefore SW/NE oriented fault could be as a result of more upthrow to the east as compared to the west leading to a normal fault. This indicates that there could have been a differential upward movement of magma chambers with the SouthEastern side moving more closer to the surface moving these rocks nearer to the surface.

From the resistivity contours at selected depths, three main resistivity regions were deduced; one low resistivity to the Northwest, the second low resistivity was observed to the Eastern side of the field. The two low resistive regions are separated by a NE-SW trending high resistive region. Resistivity decreases with depth up to a depth of 500 m a s l then it increases with depth. From the cross sections, the results reveal three main resistivity zones. The first zone was characterized by a narrow layer of higher resistivity

near the surface and was depicted to represent unaltered region. The second was a broader layer of high conductivity that was interpreted to be due to high conductive hydrothermally altered mineralogy such as zeolites. The third was a relatively higher-resistivity zone where resistivity is believed to be controlled by the formation of high temperature mineralogy at depth such as epidote.

For Decision Tree Regression (DTR), R^2 was 0.8147 implying that 81.47% coefficient of correlation between the actual and the predicted temperatures. This was the highest value in all the algorithms used in this research. Mean Absolute Error was 29.8 which was also the lowest value in all the algorithms used in this research.

Results from multiple image stacking technique by use of Pivotal Focus Algorithms in mapping geothermal reservoir indicate that uncertainty in well drilling can be minimized if results from two or more geophysical methods are merged.

5.3 Recommendations

This research only used data from one of the seven segments of the greater Olkaria geothermal field and from the available data, the set of resistivity input attributes employed in this study might not entirely capture all the influences of resistivity on temperature for the entire Olkaria field. Therefore, future studies should include data from other producing fields to increase the number of samples. The study can also be extended to other geothermal fields in the world for comparison purposes.

Other predictive models can also be applied and the results compared with this research.

REFERENCES

Abrahart, R.J., See L.M. and Solomatine, D. (2008). *Practical Hydroinformatics: Computational Intelligence and Technological Developments in Water Applications:* Springer, (1st Ed.) Water Science and Technology Library

Adebayo, A.R., Babalola, L., Hussaini, S.R., Alqubalee, A. and Babu, S.R (2019). Insight into the Pore Characteristics of a Saudi Arabian Tight Gas Sand Reservoir. *Energies*. 2019; 12(22):4302.

Ahmed, Z., Mohamed, K., Zeeshan, S. and Dong, X. (2020). Artificial Intelligence with Multi-Functional Machine Learning Platform Development for Better Healthcare and Precision Medicine. *The Journal of Biological Databases and Curatio,*. 2020

Alhakeem, A.A. (2018). *3D Seismic Attribute Analysis And Machine Learning For Reservoir Characterization in Taranaki Basin, New Zealand*. PhD Dissertation 2659, Missouri University of Science and Technology.

Alwee, R., Shamsuddin, S.M. and Sallehuddin, R. (2013). Hybrid Support Vector Regression and Autoregressive Integrated Moving Average Models Improved By Particle Swarm Optimization for Property Crime Rates Forecasting With Economic Indicators. *Scientific World Journa*l, 2013

Antonopoulosa, I., Robua,V., Courauda, B., Kirlic, D., Norbua, S., Kiprakisc, A., Flynna, D., Elizondo-Gonzalezd, S. and Wattam, S. (2020). Artificial intelligence and Machine Learning approaches to energy demand-side response: A systematic review. *Renewable and Sustainable Energy Reviews,* 130 (2020)

Aragón, A.A., Izquierdo-Montalvo, G., Aragón-Gaspar, D.O and Barreto-Rivera, D.N. (2019) Stages of Integrated Geothermal Project. *Renewable Geothermal Energy Explorations* (2019)

Arthur, M.M. (2018). *A Case Study of Three Dimensional Inversions of Electrical Resistivity to Image Geothermal System in Korosi Geothermal Prospect-Kenya*. MSc. Thesis, Dedan Kimathi University of Technology, Kenya

Barnadas, M.V. (2016). *Machine Learning Applied To Crime Prediction*. BSc. Dissertation, Universitat Politècnica de Catalunya.

Biru, M.F., (2016). Analysis of Well Testing, Temperature and Pressure in High-Temperature Wells of Aluto-Langano, Ethiopia. *Proceedings of the World Geothermal Congress.* Reykjavik, Iceland

Bravo-Osuna, A.G., Gómez-Treviño, E., Cortés-Arroyo, O.J., Delgadillo-Jauregui, N.F. and Arellano-Castro, R.F. (2021). Reframing the Magnetotelluric Phase Tensor for Monitoring Applications: Improved Accuracy and Precision in Strike Determinations. *Earth, Planets and Space (2021)* 73:34 https://doi.org/10.1186/s40623-021-01354-y

Brodie, G. (2019). Energy Transfer from Electromagnetic Fields to Materials, in K. H. Yeap, K. Hirasawa (eds.), Electromagnetic Fields and Waves, IntechOpen, London. 10.5772/intechopen.83420..

Brownlee, J. (2020). Boosting and AdaBoost for Machine Learning in Machine Learning Algorithms. *https://machinelearningmastery.com/boosting-and-adaboost-for-machine-learning/*

Cai, Y. (2019). *Design of Electrically Small Loop Antennas with Wide-Band Frequency Tuning Capability to Facilitate Mars Exploration*. MSc. Thesis, University of California.

Chaki, S. (2015). *Reservoir Characterization: A Machine Learning Approach*. MSc. Thesis, Indian Institute of Technology, Kharagpur.

Chaki, S., Verma, A.K., Routray, A., Mohanty, W.K. and Jenamani, M. (2014). Well Tops Guided Prediction of Reservoir Properties using Modular Neural Network Concept: A Case Study from Western Onshore, India. *Journal of Petroleum Science and Engineering,* 123, 155-163

Christensen, B.N. (2022). Joint Inversion of Airborne TEM Data and Surface Geoelectrical Data. The Egebjerg Case. *Journal of Applied Geophysics, Volume 196, https://doi.org/10.1016/j.jappgeo.2021.104511.*

Das, D. (2018). *Time Series Analysis & Visualization: Forecasting and Detection of the Abnormal Changes in Data*. MSc. Thesis, North Carolina State University

Dawood, G.E. (2019). *Geo-locating UEs Using Multi-output Decision Tree Regressor*. MSc. Thesis, Florida Institute of Technology, Melbourne.

Deland, S. (2018). Getting Back to the Basics: What is Machine Learning? *https://www.dataversity.net/getting-back-basics-machine-learning*

Elizondo, D.P. (2019). *The Borinquen Geothermal Area, Costa Rica: 1D and 3D Inversion of Resistivity Data, Geological/Geothermal Interpretation.* MSc Thesis University of Iceland

Enyenihi, H. J. (2018). Automated Electrically Operated Dehydration Bed Modelling: Automatic Generation of Expression Approach. *IOSR Journal of Electrical and Electronics Engineering (IOSR-JEEE)* 13.6 (2018): 43-55.

Escobedo, D., Patrier, P., Beaufort, D., Gibert, B., Levy, L., Findling, N. and Mortensen, A. (2021). Contribution of the Paragenetic Sequence of Clay Minerals to Re-Examination of the Alteration Zoning in the Krafla Geothermal System. Minerals 2021, 11, 935. https://doi.org/10.3390/min11090935

Ferroud, A., Rafini, S. and Chesnaux, R. (2019). Using Flow Dimension Sequences To Interpret Non-Uniform Aquifers With Constant-Rate Pumping-Tests: A Review. *Journal of Hydrology X,* 2

Flóvenz, Ó.G., Hersir, G.P., Saemundsson, K., Ármannsson, H., and Fridriksson T.H., (2012) Geothermal energy exploration techniques. In: Syigh, A., (ed.), *Comprehensive renewable energy, 7:* 51-95.

Flynn, K.C. (2019). *Predicting Nutrient Content, Plant Health, and Site Suitability: A Case Study of Eragrostis Tef*. PhD. Thesis, Oklahoma State University.

Franco, A. and Vaccaro, M. (2020). Sustainable Sizing of Geothermal Power Plants: Appropriate Potential Assessment Methods. *Sustainability* 12, 3844

Fulignati, P. (2020). Clay Minerals in Hydrothermal Systems. *Minerals*, 10, 919; doi:10.3390/min10100919

Gao, C., Sun, H., Wang, T., Tang, M., Bohnen, N.I., Müller, M. L. T. M., Herman, T., Giladi, N., Kalinin, A., Spino, C., Dauer, W., Hausdorff, J.M. and Dinov, I.D. (2018). Model-based and Model-free Machine Learning Techniques for Diagnostic Prediction and Classification of Clinical Outcomes in Parkinson's Disease. Sci Rep 8, 7129 (2018). https://doi.org/10.1038/s41598-018-24783-4

Gebru, T.N. (2018). Temperature and Well Test Analyses of Selected Wells of Aluto Langano Geothermal Field, Ethiopia. *Proceedings of the 7th African Rift Geothermal Conference*. Kigali, Rwanda

Gokani, J. (2017). The Evolution of Banking: *AI, MS&E 238 Blog, Leading Trends in Information Technology,* Stanford University.

Gómez-Treviño, E., Muñiz, Y., Cuellar, M. and Calderón-Moctezuma, A. (2018). Invariant TE and TM magnetotelluric impedances: application to the

BC87 dataset. *Earth Planets Space* 70, 133 (2018). https://doi.org/10.1186/s40623-018-0900-y

Gylfi, P.H. and Árnason, K. (2013). Resistivity of Rocks

Habuza, T., Navaz, A.N., Hashim, F., Alnajjar, F., Zaki, N., Serhani, M.A. and Statsenko, Y. (2021). AI applications in robotics, diagnostic image analysis and precision medicine: Current limitations, future trends, guidelines on CAD systems for medicine. *Informatics in Medicine Unlocked, ISSN: 2352-9148, Vol: 24, Page: 10059 doi.org/10.1016/j.imu.2021.100596*

Hadush, W. (2018). Magneto Telluric Methods in Geothermal Exploration at Aluto-Langano, Ethiopia. *Proceedings, 7th African Rift Geothermal Conference Kigali, Rwanda.*

Haldar, S. K. (2020). Sedimentary rocks. In *Introduction to mineralogy and petrology* (pp. 121–212). Elsevier, Amsterdam.

Hersir, G.P. (2018). Resistivity Surveying and Electromagnetic Methods. *Presented at SDG Short Course III on Exploration and Development of Geothermal Resources,*
at Lake Bogoria and Lake Naivasha, Kenya.

Herzog, S., Wörgötter, F. and Parlitz, U. (2018). Data-Driven Modeling and Prediction of Complex Spatio-Temporal Dynamics in Excitable Media. *Front. Appl. Math. Stat.* 4:60.

Hu, H., Dai, L., Sun, W., Zhuang, Y., Liu, K., Yang, L., Pu, C., Hong, M., Wang, M., Hu, Z., Jing,C., Li, C. and Yin, C. (2022). Some Remarks on the Electrical Conductivity of Hydrous Silicate Minerals in the Earth Crust, Upper Mantle and Subduction Zone at High Temperatures and High Pressures. Minerals 2022, 12, 161. https://doi.org/10.3390/min12020161

Hu, X. (2016). *A Coupled Geomechanics and Flow Modeling Study for Multistage Hydraulic Fracturing Of Horizontal Wells in Enhanced Geothermal Systems Applications*. MSc. Thesis, Colorado School of Mines

Jansen, G and Miller, S. A. (2017). On the Role of Thermal Stresses during Hydraulic Stimulation of Geothermal Reservoirs. *Hindawi Geofluids*, 2017

Jolie, E., Scott, S., Faulds, J. and Chambefort, I. (2021). Geological controls on geothermal resources for power generation. Nature Reviews Earth & Environment 2(5):1-16 DOI:10.1038/s43017-021-00154-y

Kahlen, L., Kurdziel, J.M., Day, T. and Schiefer, T. (2019). *The role of geothermal and coal in Kenya's electricity sector and implications for sustainable developmen*t. New climate institute, 2019

Kahuda, D. and Pech, P. (2020). A New Method for the Evaluation of Well Rehabilitation from the Early Portion of a Pumping Test. *Water*, 12, 744

Kandie, R., Mbuthia, P., and Stamic, J. (2016). Use of Leapfrog Geothermal Software in Data Integration and 3D Visualization: A Case Study of Olkaria Domes Geothermal System. *Proceedings of the 6th African Rift Geothermal Conference.* Addis Ababa, Ethiopia.

Kargarpour, A.M., (2017). Investigation of reservoir temperature in a gas reservoir in Middle East: Case Study. *Journal of Petroleum Exploration and Production Technology* 7, (2): 531–541

Karrar, A.M.A. and Sun, J. (2017). Artificial Intelligence: An Overview. *International Journal of Science and Research (IJSR)*, 7(6): 495-501

Kibet, M.K., Sang P.K. and Varet, J. (2019). Rock Types and Alteration Mineralogy Occurring In Olkaria Geothermal Field, Kenya. *IOSR Journal of Applied Geology and Geophysics (IOSR-JAGG)* V7, (1): 47-57

Kothapalli, A.K. (2021). *Short-Term Electricity Consumption Prediction: Elområde 4, Sweden*. MSc. Thesis, Lule a University of Technology Lule a, Sweden

Lagat, J.K., (2004). *Geology, hydrothermal alteration and fluid inclusion studies of the Olkaria Domes geothermal field, Kenya*. MSc. thesis, University of Iceland.

Lee D, Lee J, Ko J, Yoon J, Ryu K, Nam Y. Deep Learning in MR Image Processing. Investig Magn Reson Imaging. 2019 Jun;23(2):81-99. https://doi.org/10.13104/imri.2019.23.2.81

Lee, G., Kim, W., Oh, H., Youn, B.D. and Kim, N.H. (2019). Review of statistical model calibration and validation—from the perspective of uncertainty structures. *Struct Multidisc Optim,* 60: 1619–1644.

Lewis, A. (2014). *A Method Using Drawdown Derivatives to Estimate Aquifer Properties near Active Groundwater Production Well Fields*. MSc. Thesis Colorado State University

Lichoro, C.M. (2010). Joint 1-D Inversion of TEM and MT Data from Olkaria Domes Geothermal Area, Kenya. *Proceedings, 3rd East African Rift Geothermal Conference Djibouti*

Ma, Z. (2018). *Practical Integration of Data-Driven Models for Production Analysis and Inference of Reservoir Heterogeneities in SAGD Operations*. PhD. Thesis, University of Alberta

Macharia, M. W., Gachari, M. K., Kuria, D. N., & Mariita, N. O. (2017). Low cost geothermal energy indicators and exploration methods in Kenya. *Journal of Geography and Regional Planning, 10(9), 254-265.*

Mania, F.M. (2017). *Estimation of Permeability in Siliciclastic Reservoirs from Well Log Analysis and Core Plug Data; Based on the Data From an Exploration Well*

Offshore Norway. MSc. Thesis, Norwegian University of Science and Technology

Manning, D.E. (2018). Fluids of the Lower Crust: Deep Is Different. Annual Review of Earth and Planetary Sciences

Mariita, N.O. (2013). Geophysical Surveys of High Temperature Fields – A Case for Olkaria and Menengai Geothermal Fields, Kenya, *Presented at Short Course VIII on Exploration for Geothermal Resources. Lake Bogoria and Lake Naivasha, Kenya*

Marwan, Yanis, M., Idroes, R. and Ismail, N. (2019). 2D Inversion And Static Shift Of Mt And Tem Data For Imaging The Geothermal Resources Of Seulawah Agam Volcano, Indonesia. *GEOMATE Journal, 17(62), 173–180.*

Maxhuni, A. (2017). *Managing the Scarcity of Monitoring Data through Machine Learning in Healthcare Domain*. PhD. Thesis, university of Trento.

Mbithi, U.K. (2016). *Interpretation of Feed Zones to Map Sub-Surface Permeability Structures and Natural State Simulation: A Case Study of Olkaria Domes Geothermal System in Kenya.* MSc. Thesis, University of Iceland.

Mohamud, N.S. (2013). 1D Joint Inversion of Tem and Mt Data: Suswa Geothermal Field, Rift Valley, Kenya. *Geothermal Training Programme Reports No. 9.* Orkustofnun, Grensasvegur Iceland.

Mont´ans, F.J., Chinesta, F., Bombarelli, G.R., Kutz, N.J., (2019). Data-driven modeling and learning in science and engineering. *Data-Based Engineering Science and Technology* 347(11): 845-855

Mortensen, A.K. and Axelsson, G. (2013).Developing A Conceptual Model Of A Geothermal System. *Presented at "Short Course on Conceptual Modelling of Geothermal Systems", in Santa Tecla, El Salvador*.

Muther, T., Syed, F. I., Lancaster, A.T., Salsabila, F.D., Dahaghi, A.K. and Negahban, S. (2022). Geothermal 4.0: AI-Enabled Geothermal Reservoir Development-Current Status, Potentials, Limitations, and Ways Forward. *Geothermics Volume 100, https://doi.org/10.1016/j.geothermics.2022.102348*

Mwandigha, R.J. (2020). Unique, Resource-rich Olkaria in Hells Gate National Park The need to have a resource center in Olkaria. *Proceedings, 8th African Rift Geothermal conference,* Nairobi, Kenya

Mwangi, A.W., Mickus, K. and Serpa, L. (2018). Dimensionality Analysis of the Olkaria Geothermal Field, East Africa Rift. *Proceedings of the 7th African Rift Geothermal Conference*, Kigali, Rwanda

Mwaura, D.W. (2018). *Exploration and Optimized Siting of Geothermal Wells Using a Web-Based Spatial Decision Support System: A Case Study of the Olkaria Geothermal Field*. MSc. Thesis Technical University of Berlin.

Mwawasi, H.M. (2018). Geothermal Resource Assessment – Case Example, Menengai Phase I. *Presented at SDG Short Course III on Exploration and Development of Geothermal Resources, at Lake Bogoria and Lake Naivasha, Kenya.*

Nabi, A., Liu, X., Gong, Z. & Ali, A. (2020) Electrical Resistivity Imaging of Active Faults In Palaeoseismology: Case Studies From Karachi Arc, Southern Kirthar Fold Belt, Pakistan, *NRIAG Journal of Astronomy and Geophysics*, 9:(1), 116-128, DOI: 10.1080/20909977.2020.1722524

Naderi, M. (2019). Estimating Confined Aquifer Parameters Using a Simple Derivative-Based Method. *Earth Science,* Volume 5, Issue 10 doi: 10.1016/j.heliyon.2019.e02657

Neupert, T., Mark, H., Fischer, M.H., Greplova, E., Choo, K. and Denner, M. (2021). *Introduction to Machine Learning for the Sciences, Lecture Notes,* Department of Physics, University of Zurich

Ngarambe, J., Irakoze, A., Yun, Y.G. and Kim, G. (2020). Comparative Performance of Machine Learning Algorithms in the Prediction of Indoor Daylight Illuminances, *Sustainability* 12, 4471

Nguimbous-Kouoh, J.J., Ndougsa-Mbarga, T. and Manguelle-Dicoum, E. (2018). Audio-Frequency Magnetotelluric Prospecting in the Mamfe Sedimentary Basin of Southwestern Cameroon. *Int J Earth Sci Geophys* 4:020

Nisbet, R., Miner, G. andYale, K. (2018). Handbook of Statistical Analysis and Data Mining Applications: *Second Edition Academic Press, Elsevier ISBN 978-0-12-416632-5*

Nordiana, M.M., M a s linda, U., Hisham, H. and Nuramalina, M.K.A. (2016). The Effect of 2-D Resistivity on Magnetic Field, *International Research Journal of Engineering and Technology (IRJET)*, 3(1): 43-47

Nyandigisi, J.A (2020). Geothermal Wells Productivity Controls Case Study of Wells Drilled in the Greater Olkaria Geothermal Area, Kenya. *Proceedings, 8th African Rift Geothermal Conference Nairobi, Kenya.*

Nyandigisi, J.A. and Katana, C. (2016). High Temperature Hydrothermal Alteration in Active Geothermal Systems a Case Study of Olkaria Domes. *Proceedings of the 6th African Rift Geothermal Conference* Addis Ababa, Ethiopia.

Nzioka, K.F. (2017). *Technical and Economic Assessment of Adopting Under-Balanced Drilling at the Olkaria Geothermal Field*. MSc. Thesis, Jomo Kenyatta University of Agriculture and Technology

Ojha, L., Karunatillake, S., Karimi, S.and Buffo, J. (2021). Amagmatic Hydrothermal Systems on Mars from Radiogenic Heat. *Nat Commun* 12**,** 1754 (2021). https://doi.org/10.1038/s41467-021-21762-8

Okoo, J.A. (2013). Borehole Geology And Hydrothermal Alteration Mineralogy of Well OW-39A, Olkaria Geothermal Project, Naivasha, Kenya. *Geothermal Training Programme Reports 2013 Orkustofnun, Grensasvegur 9, Number 24*. Reykjavik, Iceland.

Okoo, J., Omiti, A., Kamunya, K. and Saitet, D. (2017). Updated Conceptual Model of Olkaria Geothermal Field Naivasha, Kenya. GRC Transactions, 41: 2017

Omenda, P., Mangi, P., Ofwona, C. and Mwangi, M. (2021). Country Update Report for Kenya 2015-2019. *Proceedings World Geothermal Congress Reykjavik, Iceland*

Omenda, P.A. (2010). Geothermal Exploration in Kenya. *Presented at Short Course V on Exploration for Geothermal Resources*. Lake Bogoria and Lake Naivasha, Kenya.

Orchel, M. (2011). Regression Based on Support Vector Classification. 6594. 353-362. 10.1007/978-3-642-20267-4_37.

Ortiz, R.J., Adúriz, B.A. & Greca, I.M. (2020) A Framework for Epistemological Discussion on Integrated STEM Education. Sci & Educ 29: 857–880.

Ouma, P.A. (2012). Geothermal Exploration and Development of the Olkaria Geothermal Field. *Exploration for Geothermal Resources.* Lake Bogoria and Lake Naivasha, Kenya.

Parada, A.F.M. (2016). Phases of Geothermal Development. *Sustainability and Environmental Management of Geothermal Resource Utilization and the Role of Geothermal in Combating Climate Change*. Santa Tecla, El Salvador.

Power Technology, (2018). Kenya is betting big on geothermal installations, https://www.power-technology.com

Pupale, R. (2018). Support Vector Machines (SVM) — An Overview. *https://towardsdatascience.com/https-medium-com-pupalerushikesh-svm-f4b42800e989*

Rifat, N.I. (2020). *Feature Engineering on the Cybersecurity Dataset for Deployment on Software Defined Network*. MSc. Thesis, North Dakota State University of Agriculture and Applied Science

Riley, S. N. (2018). *Numerical Modeling of Cold Water Injection into Supercritical Geothermal Reservoirs*. MSc. Thesis Reykjavík University 2018.

Rionomakal, E., Mariita, N., Varet, J and Namaswa, S. (2018). Pressure Transients Test Analysis of Well OW-918 at Olkaria Domes Geothermal Field. *International Journal of Science and Research, 8 (8)*

Rogei, D.S. (2021). *Mediating Maendeleo: Examining the nexus between geothermal extraction, wildlife conservation and community well-being in Olkaria-Suswa, Southern Kenya.* PhD Thesis Carleton University Ottawa, Ontario

Rosenkjær, G.K. (2011). Electromagnetic Methods in Geothermal Exploration. 1D and 3D Inversion of TEM and MT Data from a Synthetic Geothermal Area and the Hengill Geothermal Area, SW Iceland. *MSc. Thesis, University of Iceland*

Sakindi, G. (2015). *Three-Dimensional Inversion of Magnetotelluric Data: Geological/ Geothermal Interpretation of Asal Geothermal Field, Djibouti*. MSc. Thesis, University of Iceland.

Sapkota, R.S. (2019). *Machine Learning and Artificial Intelligence in Production Technology.* Bsc. Thesis, Oulu University of Applied Sciences.

Sarker, H.I. (2021). Machine Learning: Algorithms, Real-World Applications and Research Directions, *SN Computer Science* (2) 160 (2021). https://doi.org/10.1007/s42979-021-00592-x

Schmidt, J., Marques, M.R.G. and Botti, S. (2019). Recent advances and applications of Machine Learning in solid-state materials science. *npj Comput Mater* 5, 83 (2019).

Schölderle, F., Lipus, M., Pfrang, D., Reinsch, T., Haberer, S., Einsiedl, F. and Zossede, K. (2021). Monitoring cold water injections for reservoir characterization using a permanent fiber optic installation in a geothermal production well in the Southern German Molasse Basin. *Geotherm Energy* 9, 21 (2021). *https://doi.org/10.1186/s40517-021-00204-0*

Scowen, M., Athanasiadis, N.I., Bullock, M.J., Eigenbrod, F. and Willcock, S. (2021). The current and future uses of machine learning in ecosystem service research, *Science of The Total Environment,Volume 799, https://doi.org/10.1016/j.scitotenv.2021.149263.*

Seyedrahimi, N.M., Doulati, A.F. and Noorollahi, Y. (2017). Development of an updated geothermal reservoir conceptual model for NW Sabalan geothermal field, Iran. *Geotherm Energy* 5, 14 (2017).

Siebert, J., Joeckel, L., Heidrich, J Trendowicz, A., Nakamichi, K., Ohashi, K., Namba, I., Yamamoto, R. & Aoyama, M. (2021). Construction of a quality model for Machine Learning systems. *Software Quality Journal* (2021)

Sirma, R.C. (2019). *Analysis Of The Geothermal Potential of The North West Olkaria Prospect (Oserian Sanctuary)In Kenya, Using Transient Electromagnetic Method And Magnetotelluric Method*. MSc. Thesis Egerton University, Kenya.

Spivey, J. P., Lee, W. J., and Sheikh, M.H. (2020). Transient Volume of Investigation: Definition, Theory, and Applications. *Presented at the SPE Annual Technical Conference and Exhibition*, doi: https://doi.org/10.2118/201414-MS

Steuer, F. (2018). *Machine Learning for Public Policy Making. How to Use Data-Driven Predictive Modeling for the Social Good*. MSc. Thesis, Institut Barcelona Estudis Internacionals.

Tahir, M.A., (2019). *Proxy-model for Flow and Transport in Geothermal Reservoirs.* MSc. Thesis, Delft University of Technology.

Thanh, L.D., Jougnot, D., Do, V.P. and Nghia, A.V.N., (2019). A Physically Based Model for the Electrical Conductivity of Water-Saturated Porous Media. *Geophysical Journal International*, 219 (2): 866–876.

Vehling, F., Hasenclever, J. & Rüpke, L. (2020). Brine Formation and Mobilization in Submarine Hydrothermal Systems: Insights from a Novel Multiphase Hydrothermal Flow Model in the System H2O–NaCl. *Transp Porous Med* **136,** 65–102 (2021). https://doi.org/10.1007/s11242-020-01499-6

Venkatasubramaniam, A., Wolfson, J., Mitchell, N. Barnes, T., JaKa, M. and French, S (2017). Decision trees in epidemiological research. *Emerg Themes Epidemiol* 14, 11 (2017). https://doi.org/10.1186/s12982-017-0064-4

Vihma, T., Kilpeläinen, T., Manninen, M., Sjöblom,A., Jakobson, E., Palo, T., Jaagus,J. and Maturilli, M. (2011). Characteristics of Temperature and Humidity Inversions and Low-Level Jets over Svalbard Fjords in spring *Hindawi Publishing Corporation Advances in Meteorology Volume 2011*

Wamalwa, R. , Nyamai, C. , Ambusso, W. , Mulwa, J. and Waswa, A. (2016) Structural Controls on the Geochemistry and Output of the Wells in the Olkaria Geothermal

Field of the Kenyan Rift Valley. *International Journal of Geosciences*, 7, 1299-1309. doi: 10.4236/ijg.2016.711094

Wamriew, D.S. (2019). *Magnetotelluric and Transient Electromagnetic Imaging for Geothermal Resources in Arus-Bogoria Area in Kenya*. MSc. Thesis, Kenyatta University.

Wang, H. and Sharma, M.M. (2018). Estimating Unpropped-Fracture Conductivity and Fracture Compliance from Diagnostic Fracture-Injection Tests. *SPE Journal*, 23(05):1648-1668

Wanjohi, A.W. (2014). Geophysical survey of a high-temperature field, Olkaria. *Exploration for Geothermal Resources.* Kenya.

Warega, J.A. (2019). *Geothermal Prospecting of Olkaria Dome Areas in Naivasha, Nakuru County Kenya Using Gravity Method*. MSc. Thesis, Kenyatta University.

Weisenberger, T.B., Ingimarsson, H., Hersir, G.P. and Flóvenz, O.G. (2020). Cation-Exchange Capacity Distribution within Hydrothermal Systems and Its Relation to the Alteration Mineralogy and Electrical Resistivity. Energies 2020,13, 5730

Willmott, C.J. and Matsuura, K. (2005). Advantages of the Mean Absolute Error (MAE) Over the Root Mean Square Error (RMSE) in Assessing Average Model Performance. *Climate Research* 30 (1): 79-82

World Bank, (2016). World Development Indicators, 2016. *Washington D.C: The World Bank*

Yakovlev, D., Yakovlev, A. and Valyasina, O. (2018). Permafrost Study In The Northern Margin Of The Siberian Platform Based On Regional Geoelectric Survey Data In Russian. *Earth's Cryosphere*. 5:77-95.

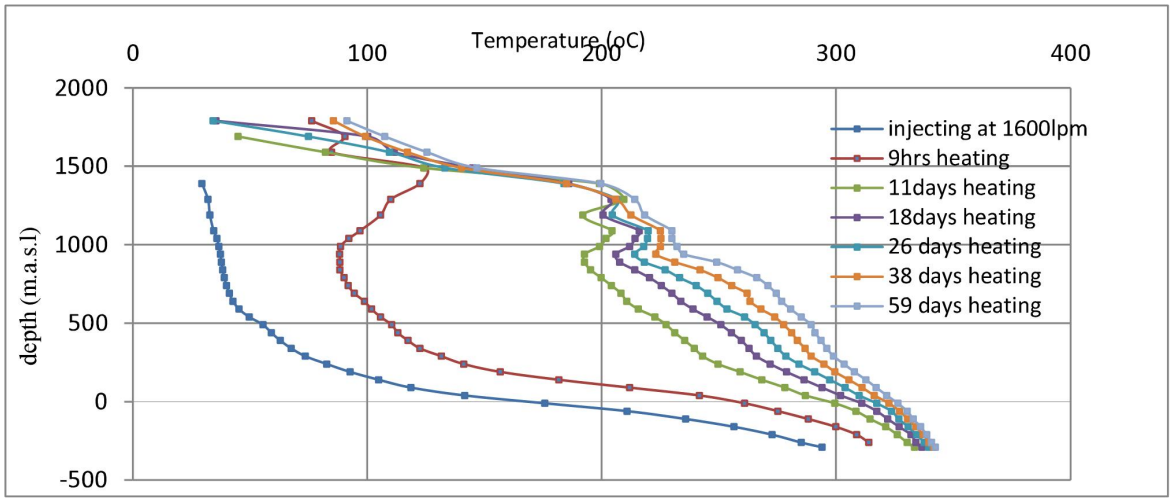
Zarrouk, S.J. and McLean, K. (2019). Advanced analytical pressure-transient analysis relevant to geothermal wells. Geothermal Well Test Analysis.

Zhu, Z., Lei, X., Xu, N., Shao, D., Jiang, X and Wu, X. (2020). Integration of 3D Geological Modeling and Geothermal Field Analysis for the Evaluation of Geothermal Reserves in the Northwest of Beijing Plain, China. *Water 2020*,12, 638

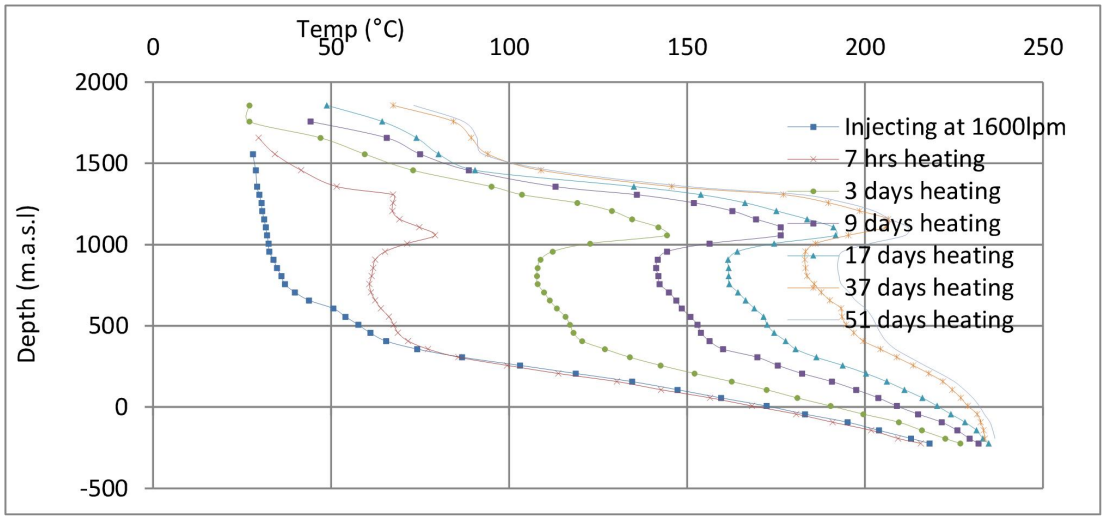
APPENDICES

Appendix I: Well Recovery Profiles

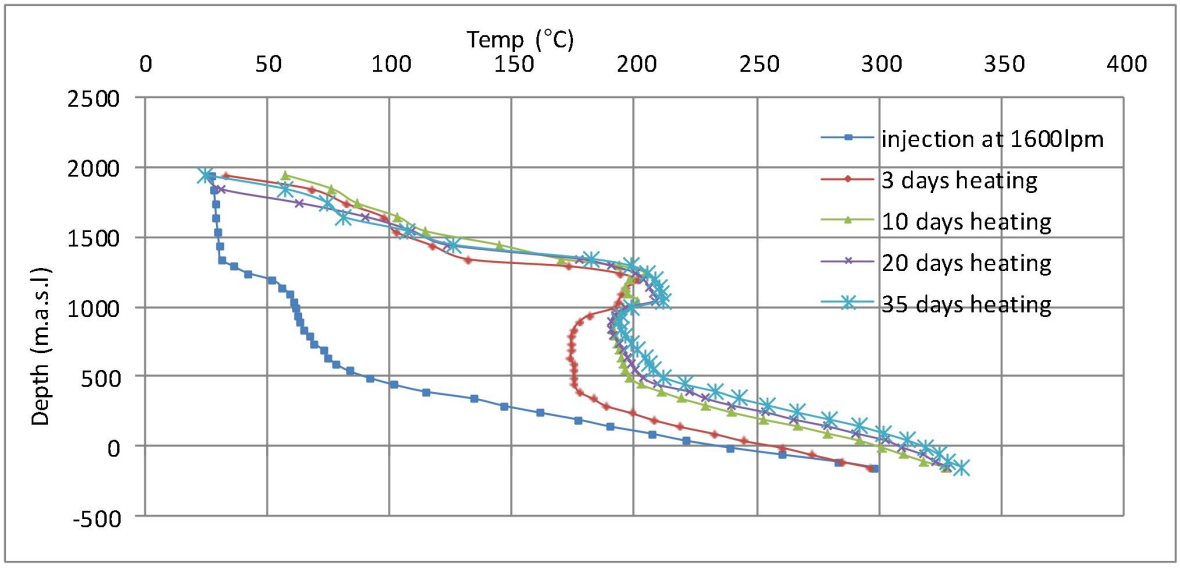
Appendix Ia: Temperature recovery profiles in OW-901



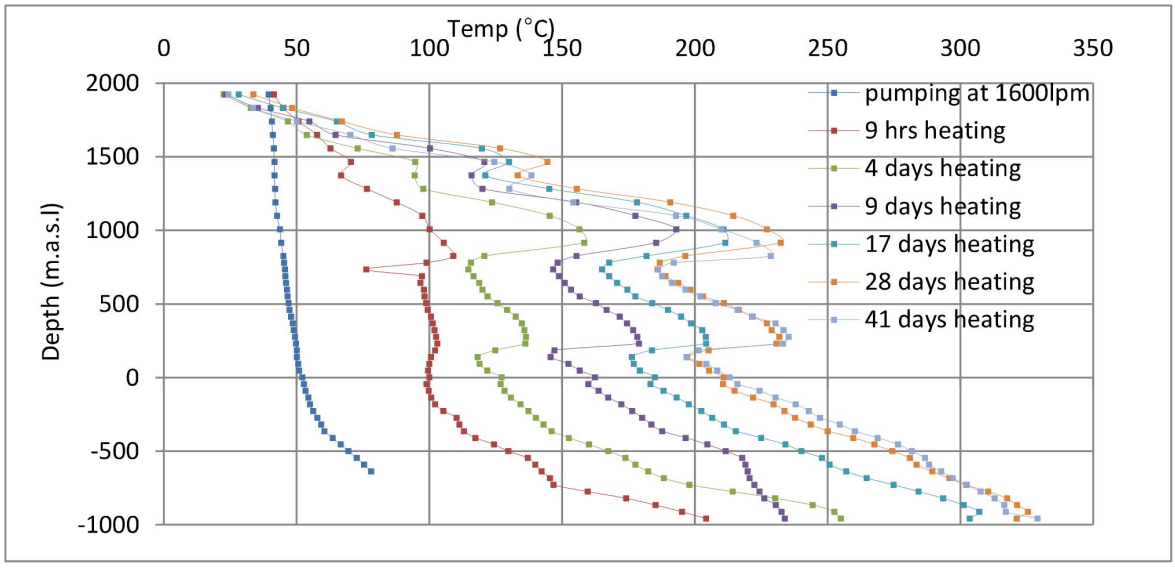
Appendix Ib: Temperature recovery profiles in OW-902B



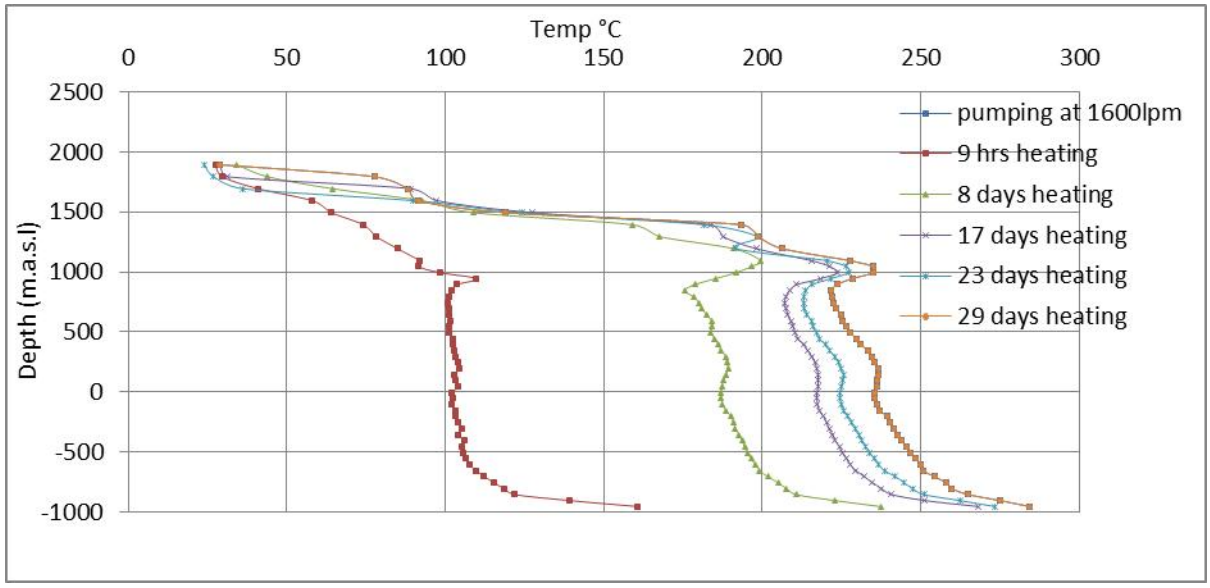
Appendix Ic: Temperature recovery profiles in OW-903



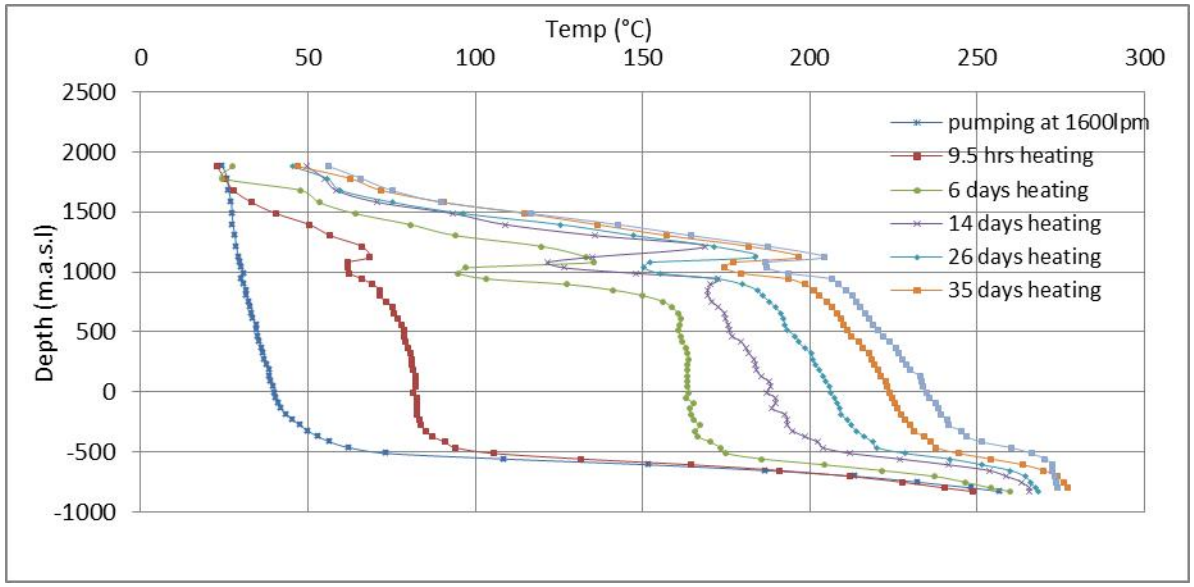
Appendix Id: Temperature recovery profiles in OW-908



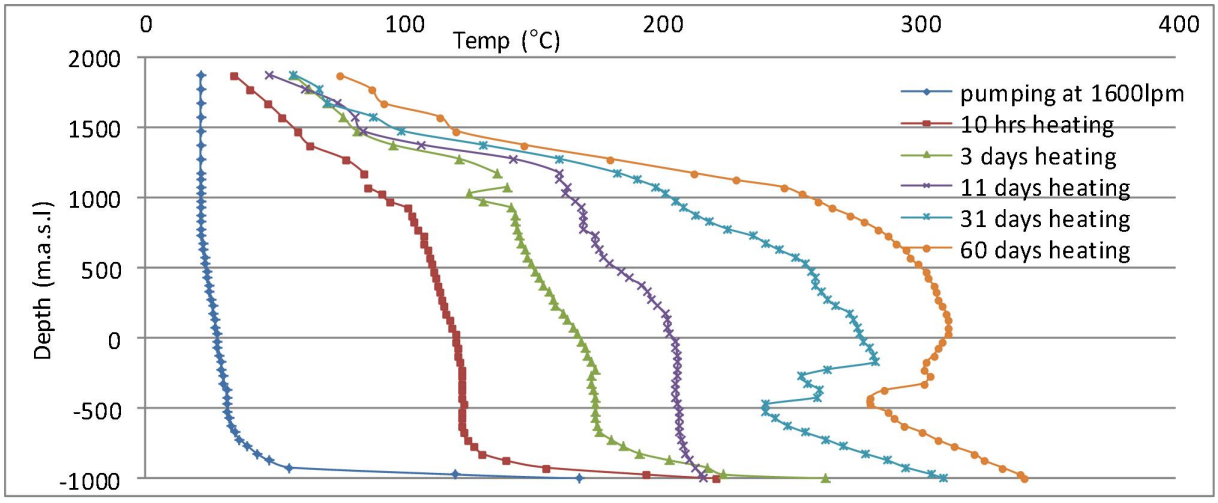
Appendix Ie: Temperature recovery profiles in OW-910



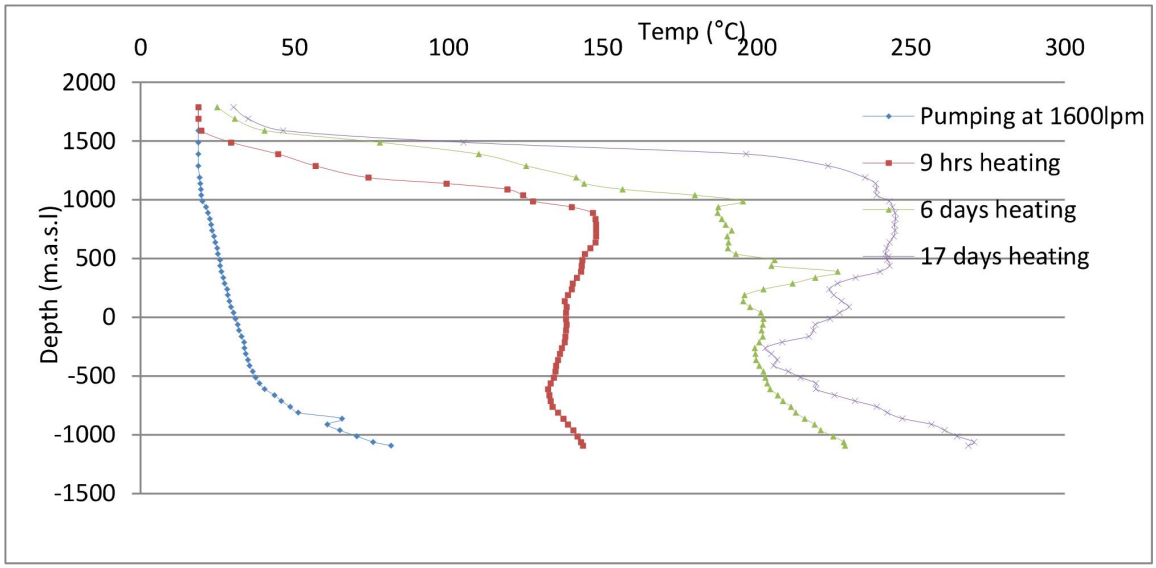
Appendix If: Temperature recovery profiles in OW-911A



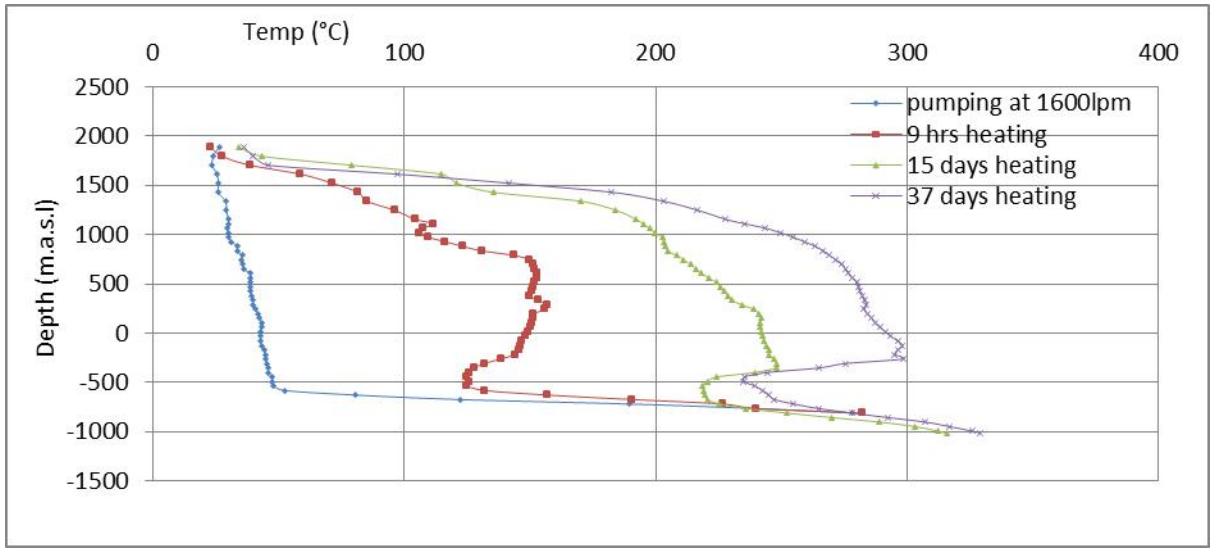
Appendix Ig: Temperature recovery profiles in OW-912



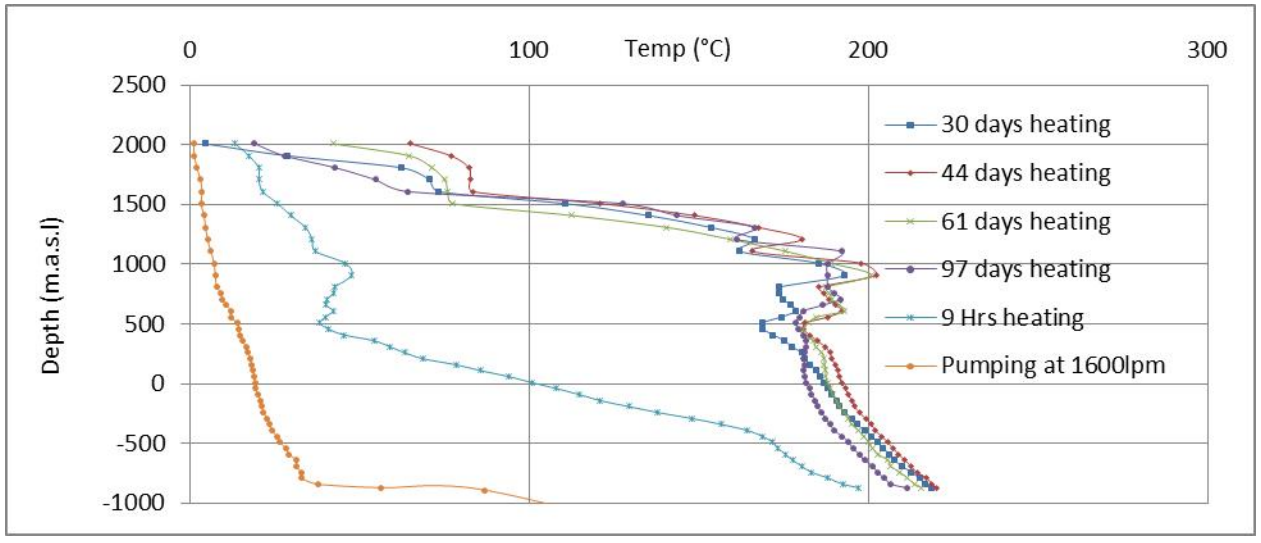
Appendix Ih: Temperature recovery profiles in OW-914B



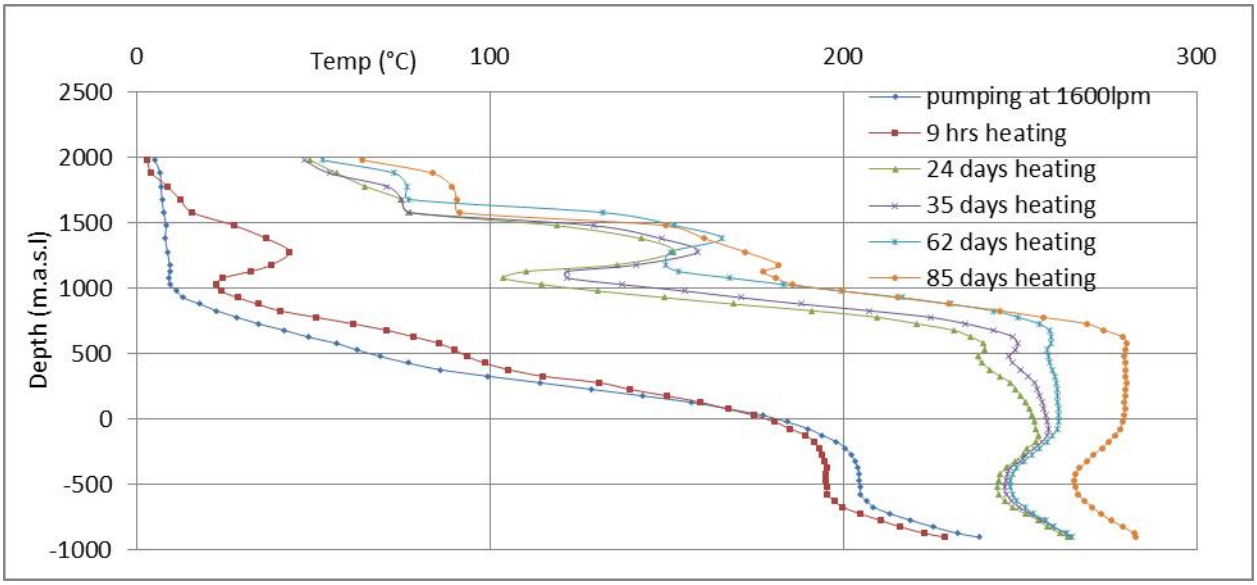
Appendix Ii: Temperature recovery profiles in OW-915



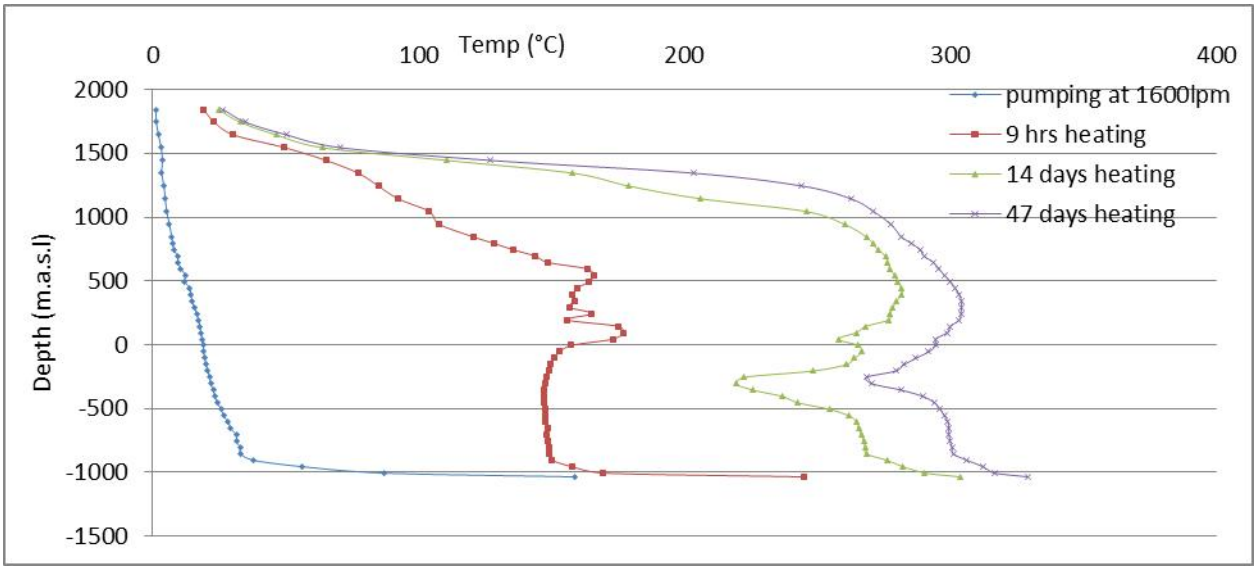
Appendix Ij: Temperature recovery profiles in OW-917



Appendix Ik: Temperature recovery profiles in OW-918

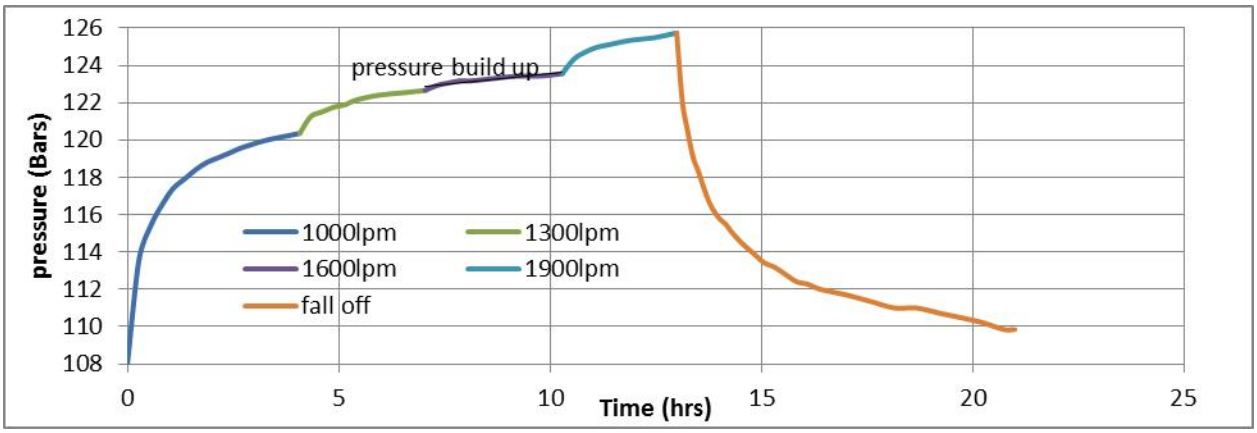


Appendix Il: Temperature recovery profiles in OW-921

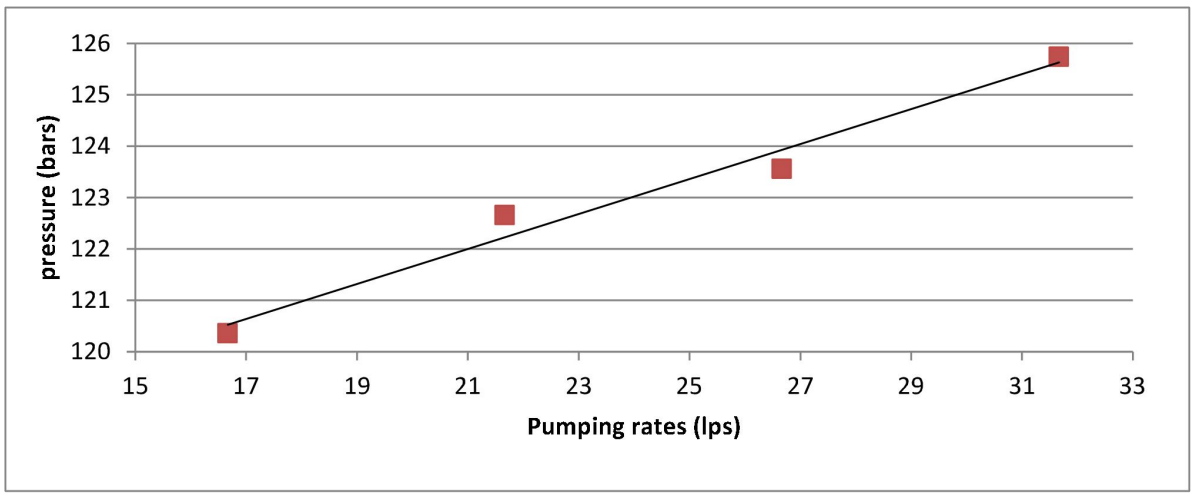


Appendix II: Injection and fall off Graphs

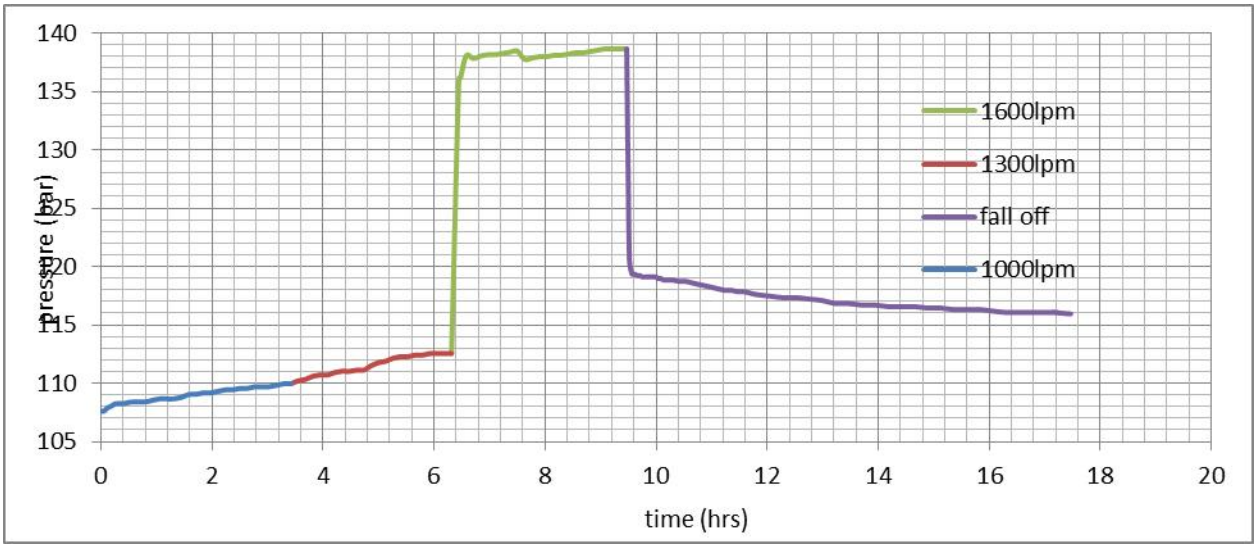
Appendix IIa (i): Pressure Build Up in OW-902B



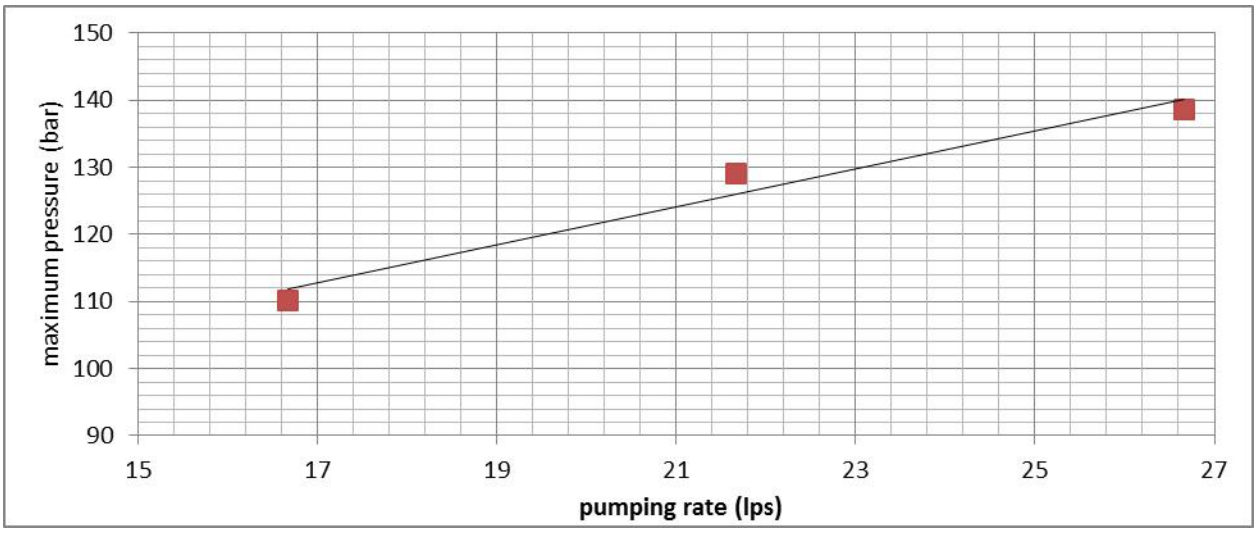
Appendix IIa (ii): Injectivity in OW-902B



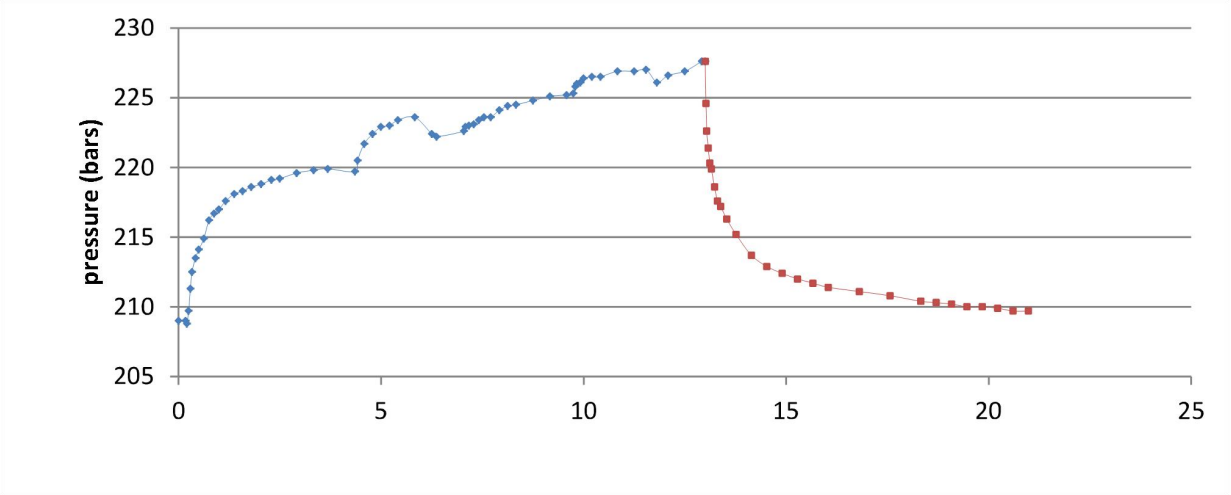
Appendix IIb (i): Pressure Build Up in OW-903



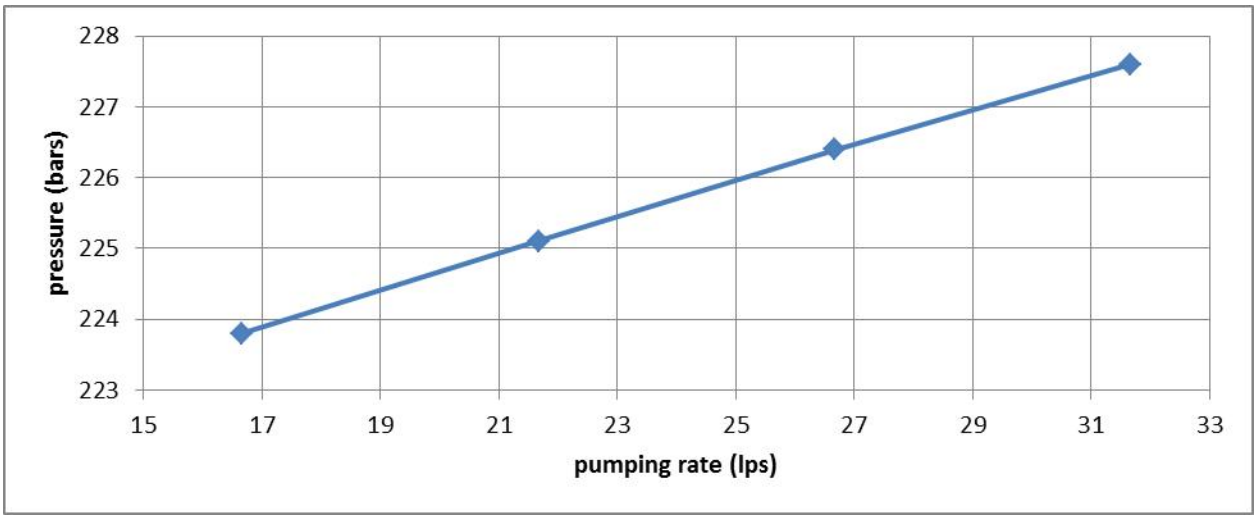
Appendix IIb (ii): Injectivity in OW-903



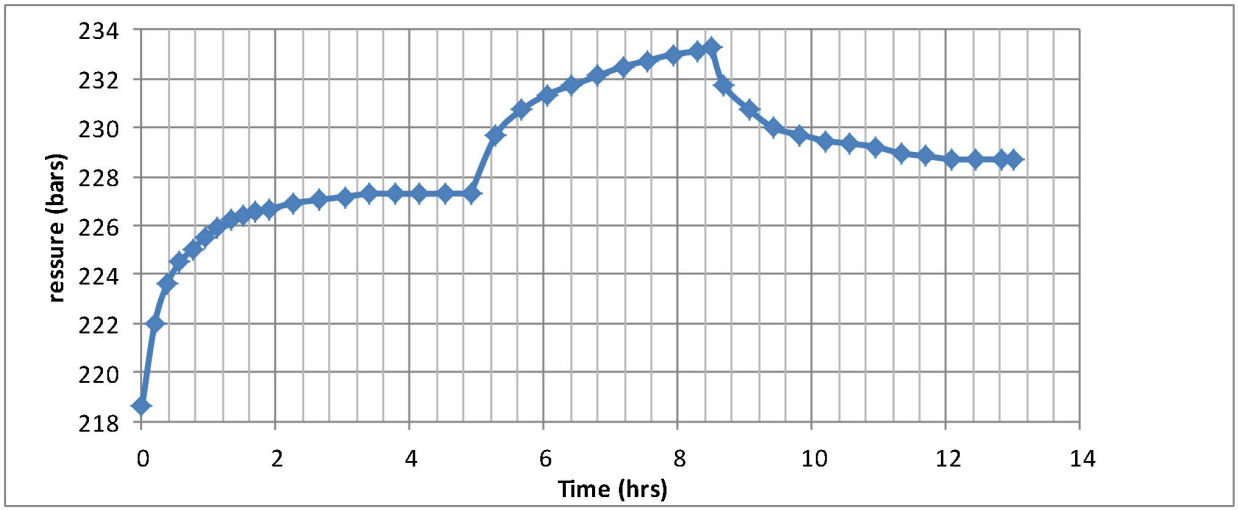
Appendix IIc (i): Pressure Build Up in OW-905A



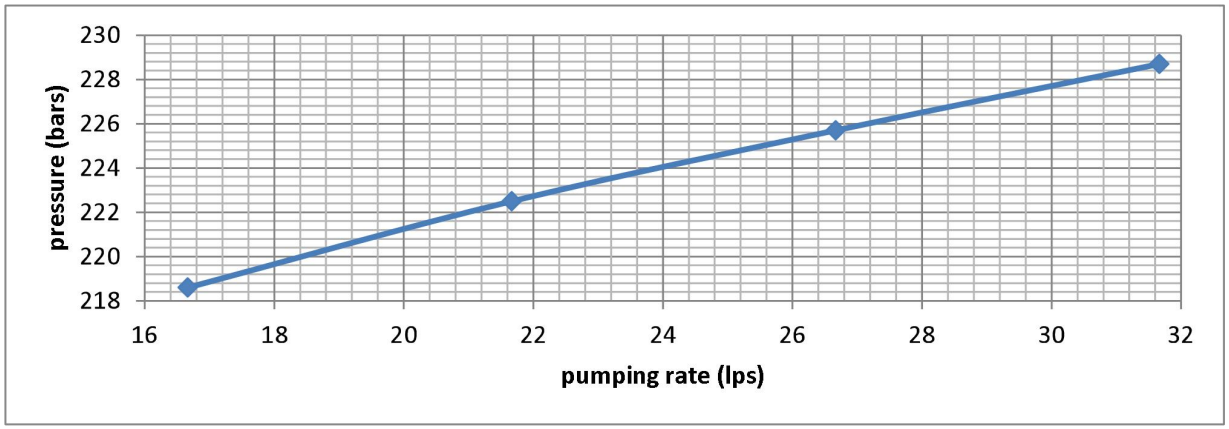
Appendix IIc (ii): Injectivity in OW-905A



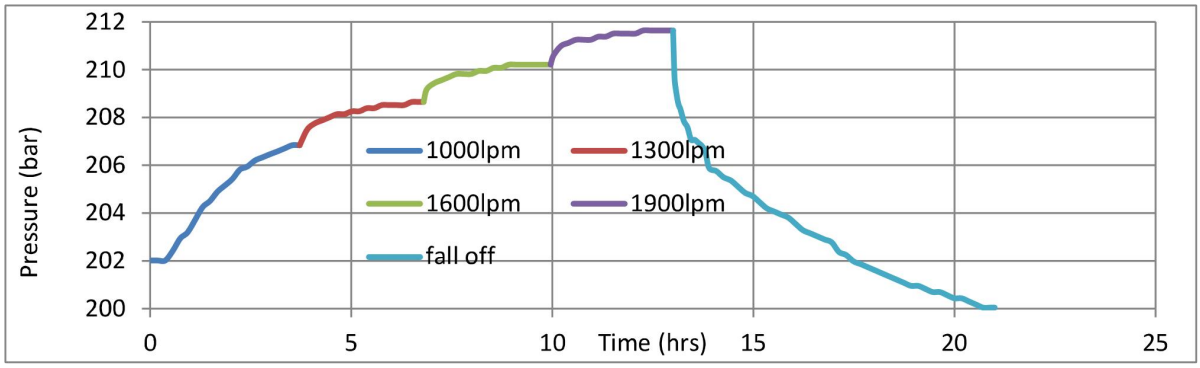
Appendix IId (i): Pressure Build Up in OW-908



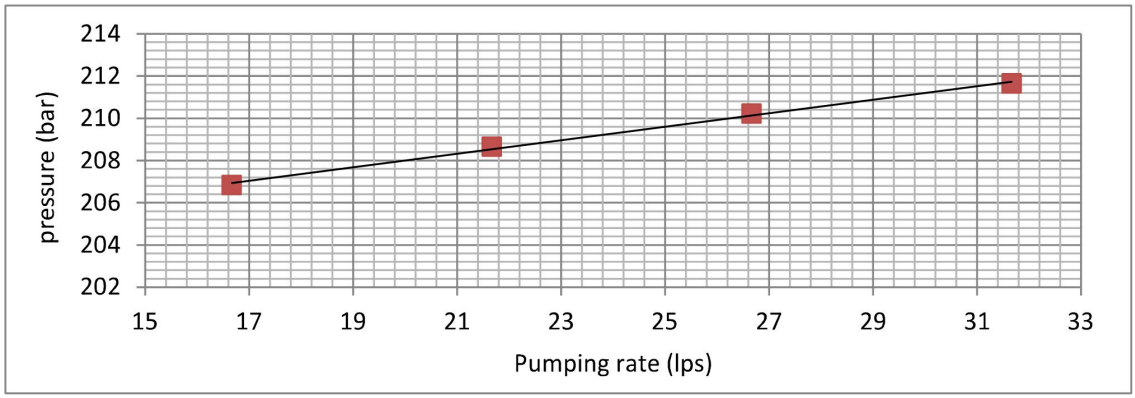
Appendix IId (ii): Pressure Build Up in OW-908



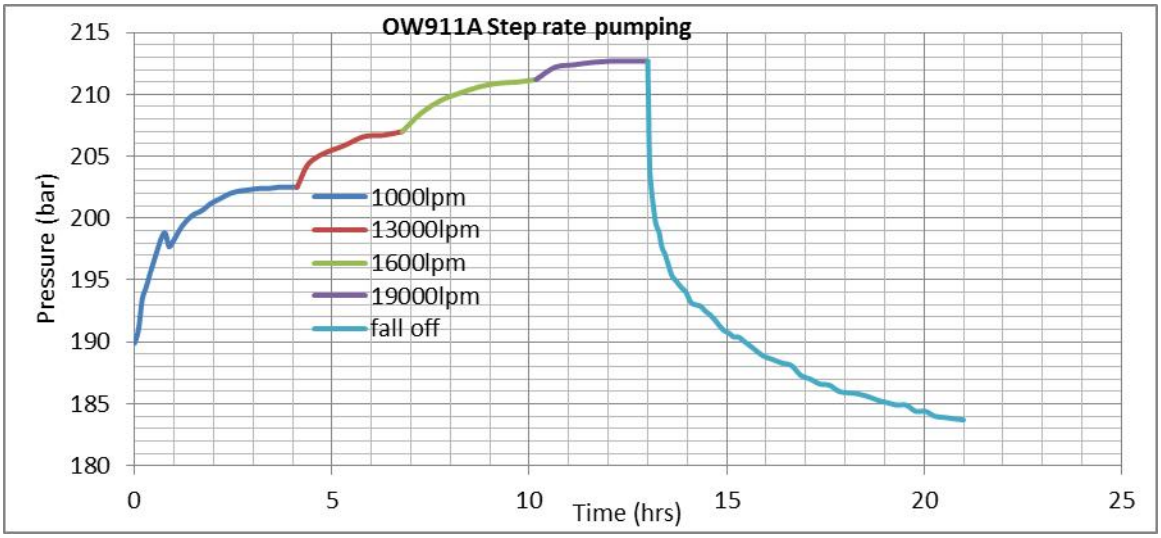
Appendix IIe (i): Pressure Build Up in OW-910



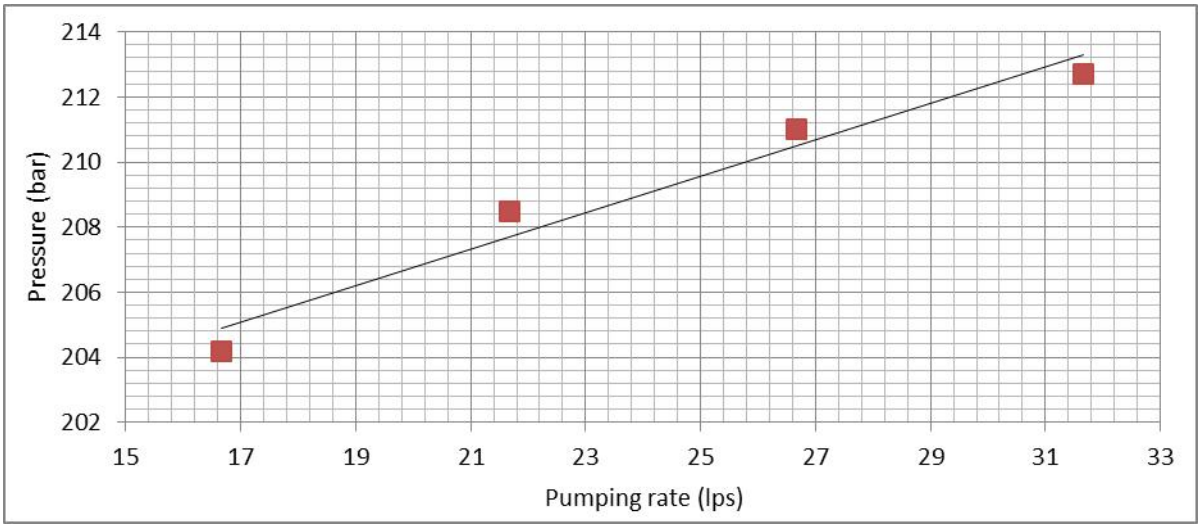
Appendix IIe (ii): Injectivity in OW-910



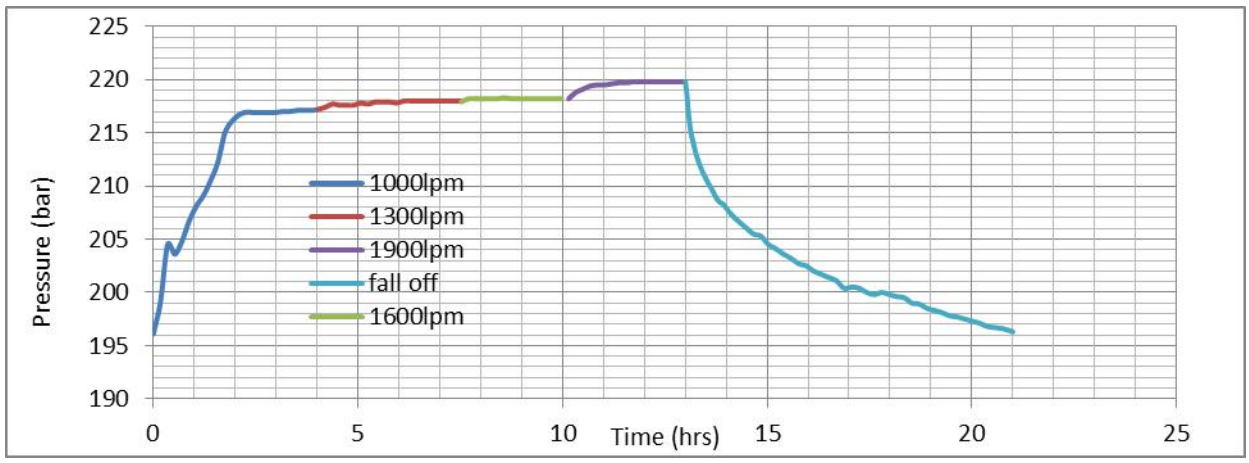
Appendix IIf (i): Pressure Build Up in OW-911A



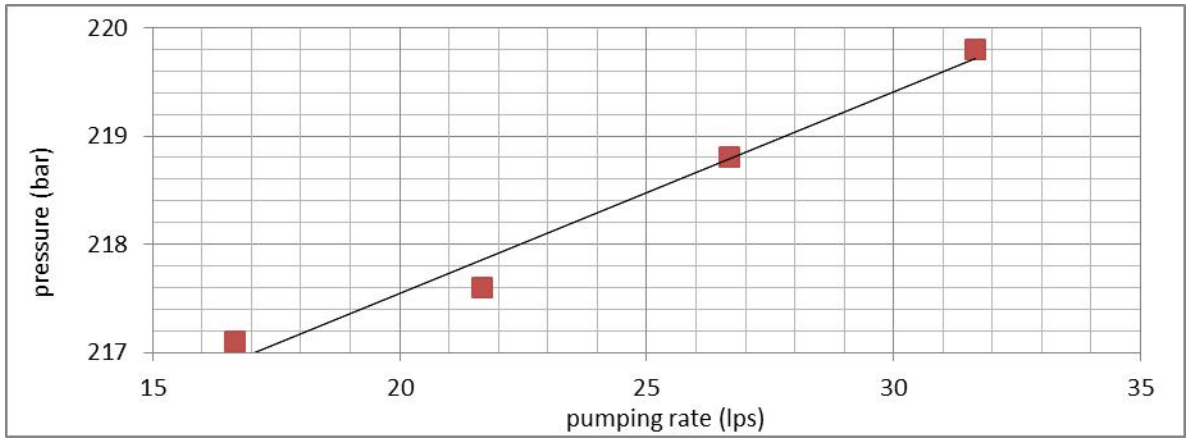
Appendix IIf (ii): Injectivity in OW-911A



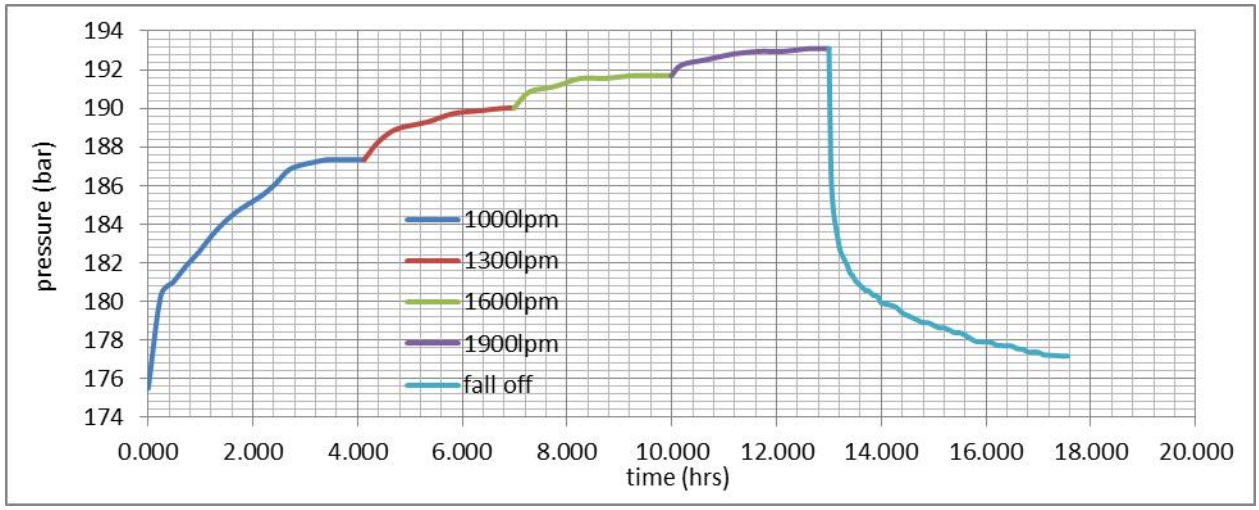
Appendix IIg (i): Pressure Build Up in OW-912



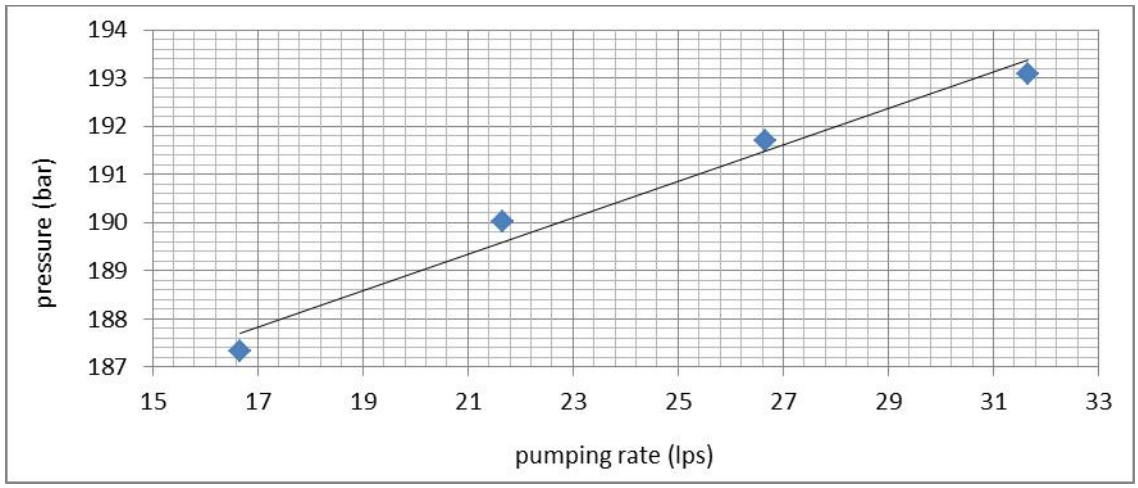
Appendix IIg (ii): Injectivity in OW-912



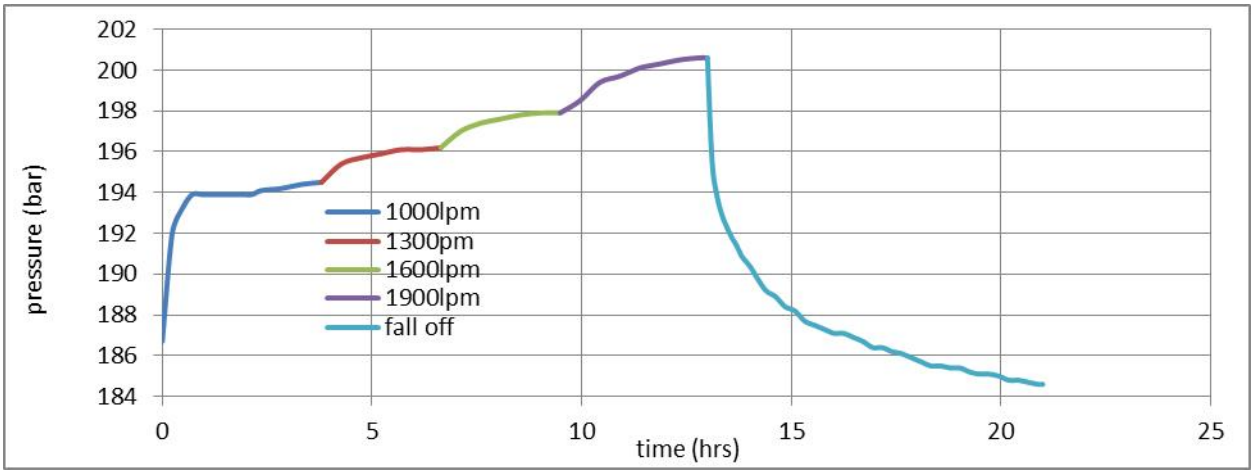
Appendix IIh (i): Pressure Build Up in OW-914B



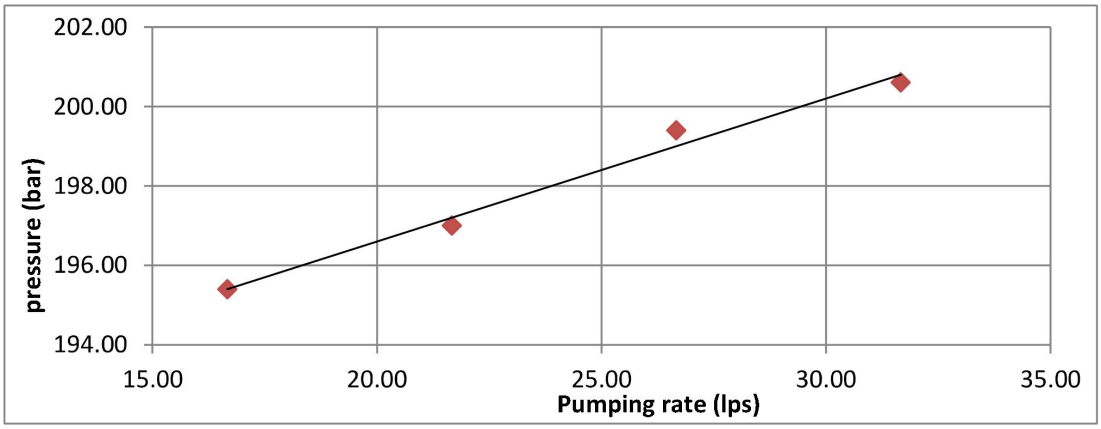
Appendix IIh (ii): Injectivity in OW-914B



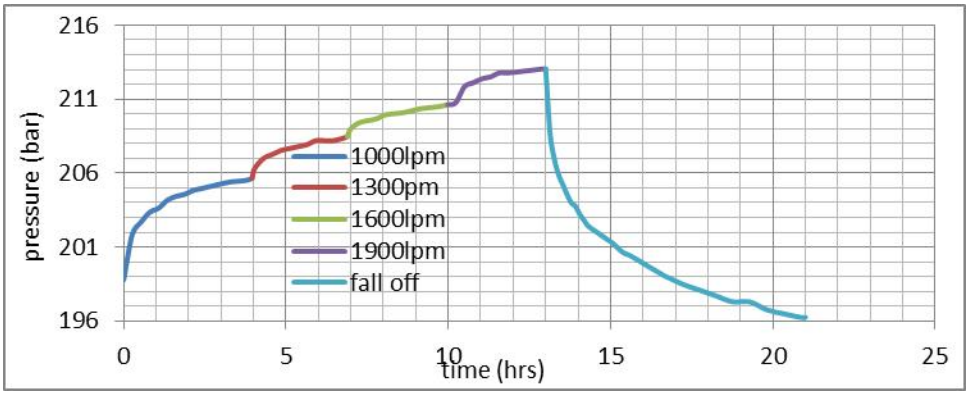
Appendix IIi (i): Pressure Build Up in OW-915



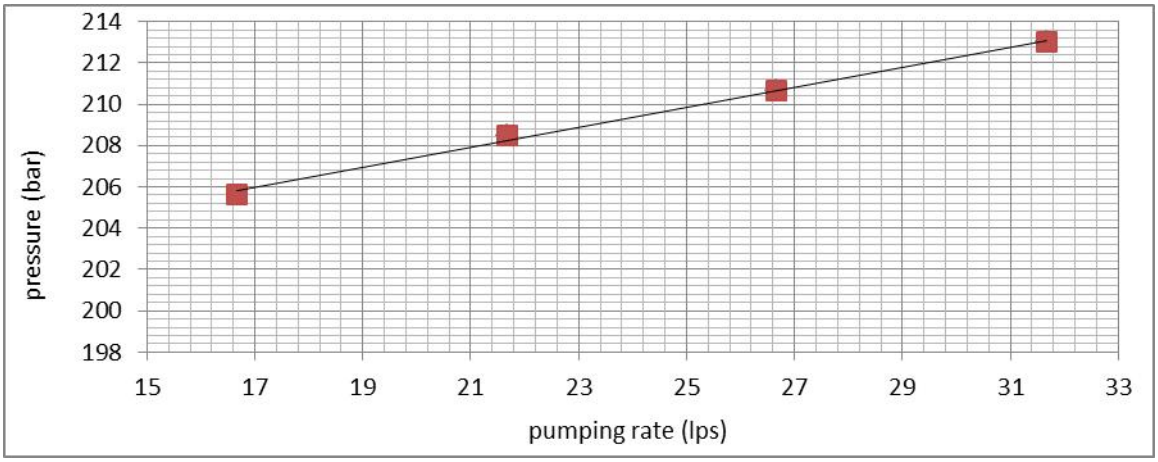
Appendix IIi (ii): Injectivity in OW-915



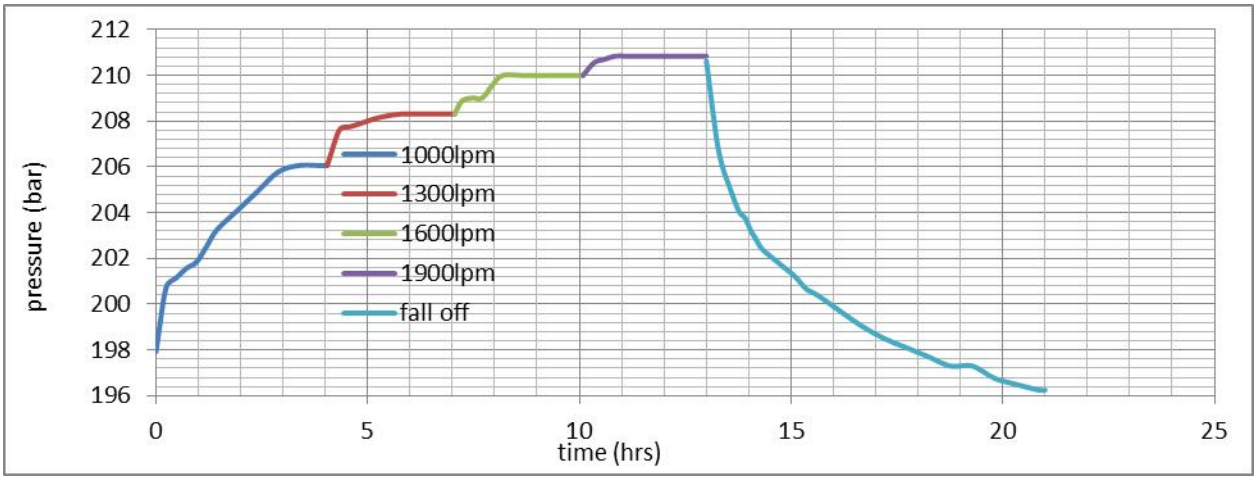
Appendix IIj (i): Pressure Build Up in OW-918



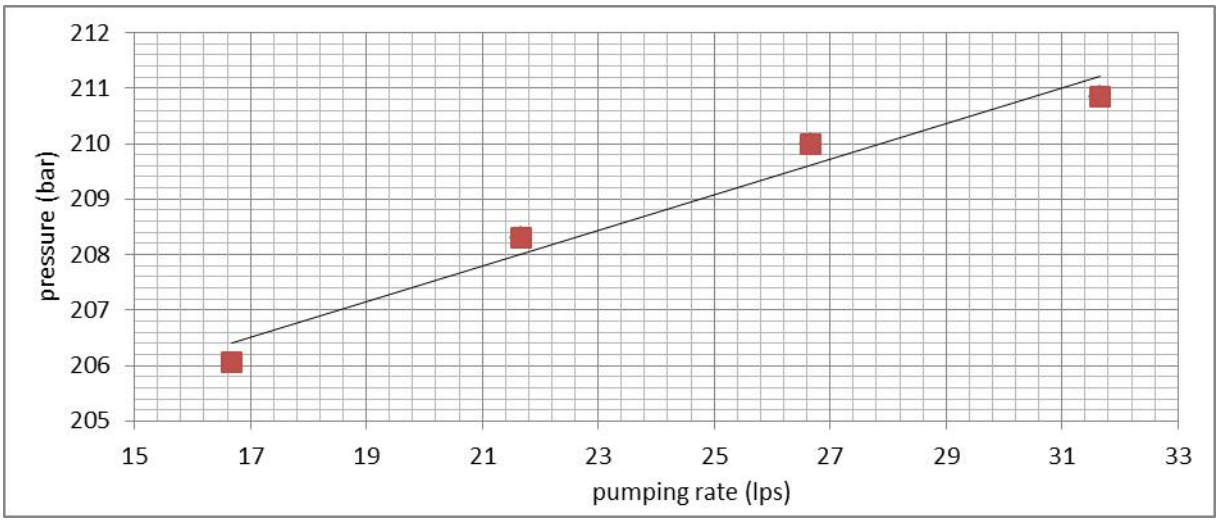
Appendix IIj (ii): Injectivity in OW-918



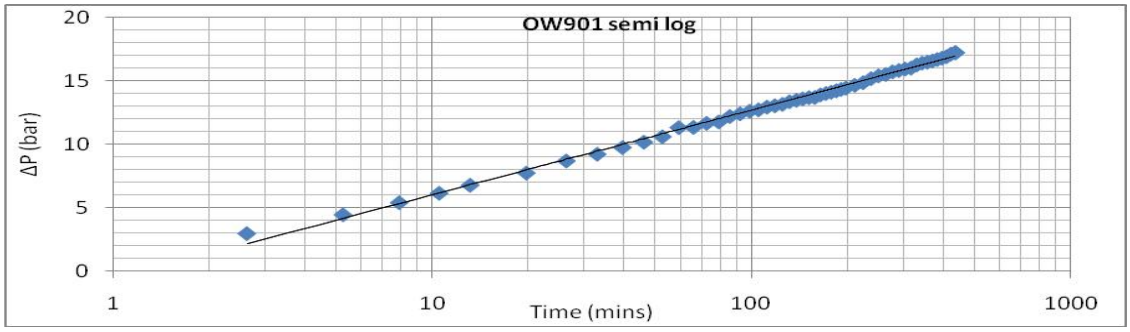
Appendix IIk (i): Pressure Build Up in OW-921

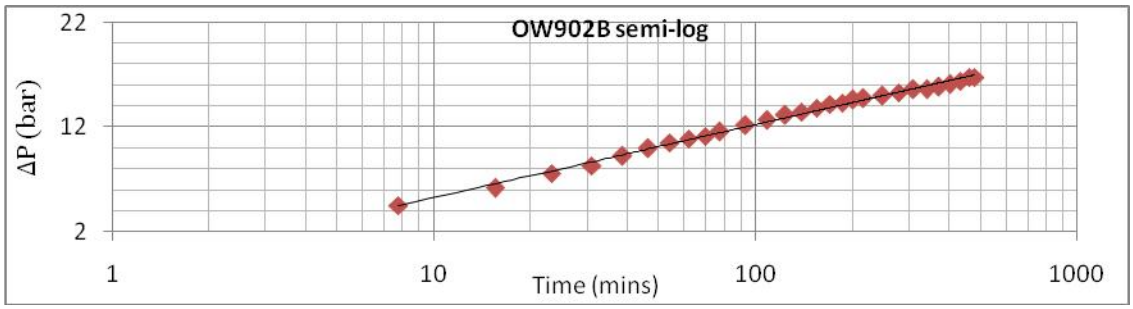


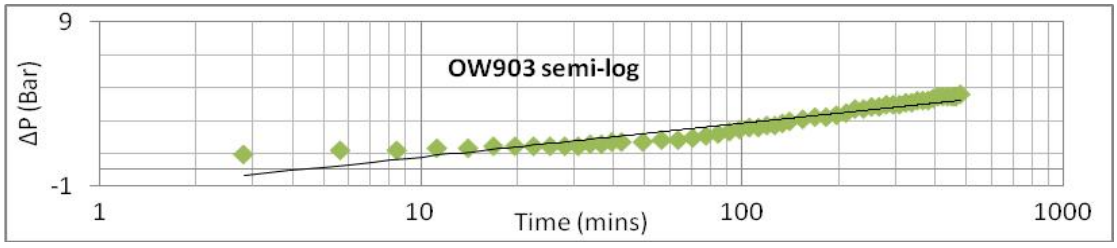
Appendix IIk (ii): Injectivity in OW-921

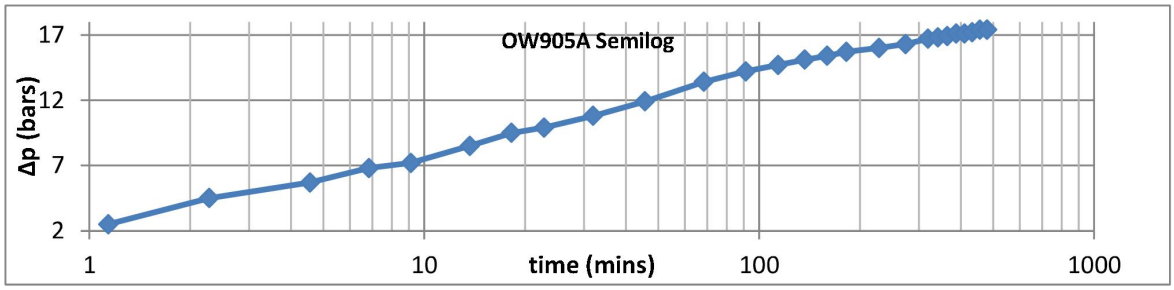


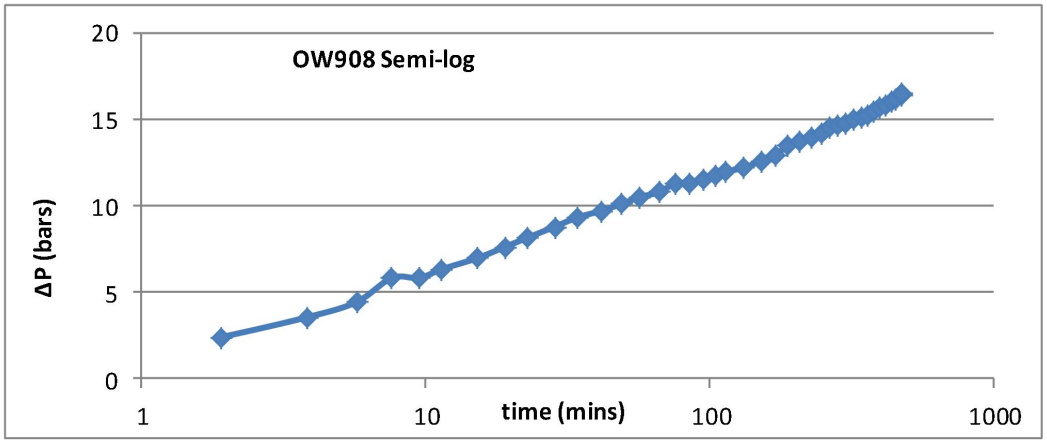
Appendix III: SemiLog pressure Fall off Graphs

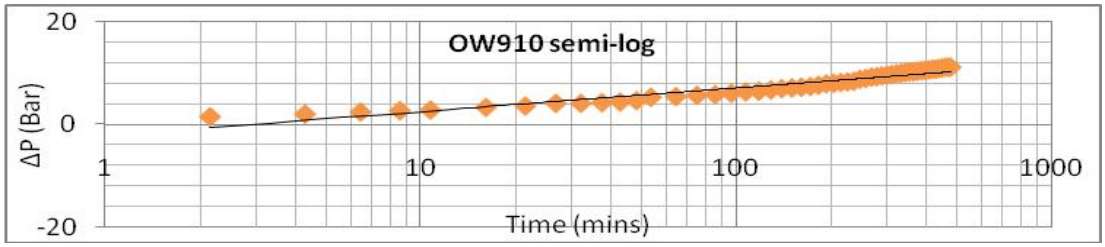


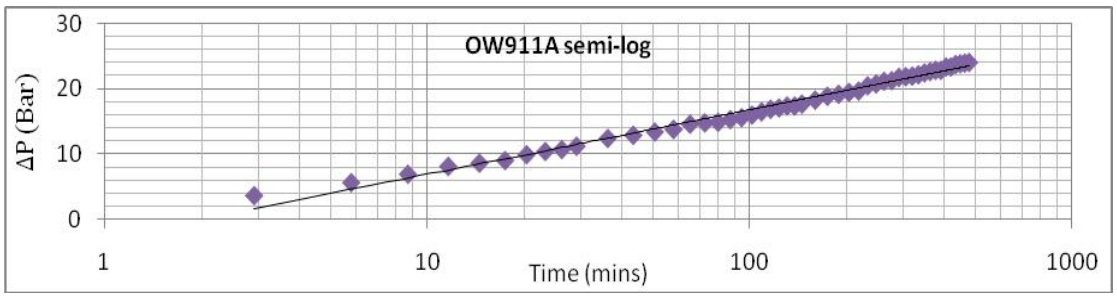


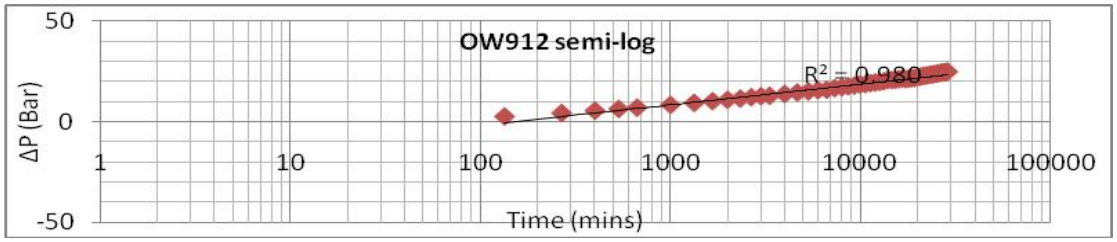


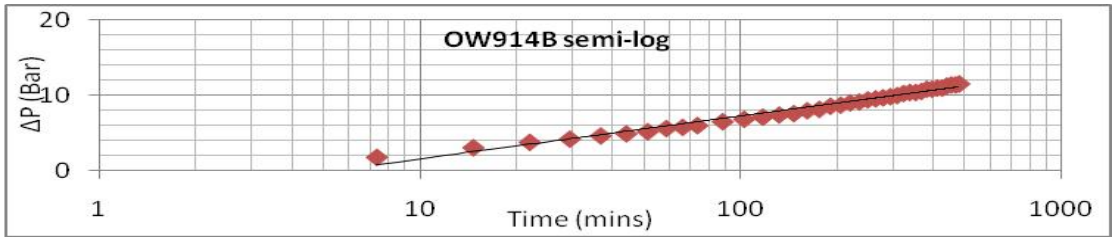

OW905A Semilog
Δp (bars)
17
12
7
2
1
10
time (mins)
100
1000

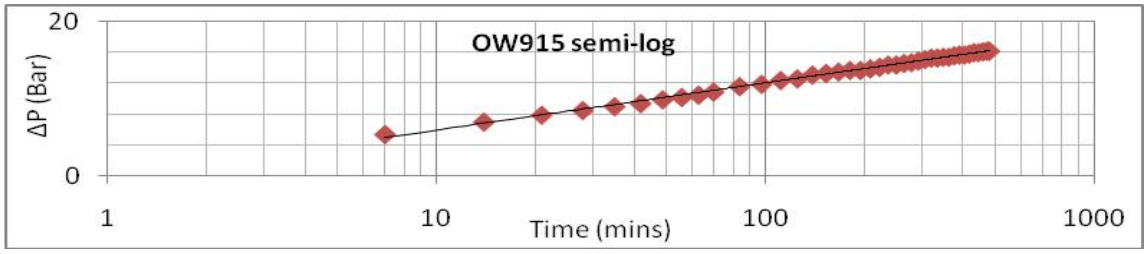

OW908 Semi-log
ΔP (bars)
20
15
10
5
0
1
10
time (mins)
100
1000

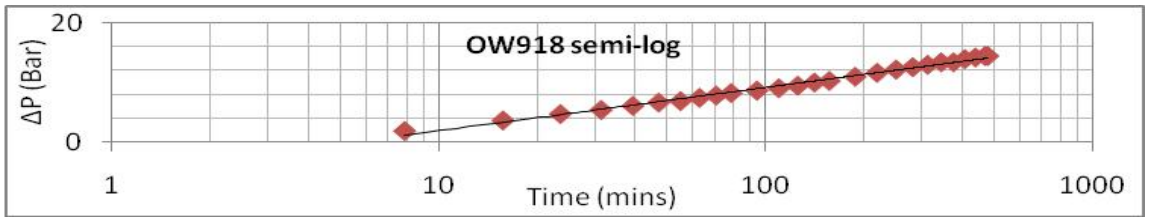

OW910 semi-log
ΔP (Bar)
20
0
-20
1
10
100
1000
Time (mins)

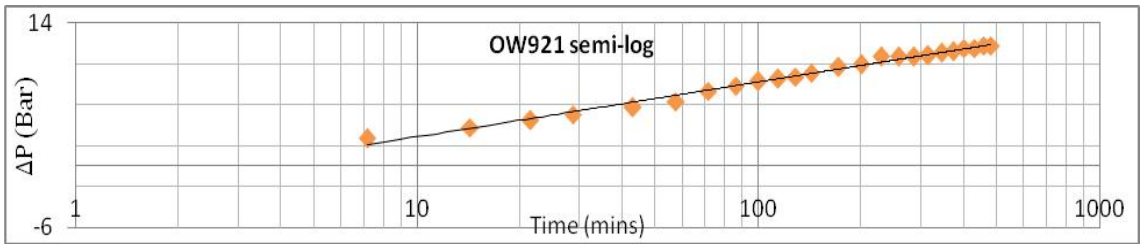

OW911A semi-log
ΔP (Bar)
30
20
10
0
1
10
Time (mins)
100
1000


OW912 semi-log
R² = 0.980
ΔP (Bar)
50
0
-50
1
10
100
1000
10000
100000
Time (mins)


OW914B semi-log
ΔP (Bar)
20
10
0
1
10
100
1000
Time (mins)


OW915 semi-log
ΔP (Bar)
20
0
1
10
100
1000
Time (mins)


OW918 semi-log
ΔP (Bar)
20
0
1
10
100
1000
Time (mins)


OW921 semi-log
ΔP (Bar)
14
-6
1
10
100
1000
Time (mins)

Appendix IV: Program Codes

Appendix IVa: Polynomial Regression Order 4

```
# Step 1: Importing necessary libraries
import numpy
import pandas as dfpd
import matplotlib.pyplot as plt
from sklearn.metrics import mean_absolute_error as mae
import math
#Step 2: Reading the dataset
df = dfpd.read_csv("wellsdata.csv")
print(df.shape)
temp_data = df['Temp']
rest_data = df['Resistivity']

# Step 3: Fit regression model
geo_model = numpy.poly1d(numpy.polyfit(rest_data, temp_data, 4))
Model_fit_line = numpy.linspace(0,630 , 297)

#Step 4. Predicting a new result
for resist in range(40):
    print("-------------------------------------------------------------------------")
    resistivity=float(input("Enter the Resistivity to Predict the Temparature : "))
    print("-------------------------------------------------------------------------")
    rest_model= geo_model(resistivity)
    print(resist+1," : ", rest model)
```

```
#Step 5: Visualize data
plt.plot(y_actual,y_predicted, 'bo', color='red')
plt.title('AddBoost Regression Predicted vs Actual temperature Scatter Plot \n')
plt.ylabel('Predicted Temp oC')
plt.xlabel('Actual temp oC')
plt.show()

#step 6: coefficient of correlation
y_actual = df['Actual Temp']
y_predicted = df['Predicted Temp']

np.corrcoef(y_actual, y_predicted)

#step 7: determination of MAE
y_actual = df['Actual Temp']
y_predicted = df['Predicted Temp']
mae(y_actual, y_predicted)
```

Appendix IVb: AdaBoost Regressor

```
# Step 1: Importing necessary libraries
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.model_selection import RandomizedSearchCV
from sklearn.ensemble import AdaBoostRegressor
from sklearn.tree import DecisionTreeRegressor
from sklearn.metrics import mean_absolute_error as mae
import math
#Step 2: Reading the dataset
dataset = pd.read_csv('wellsdata.csv')
X = dataset.iloc[:, 0].values
y = dataset.iloc[:, 1].values
X = np.array(X).reshape(-1,1)
y = np.array(y).reshape(-1,1).ravel()
# Step 3: Fit regression model
param_grid = {"learning_rate" : [0.01, 0.1, 0.3],
              "loss"          : ['linear', 'square', 'exponential']
             }
ada_reg = RandomizedSearchCV(AdaBoostRegressor( DecisionTreeRegressor(),
                                                n_estimators=500),
                             param_distributions = param_grid, n_iter = 10,
                             verbose = 2, n_jobs = -1)
```

```
ada_reg.fit(X, y)
#Step 4. Predicting a new result
for count in range(40):
    y_pred=float(input("Enter Resistivity \t"))
    y_pred = clf.predict([[y_pred ]])
    print("Predicted Temparature: % d\n \t"% y_pred)
    print('-----------------------------------------')
y_pred = regressor.predict(X_test)
y_pred = sc_y.inverse_transform(y_pred)
#Step 5: Visualize data
plt.plot(y_actual,y_predicted, 'bo', color='red')
plt.title('AddBoost Regression Predicted vs Actual temperature Scatter Plot \n')
plt.ylabel('Predicted Temp oC')
plt.xlabel('Actual temp oC')
plt.show()

#step 6: coefficient of correlation
y_actual = df['Actual Temp']
y_predicted = df['Predicted Temp']
np.corrcoef(y_actual, y_predicted)

#step 7: determination of MAE
y_actual = df['Actual Temp']
y_predicted = df['Predicted Temp']
mae(y_actual, y_predicted)
```

Appendix IVc: Random Forest Regression

```
# Step 1: Importing necessary libraries
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.model_selection import RandomizedSearchCV
from sklearn.ensemble import RandomForestRegressor
param_grid = {'max_features': ['auto', 'sqrt']}
from sklearn.metrics import mean_absolute_error as mae
import math
#Step 2: Reading the dataset
dataset = pd.read_csv('wellsdata.csv')
X = dataset.iloc[:, 0].values
y = dataset.iloc[:, 1].values
X = np.array(X).reshape(-1,1)
y = np.array(y).reshape(-1,1).ravel()
# Step 3: Fit regression model
rfor_reg = RandomizedSearchCV(RandomForestRegressor(), param_distributions = param_grid,
                              n_iter = 100, verbose = 2, n_jobs = -1)
rfor_reg.fit(X, y)
#Step 4. Predicting a new result
for count in range(40):
    y_pred=float(input("Enter Resistivity \t"))
    y_pred = clf.predict([[y_pred ]])
    print("Predicted Temparature: % d\n \t"% y_pred)
    print('-----------------------------------------')
y_pred = regressor.predict(X_test)
y_pred = sc_y.inverse_transform(y_pred)
#Step 5: Visualize data
plt.plot(y_actual,y_predicted, 'bo', color='red')
plt.title('RFR Regression Predicted vs Actual temperature Scatter Plot \n')
plt.ylabel('Predicted Temp oC')
plt.xlabel('Actual temp oC')
plt.show()

#step 6: coefficient of correlation
y_actual = df['Actual Temp']
y_predicted = df['Predicted Temp']
np.corrcoef(y_actual, y_predicted)

#step 7: determination of MAE
y_actual = df['Actual Temp']
y_predicted = df['Predicted Temp']
mae(y_actual, y_predicted)
```

Appendix IVd: Support Vector Regression

```
# Step 1: Importing necessary libraries
import numpy as np
from sklearn.svm import SVR
import matplotlib.pyplot as plt
import pandas as pd
from sklearn.metrics import mean_absolute_error as mae
import math

#Step 2: Reading the dataset
dataset = pd.read_csv('wellsdata.csv')
X = dataset.iloc[:, 0].values
y = dataset.iloc[:, 1].values
X = np.array(X).reshape(-1,1)
y = np.array(y).reshape(-1,1).ravel()
Fit regression model
# Step 3: Fit regression model
svr_rbf = SVR(kernel='rbf', C=1000, gama=0.1, epsilon=10)
#regressor = SVR(kernel = 'rbf')
svr_rbf.fit(X, y)
#Step 4 Predicting a new result
for count in range(40):
    y_pred=float(input("Enter Resistivity \t"))
    y_pred = svr_rbf.predict([[y_pred ]])
```

```
    print("Predicted Temparature: % d\n \t"% y_pred)
    print('-----------------------------------------')
y_pred = regressor.predict(X_testy_pred)
y_pred = sc_y.inverse_transform(y_pred)
#Step 5: Visualize data
plt.plot(y_actual,y_predicted, 'bo', color='red')
plt.title('SVR Regression Predicted vs Actual temperature Scatter Plot \n')
plt.ylabel('Predicted Temp oC')
plt.xlabel('Actual temp oC')
plt.show()
#step 6: coefficient of correlation
y_actual = df['Actual Temp']
y_predicted = df['Predicted Temp']
np.corrcoef(y_actual, y_predicted)
#step 7: determination of MAE
y_actual = df['Actual Temp']
y_predicted = df['Predicted Temp']
mae(y_actual, y_predicted)
```

Appendix IVe: Decision Tree Regression

```
# Step 1: Importing necessary libraries
from sklearn.tree import DecisionTreeRegressor
from sklearn.metrics import mean_absolute_error as mae
import math
import numpy as np
import matplotlib.pyplot as plt
import pandas as pd
#Step 2: Reading the dataset
dataset = pd.read_csv('wellsdata.csv')
print(dataset.columns)
print(" __________Dataset Information_____________\n" )
print(dataset.info())
print(" ++++++++++++++++++Dataset Description+++++++++++++++\n" )
print(dataset.describe())
X = dataset.iloc[:, 0].values
y = dataset.iloc[:, 1].values
X = np.array(X).reshape(-1,1)
y = np.array(y).reshape(-1,1)
# Step 3: Fit regression model
regressor = DecisionTreeRegressor(random_state = 0)
print(regressor.fit(X, y))
print('_____________________________________________________')
#Step 4 Predicting a new result
```

```
for count in range(40):
    y_pred=float(input("Enter Resistivity \t"))
    y_pred = regressor.predict([[y_pred ]])
    print("Predicted Temparature: % d\n \t"% y_pred)
    print('-----------------------------------------')
#Step 5: Visualize data
Model_fit_line = numpy.linspace(1,630 , 297)
plt.plot(Model_fit_line, geo_model(Model_fit_line))
plt.title('Geothermal Wells Temparature vs Resistivity Predictive Model \n')
plt.ylabel('Temparature (Degree Celsius)')
plt.xlabel('Resistivity (Ohm)')
plt.scatter(x, y)
plt.show()
#step 6: coefficient of correlation
y_actual = df['Actual Temp']
y_predicted = df['Predicted Temp']
np.corrcoef(y_actual, y_predicted)
#step 7: determination of MAE
y_actual = df['Actual Temp']
y_predicted = df['Predicted Temp']
mae(y_actual, y_predicted)
```

Appendix IVf: Image Stacking

```
#First Step was to install Python Imaging Library into anaconda framework
from PIL import Image
#Variables for resizing the original image for Preprocessing
new_width1=500
new_height1=500
im1 = Image.open('resistivity_at_200_masl.jpg')
im1= im1.resize((new_width1, new_height1), Image.ANTIALIAS)
#Stacking the images together and doing interpolation
im2 = Image.open('Temp_at_200_masl.jpg')
im2= im2.resize((new_width1, new_height1), Image.ANTIALIAS)

#Creates a new image by interpolating between two input images, using a constant alpha
stack1=Image.blend(im1,im2,0.5).show()
stack1=(Image.blend(im1,im2,0.5))
stack1.save("Method1_res_plus_temp200.png")
#opening the image for processing
im3 = Image.open('resistivity_500_masl.jpg')
im3= im3.resize((new_width1, new_height1), Image.ANTIALIAS)
im4 = Image.open('Temp_500_masl.jpg')
im4= im4.resize((new_width1, new_height1), Image.ANTIALIAS)
stack2=Image.blend(im3,im4,0.5).show()
stack2=(Image.blend(im3,im4,0.5))
stack2.save("Method1_res_plus_temp500.png")
```

Appendix V: Predicted Vs Actual Data

Appendix Va: OW901 Predicted Vs Actual Data

Depth	Temp (°C)	Resistivity	2-Degree	3-Degree	4-Degree	5-Degree	SVR	Random F	AdaBoost	DTR
100	91.3	500	49.93	50.65	72.96	9.09	91	54	91	91
200	107.5	320	94.42	95.32	87.11	116.07	107	160	177	153
300	125.5	200	140.04	139.94	130.85	119.68	84	90	96	73
400	146.8	157	159.5	159.12	154.92	139.12	146	157	157	146
500	199.7	125	175.04	174.55	174.39	162.35	166	165	125	122
600	214.1	74	201.69	201.25	205.41	206.97	214	191	195	214
700	218.4	50	215.02	214.74	218.75	224.88	221	233	241	207
800	230	30	226.53	226.43	228.59	233.99	233	240	232	254
900	232.2	25	229.46	229.42	230.82	235.03	263	247	227	248
1000	249	20	232.41	232.44	232.94	235.46	248	247	233	244
1100	266.2	15	235.39	235.49	234.95	235.21	248	243	218	234
1200	274.4	14	235.99	236.1	235.33	235.08	245	244	220	254
1300	280.6	12	237.19	237.32	236.09	234.71	213	227	227	220
1400	289.4	11	237.79	237.94	236.46	234.49	218	233	228	234
1500	293.6	12.5	236.89	237.02	235.9	234.82	218	232	227	293
1600	298.9	13	236.59	236.71	235.71	234.91	227	243	227	255
1700	307.8	15	235.39	235.49	234.95	235.21	248	243	218	234
1800	317.1	16	234.79	234.87	234.55	235.32	240	247	222	255
1900	326.3	19.3	232.83	232.87	233.23	235.47	250	247	234	326
2000	332.9	20	232.41	232.44	232.94	235.46	248	247	233	244
2100	338.6	22.9	230.7	230.69	231.72	235.29	244	249	224	338
2200	342.3	26	228.87	228.82	230.38	234.87	257	248	231	342

Appendix Vb: OW902B Predicted Vs Actual Data

Depth (m)	Temp (°C)	Resistivity	2 degree	3 Degrees	4 Degrees	5 Degrees	SVR	RFR	AdaBoost	DTR
100	31.1	630	24.066	32.78	11.37	62.48	31	37	86	31
200	82.8	400	69.426	72.38	78.01	96.85	65	52	85	53
300	85.4	320	94.206	95.32	87.11	116.07	107	160	177	153
400	91.6	219	132.126	132.02	121.39	115.46	91	77	91	91
500	113.8	160	157.7	157.73	153.14	137.3	113	135	154	141
600	166.7	125	174.07	174.55	174.39	162.35	166	129	122	122
700	207.1	100	186.307	187.31	189.82	184.02	142	108	135	133
800	217.5	70	201.589	203.46	207.72	210.29	217	187	184	195
900	221.1	50	212.139	214.74	218.75	224.88	221	206	216	207
1000	210.3	40	217.524	220.53	223.84	230.32	237	222	216	215
1100	211.1	38	218.609	221.7	224.82	231.21	211	236	216	238
1200	208.7	36	219.698	222.88	225.79	232.02	208	204	211	208
1300	212.6	34	220.789	224.065	226.74	232.76	226	254	216	214
1400	216.6	33	221.336	224.65	227.21	233.1	224	211	212	216
1500	217.9	32	221.883	225.25	227.67	233.42	223	198	217	198
1600	220.5	20	228.51	232.44	232.94	235.466	248	245	233	244
1700	225.6	22	227.398	231.23	232.1	235.37	236	241	231	242
1800	230.3	28	224.08	227.63	229.49	234.48	230	224	233	226
1900	232.2	30	222.98	226.43	228.59	233.99	233	253	237	254
2000	237.9	32	221.88	225.25	227.67	233.42	223	198	217	198
2100	237.1	40	217.52	220.53	223.84	230.32	237	222	216	215

Appendix Vc: OW903 Predicted Vs Actual Data

depth	temp	Resistivity	2 Degee	3 Degree	4 Degree	5 Degrees	SVR	RFR	AdaBoost	DTR
100	24.5	600	27.8	35.94	36.93	50.95	24	34	91	24
200	57.6	200	140.087	139.94	130.85	119.68	84	76	95	73
300	75	180	148.75	148.64	141.64	126.94	116	109	150	95
400	81.2	70	201.589	203.46	207.72	210.29	217	191	184	195
500	107.4	25	225.73	229.42	230.82	235.03	266	247	227	248
600	126.4	18	229.6	233.66	233.76	235.45	243	234	236	232
700	182.3	10	234.11	238.56	236.83	234.22	241	228	228	227
800	205.3	14	231.86	236.1	235.33	235.08	252	253	221	254
900	210.3	12	232.98	237.32	236.09	234.71	207	219	227	220
1000	211.5	10	234.11	238.56	236.83	234.22	241	228	228	227
1100	194.6	9	234.67	239.17	237.19	233.93	239	245	230	243
1200	194.8	5	236.94	241.65	238.59	232.43	226	213	227	212
1300	198.6	8	235.24	239.79	237.55	233.61	221	213	227	212
1400	204.3	9	234.67	239.17	237.19	233.93	226	245	230	243
1500	208.1	9.1	234.62	239.11	237.16	233.96	240	223	228	208
1600	220.8	9.3	234.5	238.99	237.08	234.02	242	236	228	220
1700	242.6	9.9	234.16	238.62	236.87	234.2	242	273	233	285
1800	266.6	10	234.11	238.56	236.83	234.22	241	228	228	227
1900	292.2	10	234.11	238.56	236.83	234.22	241	228	228	227
2000	311.6	10	234.11	238.56	236.83	234.91	241	228	228	227
2100	324.3	13	232.42	236.71	235.71	234.91	227	251	232	255
2200	333.8	16	230.74	234.87	234.55	235.32	238	256	222	255

Appendix Vd: OW905A Predicted Vs Actual Data

Depth (m)	Temp	Resistivity	2 Degee	3 Degee	4 Degee	5 Degrees	SVR	RFR	AddBoost	DTR
100	34.7	400	69.42	72.38	78.01	96.85	65	53	85	53
200	48.1	200	140.08	139.94	130.85	119.68	84	74	95	73
300	63.9	100	186.3	187.31	189.82	184.02	64	105	135	103
400	70.3	67	203.15	205.12	209.44	212.71	116	90	184	70
500	126.7	50	212.13	214.74	218.75	224.88	221	211	216	207
600	134.5	37	219.15	222.29	225.31	231.62	210	168	212	134
700	175	25	225.73	229.42	230.82	235.03	266	250	227	230
800	187.6	18	229.62	233.66	233.76	235.45	243	231	236	232
900	196.7	10	234.112	238.56	236.83	234.22	241	227	228	227
1000	208.8	8	235.24	239.79	237.55	233.61	221	213	227	212
1100	209.2	9	234.67	239.17	237.19	233.93	239	238	232	243
1200	201.6	10	234.112	238.56	236.83	234.22	241	227	228	227
1300	191.2	10.5	233.83	238.25	236.65	234.36	231	223	227	191
1400	185	11	233.54	237.94	236.46	234.49	219	234	228	234
1500	183.3	12	232.98	237.32	236.09	234.71	207	220	227	220
1600	182.9	12.6	232.64	236.96	235.87	234.84	214	211	227	182
1700	183.5	12	232.98	237.32	236.09	234.71	207	220	227	220
1800	183.7	11.8	233.09	237.45	236.17	234.67	207	188	227	183
1900	184.5	12	232.98	237.32	236.09	234.71	207	220	227	220
2000	184.8	13	232.42	236.71	235.71	234.91	227	256	236	216
2100	186.6	12.3	232.81	237.14	235.98	234.78	210	226	227	186
2200	188.2	12	232.98	237.32	236.09	234.71	207	220	227	220
2300	192.6	11.6	233.2	237.57	236.24	234.63	208	191	227	192
2400	196.3	11.9	233.04	237.39	236.13	234.69	206	194	227	196
2500	207.2	12	232.98	237.32	236.09	234.71	207	220	227	220

Appendix Ve: OW906A Predicted Vs Actual Data

Depth (m)	Resistivity	temp	2 Degree	3 Degree	4 Degree	5 Degrees	SVR	RFR	AdaBoost	DTR
100	200	84.4	140.08	139.94	130.85	119.68	84	72	98	73
200	125	95.1	174.07	174.55	174.39	162.35	166	124	122	122
300	79	98.1	196.93	198.52	202.48	202.7	98	116	184	98
400	57	99.4	208.41	210.74	215	220.24	259	149	239	99
500	40	132.7	217.52	220.53	223.84	230.32	237	211	216	215
600	28	164.8	224.08	227.63	229.49	234.48	230	226	233	226
700	20	192	228.51	232.44	232.94	235.46	248	246	233	244
800	15	210	231.3	235.49	234.95	235.21	252	231	218	234
900	11	215.7	233.54	237.94	236.46	234.49	219	232	228	234
1000	8	207.3	235.24	239.79	237.55	233.61	221	214	227	212
1100	7.8	207.3	235.35	239.92	237.62	233.54	219	213	227	212
1200	7.3	212.6	235.63	240.23	237.79	233.36	215	229	228	233
1300	8	233.1	235.24	239.79	237.55	233.61	221	214	227	212
1400	8.2	244.7	235.12	239.67	237.48	233.67	224	249	228	265
1500	9	265.4	234.67	239.17	237.19	233.93	239	246	230	243
1600	9	279.3	234.67	239.17	237.19	233.93	239	246	230	243
1700	9.4	293.8	234.45	238.93	237.05	234.05	243	279	233	302
1800	9.6	302.4	234.33	238.8	236.98	234.11	244	296	239	315
1900	9.9	315.3	234.16	238.62	236.87	234.2	242	270	234	285
2000	10	328.6	234.11	238.56	236.83	234.22	241	225	228	327
2100	12	336.2	232.98	237.32	236.09	234.71	207	219	227	320
2200	15	340.9	231.3	235.49	234.95	235.21	252	231	218	334
2300	18	344.7	229.62	233.66	233.76	235.45	243	231	236	332
2400	20	347.2	228.51	232.44	232.94	235.46	248	246	233	344
2500	25	348.4	225.73	229.42	230.82	235.03	266	245	227	348
2600	30	348.4	222.98	226.43	228.59	233.99	242	256	237	334
2700	40	345.6	217.52	220.53	223.84	230.32	237	211	216	315
2800	48	341.1	213.21	215.89	219.79	226.09	277	297	221	336
2900	59	336.2	207.35	209.61	213.91	218.81	322	333	244	343
3000	68	343.9	202.63	204.57	208.87	211.91	304	262	258	324

Appendix Vf: OW908 Predicted Vs Actual Data

Depth (m)	Temp (oC)	Resistivity	2 Degree	3 Degree	4 Degree	5 Degree	SVR	RFR	AdaBoost	DTR
100	28.3	450	56.29	60.59	76.36	59.5	56	45	86	42
200	36.5	225	129.6	129.58	118.59	114.62	36	70	116	36
300	59.6	65	204.19	206.24	210.57	214.29	31	57	87	59
400	86.1	39	218.06	221.12	224.33	230.77	213	130	213	86
500	100.4	23	226.84	230.63	231.68	235.28	238	202	215	199
600	138.4	15	231.3	235.49	234.95	235.21	252	234	219	234
700	133.2	10	234.11	238.56	236.83	234.22	241	227	228	227
800	154.8	7	235.8	240.41	237.9	233.25	216	182	227	183
900	192.8	6.7	235.97	240.6	238	233.13	218	205	227	192
1000	209.9	6	236.37	241.03	238.25	232.85	223	214	227	213
1100	223.2	5	236.94	241.65	238.59	232.43	226	211	227	212
1200	228.5	4	237.5	242.28	238.92	231.96	223	224	227	225
1300	186.1	5	236.94	241.65	238.59	232.43	226	211	227	212
1400	196.4	5.3	236.77	241.47	238.48	232.56	226	207	227	196
1500	207.8	5	236.94	241.65	238.59	232.43	226	211	227	212
1600	221.8	5.4	236.71	241.4	238.45	232.6	226	218	227	221
1700	233.3	5.5	236.65	241.34	238.42	232.64	226	218	227	233
1800	233.1	5.5	236.65	241.34	238.42	232.64	226	218	227	233
1900	204.3	6	236.37	241.03	238.25	232.85	223	214	227	213
2000	220.5	6.4	236.14	240.78	238.11	233.02	220	226	227	228
2100	224.3	6.6	236.03	240.66	238.04	233.09	218	219	227	224
2200	237.9	7.1	235.75	240.35	237.86	233.28	216	216	228	237
2300	247	8.4	235.01	239.54	237.41	233.74	228	241	228	247
2400	260.2	8.8	234.78	239.3	237.26	233.87	236	254	237	260
2500	287.2	9	234.67	239.17	237.19	233.93	239	246	233	243
2600	288.1	10.8	233.66	238.06	236.54	234.44	224	269	230	288
2700	297.1	14	231.86	236.1	235.33	235.08	252	255	222	254
2800	307.5	17	230.18	234.26	234.16	235.4	266	265	247	262
2900	316.4	21	227.95	231.84	232.52	235.43	295	293	234	316

Appendix Vg: OW910 Predicted Vs Actual Data

Depth(m)	Temp	Resistivity	2 Degree	3 Degree	4 Degree	5 Degree	SVR	random fo	AddBoost	DTR
100	29.07	400	69.426	72.38	78.01	96.85	65	50	85	53
200	77.59	160	157.7	157.73	153.14	137.3	113	145	163	141
300	88.14	80	196.4	197.98	201.89	201.84	123	112	184	88
400	91.4	40	217.52	220.53	223.84	230.32	237	212	217	215
500	118.83	20	228.51	232.44	232.94	235.46	248	241	232	224
600	193.44	10	234.11	238.56	236.83	234.22	241	228	228	227
700	198.81	7	235.8	240.41	237.9	233.25	216	182	227	183
800	206.29	6	236.37	241.03	238.25	232.85	223	213	227	213
900	227.58	5.6	236.6	241.28	238.38	232.69	225	229	228	227
1000	235.06	5	236.94	241.65	238.59	232.43	226	212	227	212
1100	223.55	4	237.5	242.28	238.92	231.96	223	225	227	225
1200	222.02	3	238.07	242.9	239.25	231.46	222	223	227	222
1300	223.36	4	237.5	242.28	238.92	231.96	223	225	227	225
1400	225.28	4.5	237.2	241.97	238.75	232.2	225	225	227	225
1500	227.77	5	236.94	241.65	238.59	232.43	226	212	227	212
1600	231.03	5.1	236.88	241.59	238.55	232.47	226	223	228	231
1700	234.68	6	236.37	241.03	238.25	232.85	223	213	227	213
1800	236.59	6.4	236.14	240.78	238.11	233.02	220	227	228	228
1900	236.4	6.9	235.86	240.47	237.93	233.21	216	215	228	236
2000	235.44	7	235.8	240.41	237.9	233.25	216	182	227	220
2100	236.21	8	235.24	239.79	237.55	233.61	221	211	228	212
2200	239.28	9	234.67	239.17	237.19	233.93	239	241	234	243
2300	241.39	10	234.11	238.56	236.83	234.22	241	228	228	227
2400	243.88	11	233.54	237.94	236.46	234.49	219	231	228	234
2500	246.76	12	232.98	237.32	236.09	234.71	207	219	221	220
2600	249.83	14	231.86	236.1	235.33	235.08	252	252	221	254
2700	254.24	17	230.18	234.26	234.16	235.4	235	262	248	262
2800	259.8	19	229.06	233.05	233.35	235.47	249	238	236	234
2900	274.95	22	227.39	231.23	232.1	235.37	236	243	226	242
3000	284.35	25	225.73	229.42	230.82	235.03	266	252	227	248

Appendix Vh: OW911A Predicted Vs Actual Data

Depth (m)	Temp. (oC)	resistivity	2 Degree	3 Degree	4 Degree	5 Degree	SVR	RFR	AddBoost	DTR
100	56.2	450	56.29	60.59	76.36	59.5	56	45	91	42
200	65.9	400	69.42	72.38	78.01	96.85	65	51	96	53
300	75.1	380	75.18	77.63	79.19	106.05	75	66	114	75
400	90	200	140.08	139.94	130.85	119.68	84	72	99	73
500	116.6	180	148.75	148.64	141.64	126.94	116	109	143	95
600	142.7	88	192.34	193.66	197.12	194.78	142	148	142	142
700	164.4	50	212.13	214.74	218.75	224.88	221	205	214	207
800	187.6	30	222.98	226.43	228.59	233.99	242	250	231	224
900	186.8	23	226.84	230.63	231.68	235.28	238	210	214	199
1000	193.5	19	229.06	233.05	233.35	235.47	249	234	236	234
1000	208.4	18	229.62	233.66	233.76	235.45	243	232	236	232
1100	210.8	18	229.62	233.66	233.76	235.45	243	232	236	232
1200	212.7	17	230.18	234.26	234.16	235.4	235	261	236	262
1300	216.5	16	230.74	234.87	234.55	235.32	238	252	222	255
1400	218.8	14	231.86	236.1	235.33	235.08	252	253	221	254
1500	222	16	230.74	234.87	234.55	235.32	238	252	222	255
1600	225.7	15	231.3	235.49	234.95	235.218	252	236	218	234
1700	227.7	13	232.4	236.71	235.71	234.91	227	256	218	255
1800	230	12	232.98	237.32	236.09	234.71	207	217	227	220
1900	233.4	11	233.54	237.94	236.46	234.49	219	234	228	234
2000	234.9	10	234.11	238.56	236.83	234.22	241	224	227	227
2100	238.5	11	233.54	237.94	236.46	234.49	219	234	228	234
2200	241.2	13	232.4	236.71	235.71	234.91	227	256	227	255
2300	245.4	15.5	231.02	235.18	234.75	235.27	245	245	218	245
2400	251.4	19	229.06	233.05	233.35	235.47	249	234	236	234
2500	266.2	25	225.73	229.42	230.82	235.03	266	247	227	248
2600	272.5	30	222.98	226.43	228.59	233.99	242	250	231	254
2700	273.3	42	216.44	219.37	222.85	229.36	273	282	218	285

Appendix Vi: OW912 Predicted Vs Actual Data

Depth	Temp	Resistivity	2 Degree	3 Degree	4 Degree	5 Degree	SVR	RFR	AdaBoost	DTR
100	75.1	150	162.29	162.41	159.1	143.62	181	151	154	146
200	87.6	63	205.24	207.36	211.69	215.83	126	126	204	87
300	92	15	231.3	235.49	234.95	235.21	252	229	218	234
400	114.1	12	232.98	237.32	236.09	234.71	207	219	227	220
500	120.6	10	234.11	238.56	236.83	234.22	241	225	227	227
600	146.9	7	235.8	240.41	237.9	233.25	216	185	227	183
700	180.2	8	235.24	239.79	237.55	233.61	221	211	226	212
800	212.3	8.5	234.95	239.48	237.37	233.77	230	227	227	212
900	247.2	9.7	234.28	238.74	236.94	234.14	243	266	240	247
1000	261	10	234.11	238.56	236.83	234.22	241	255	227	227
1100	273.1	12	232.98	237.32	236.09	234.71	207	219	227	220
1200	283.9	14	231.86	236.1	235.33	235.08	252	257	220	254
1300	291	15	231.3	235.49	234.95	235.21	252	229	218	234
1400	296.8	17	230.18	234.26	234.16	235.4	235	265	236	262
1500	302.5	18	229.62	233.66	233.76	235.45	243	238	236	232
1600	305.7	20	228.51	232.44	232.94	235.46	248	245	233	244
1700	307.6	20	228.51	232.44	232.94	235.46	248	245	233	244
1800	310.8	23	226.84	230.63	231.68	235.28	238	211	212	250
1900	311.5	25	225.73	229.42	230.82	235.03	266	245	227	248
2000	309	28	224.08	227.63	229.499	234.48	230	231	234	226
2100	305.7	34	220.78	224.06	226.74	232.76	226	249	216	254
2200	302.1	42	216.44	219.37	222.85	229.36	273	281	219	285
2300	301.9	51	211.6	214.16	218.22	224.25	267	292	224	295
2400	281	60	206.82	209.05	213.36	218.08	232	305	239	305
2500	287.6	70	201.58	203.46	207.72	210.29	217	205	184	300
2600	294.5	82	195.39	196.89	200.71	200.09	294	254	184	294
2700	307	95	188.8	189.94	192.88	188.51	306	234	207	307
2800	321.5	108	182.34	183.16	184.9	176.9	321	254	207	321
2900	332.5	121	175.99	176.55	176.86	165.67	332	244	219	332
3000	340.8	135	169.3	169.62	168.23	154.4	340	263	219	340

Appendix Vj: OW914B Predicted Vs Actual Data

Depth (m)	Temp (°C)	Resistivity	2 Degree	3 Degree	4 Degree	5 Degree	SVR	RFR	AdaBoost	DTR
100	30.2	125	174.07	174.55	174.39	162.35	166	124	121	122
200	35	90	191.32	192.59	195.91	193	35	96	142	35
300	46.3	55	209.47	211.88	216.08	221.63	235	144	162	142
400	104.8	44	215.36	218.2	221.84	228.33	197	151	214	104
500	196.6	35	220.24	223.47	226.26	232.4	216	216	213	196
600	223.2	22	227.39	231.23	232.1	235.37	236	245	233	242
700	235.2	16	230.74	234.87	234.55	235.32	238	256	222	255
800	238.8	16	230.74	234.87	234.55	235.32	238	256	222	255
900	243	18	229.62	233.66	233.76	235.45	243	226	236	232
1000	244.9	20	228.51	232.44	232.94	235.46	248	242	233	244
1100	244.7	20.7	228.12	232.02	232.65	235.44	245	263	237	244
1200	244.5	22	227.39	231.23	232.1	235.37	236	245	233	242
1300	242.2	27	224.63	228.23	229.94	234.69	242	273	232	242
1400	242.4	30	222.98	226.43	228.59	233.99	242	258	234	254
1500	240	32	221.88	225.25	227.67	233.42	223	203	219	198
1600	226.3	34	220.78	224.06	226.74	232.76	226	253	217	254
1700	225.1	38	218.6	221.7	224.82	231.21	211	236	214	238
1800	229.9	43	215.04	218.78	222.35	228.85	230	246	216	229
1900	223.8	47	213.74	216.46	220.31	226.67	255	248	223	223
2000	218.4	50	212.13	214.74	218.75	224.88	221	207	213	207
2100	208.3	62	205.77	207.92	212.25	216.59	208	197	208	208
2200	204.8	68	202.63	204.57	208.87	211.91	204	237	284	210
2300	205.5	81	195.91	197.43	201.3	200.96	205	186	184	205
2400	214.3	96	188.3	189.41	192.27	187.61	214	221	214	214
2500	219.2	118	177.45	178.06	178.72	168.21	219	187	207	219
2600	231.9	160	157.7	157.73	153.14	137.3	211	149	172	225
2700	242.4	198	140.94	140.79	131.89	120.27	216	178	184	242
2800	256.8	240	123.63	123.64	111.97	113.4	256	240	199	256
2900	265.1	277	109.44	109.85	98.26	114.12	264	253	219	265
3000	268.7	320	94.2	95.32	87.11	116.07	290	155	177	270

Appendix Vk: OW915 Predicted Vs Actual Data

Depth(m)	Temp	resistivity	2 Degree	3 Degree	4 Degree	5 Degree	SVR	random fc	AdaBoost	DTR
100	36.1	200	140.04	139.94	130.85	119.68	84	74	104	73
200	39.9	100	187.81	187.31	189.82	184.02	64	120	140	103
300	45.8	66	206.08	205.68	210	213.51	45	56	142	45
400	97.3	32	225.36	225.25	227.67	233.42	223	200	217	198
500	141.6	20	232.41	232.44	232.94	235.46	248	240	233	244
600	203	28	227.7	227.63	229.49	234.48	230	226	233	226
700	216.7	20	232.41	232.44	232.94	235.46	248	240	222	244
800	227.7	16	234.79	234.87	234.55	235.32	238	256	236	255
900	243.7	17	234.2	234.26	234.16	235.32	235	261	236	262
1000	254.7	20	232.41	232.44	232.94	235.46	248	240	233	244
1100	263.5	25	229.46	229.42	230.82	235.03	266	250	227	248
1200	269.1	30	226.53	226.43	228.59	233.99	242	255	235	254
1300	273.9	34	224.2	224.06	226.74	232.76	226	256	217	254
1400	278.3	38	221.88	221.7	224.82	231.21	211	230	216	238
1500	280.7	40	220.73	220.53	223.84	230.32	237	216	216	215
1600	282.1	42	219.58	219.37	222.85	229.36	273	282	219	285
1700	283.8	45	217.87	217.62	221.33	227.79	223	227	220	283
1800	284.3	46	217.3	217.04	220.82	227.24	284	272	223	284
1900	287.4	49	215.59	215.31	219.27	225.49	224	291	222	287
2000	293.5	50	215.02	214.74	218.75	224.88	221	205	214	277
2100	298.2	51	214.46	214.16	218.22	224.25	267	294	226	300
2200	295.1	51.7	214.06	213.76	217.85	223.81	295	295	220	295
2300	275.5	53	213.33	213.02	217.16	222.96	283	271	214	275
2400	244.3	54.2	212.65	212.33	216.52	222.17	244	206	191	244
2500	239.5	55	212.2	211.88	216.08	221.63	235	158	142	249
2600	244.8	56	211.64	211.31	215.54	220.94	241	228	200	244
2700	254.9	56.8	211.19	210.86	215.11	220.38	254	214	200	254
2800	278.1	57.4	210.86	210.52	214.78	219.96	269	222	229	278
2900	307.2	58.5	210.24	209.89	214.18	219.17	307	304	232	307
3000	329.1	60	209.4	209.05	213.36	218.08	328	305	239	305

Appendix VI: OW917 Predicted Vs Actual Data

Depth (m)	TEMP	Resistivity	2 Degree	3 Degree	4 Degree	5 Degree	SVR	RFR	AdaBoost	DTR
100	88.5	36	223.04	222.88	225.79	232.02	208	204	208	208
200	97.8	16	234.79	234.87	234.55	235.32	238	259	223	255
300	101.6	8	239.59	239.79	237.55	233.61	221	213	227	212
400	103.4	4	242.02	242.28	238.92	231.96	223	224	226	225
500	105	5	241.41	241.65	238.59	232.43	226	212	223	212
600	142.9	4	242.02	242.28	238.92	231.96	223	224	226	225
700	178.1	3	242.62	242.9	239.25	231.46	222	223	224	222
800	196.6	2	243.23	243.53	239.57	230.93	221	223	224	222
900	211.1	2.4	242.99	243.28	239.45	231.15	221	223	224	222
1000	215.2	4	242.02	242.28	238.92	231.96	223	224	226	225
1100	227.9	6	240.8	241.03	238.25	232.85	223	213	225	213
1200	237.3	9	238.99	239.17	237.19	233.93	239	243	230	243
1300	240.8	11	237.99	237.94	236.46	234.49	219	233	228	234
1400	243.6	15	235.39	235.49	234.95	235.21	252	232	217	234
1500	239.3	16	234.79	234.87	234.55	235.32	238	259	223	255
1600	231.5	18	233.6	233.66	233.76	235.45	243	227	236	232
1700	227.3	25	229.46	229.42	230.82	235.03	266	251	227	248
1800	224.6	32	225.36	225.25	227.67	233.42	223	198	216	198
1900	222	40	220.73	220.53	223.84	230.32	237	210	216	215
2000	220.4	50	215.02	214.74	218.75	224.88	221	209	212	207
2100	220.4	63	207.74	207.36	211.69	215.83	126	118	204	230
2200	221.6	87	194.67	194.2	197.72	195.67	210	128	142	247
2300	222.8	123	176.04	175.55	175.62	164	253	231	199	232
2400	224.2	158	159.03	158.66	154.33	138.51	130	147	154	241
2500	227.7	200	140.04	139.94	130.85	119.68	184	174	195	280
2600	229.7	260	115.64	116.04	104.09	113.32	181	231	177	265
2700	232.5	312	97.07	97.91	88.81	115.98	181	252	177	251

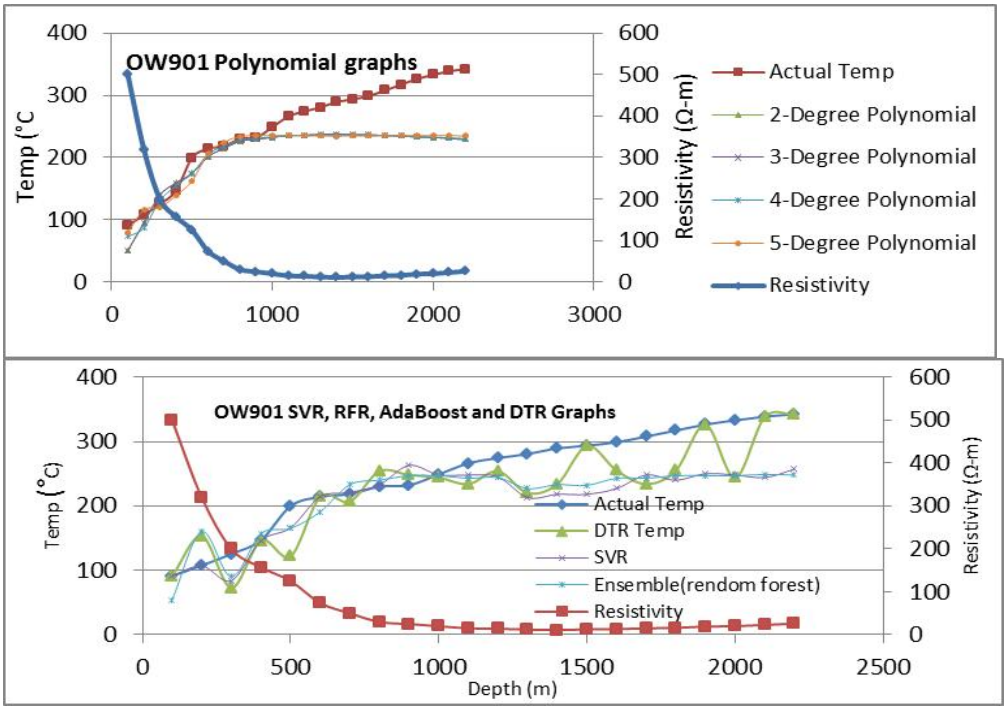
Appendix Vm: OW918 Predicted Vs Actual Data

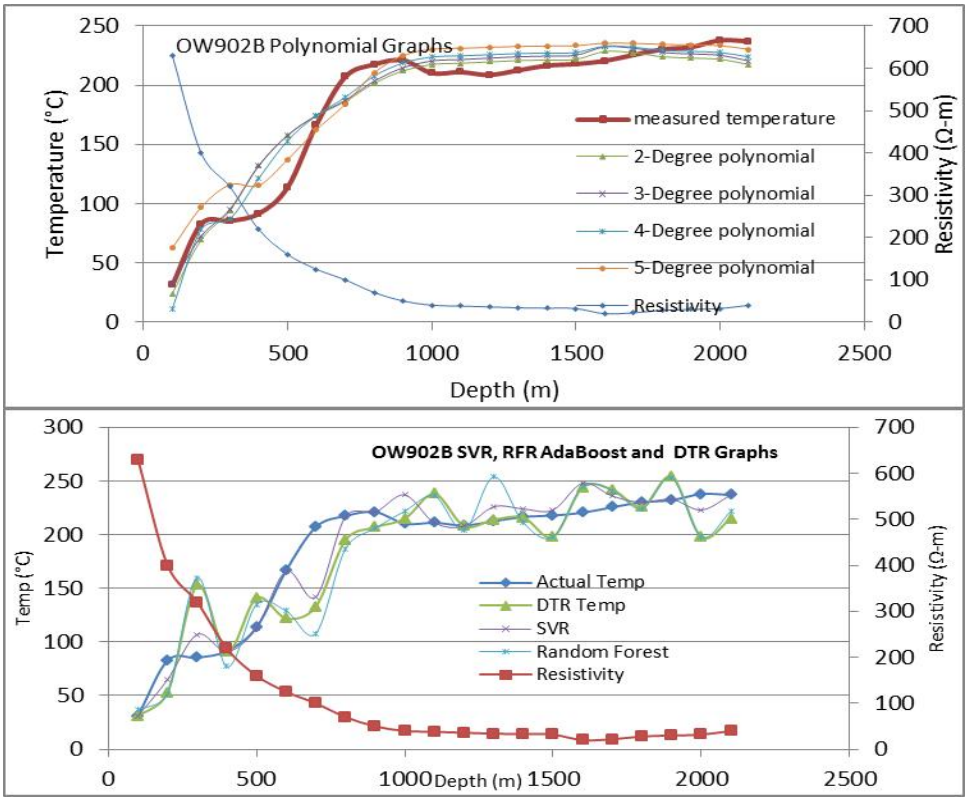
Depth	TEMP	Resistivity	2 Degree	3 Degree	4 Degree	5 Degree	SVR	RFR	AdAboost	DTR
100	64	90	193.08	192.59	195.91	193	45	93	142	35
200	83.7	50	215.02	214.74	218.75	224.88	228	210	214	207
300	89.2	20	232.41	232.44	232.94	235.46	252	246	233	244
400	90.5	18	233.6	233.66	233.76	235.45	233	239	236	232
500	91.3	16	234.79	234.87	234.55	235.32	242	257	224	255
600	149.8	14	235.99	236.1	235.33	235.08	254	251	223	254
700	160.7	12	237.19	237.32	236.09	234.71	212	223	227	220
800	172.1	10	238.39	238.56	236.83	234.22	244	224	228	227
900	181.7	13	236.59	236.71	235.71	234.91	231	261	232	255
1000	180.9	16	234.79	234.87	234.55	235.32	242	257	224	255
1100	199.8	20	232.41	234.87	232.94	235.46	252	244	233	244
1200	229.9	25	229.46	229.42	230.82	235.03	273	253	227	248
1300	256.6	29	227.11	227.03	229.04	234.25	226	231	230	226
1400	273.8	30.4	226.29	226.2	228.41	233.89	243	253	237	254
1500	280.2	32	225.36	225.25	227.67	233.42	227	202	219	278
1600	279.3	35	223.62	223.47	226.26	232.4	206	220	211	280
1700	279.9	42.1	219.53	219.31	222.8	229.31	270	283	217	285
1800	280.1	46	217.3	217.04	220.82	227.24	274	270	222	284
1900	279.7	51	214.46	214.16	218.22	224.25	258	292	223	300
2000	279.7	63	207.74	207.36	211.69	215.83	131	115	204	309
2100	279.1	100	187.81	187.31	189.82	184.02	173	104	129	320
2200	277	140	167.64	167.2	165.17	150.65	192	260	204	340
2300	273.2	180	148.89	148.64	141.64	126.94	106	110	125	300
2400	269	212.5	134.7	134.69	124.53	116.63	179	185	191	250
2500	265.8	250	119.48	119.79	107.89	113.17	180	190	199	256
2600	265.6	280	108.21	108.79	97.32	114.3	210	243	225	265

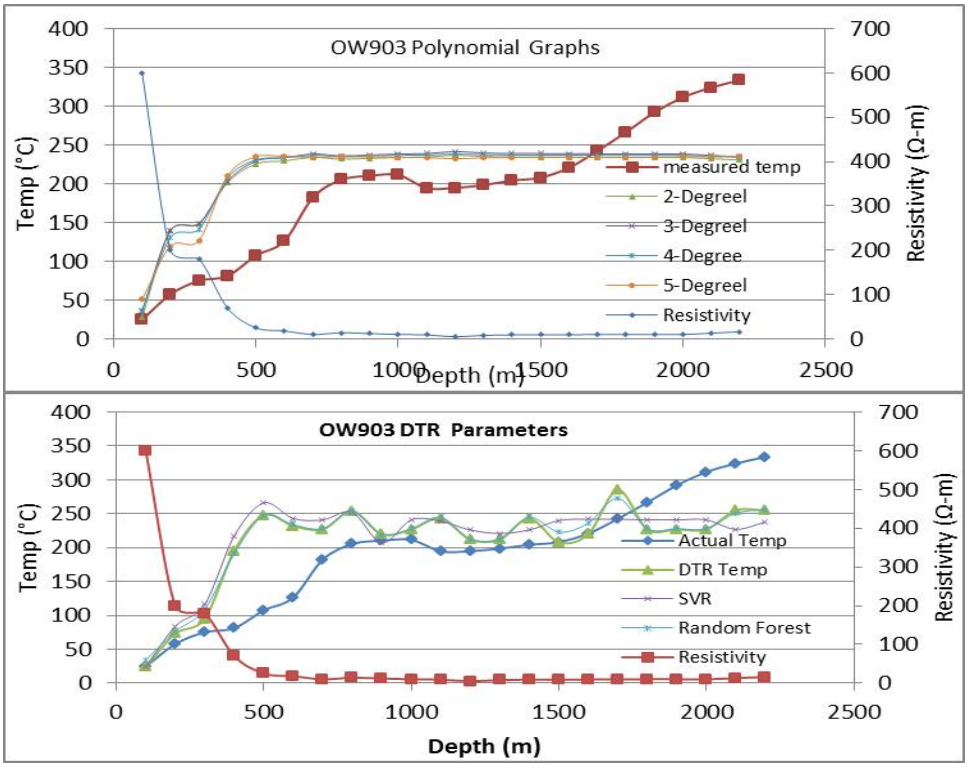
Appendix Vn: OW921 Predicted Vs Actual Data

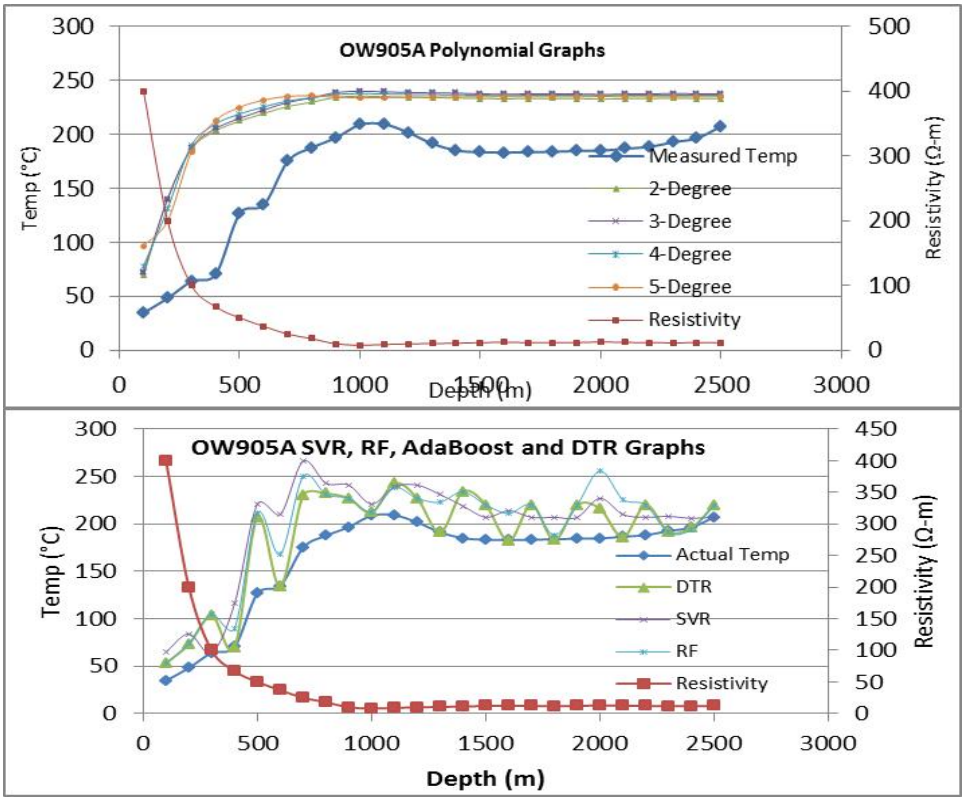
DEPTH	TEMP	Resistivity	2 Degree	3 Degree	4 Degree	5 Degree	SVR	RFR	AdaBoost	DTR
100	26.5	50	215.02	214.74	218.75	224.88	228	211	210	207
200	34.9	32	225.36	225.25	227.67	233.42	227	197	216	198
300	50.4	26	228.87	228.82	230.38	234.87	278	303	232	342
400	70.6	20	232.41	232.44	232.94	235.46	252	242	233	244
500	127.1	16	234.79	234.87	234.55	235.32	242	254	223	255
600	203.3	12	237.19	237.32	236.09	234.71	212	220	227	220
700	244.1	7	240.2	240.41	237.9	233.25	214	182	227	183
800	262.5	6	240.8	241.03	238.25	232.85	216	214	227	213
900	271.1	4	242.01	242.28	238.92	231.96	218	224	227	225
1000	277.5	3	242.62	242.9	239.25	231.46	219	223	227	222
1100	281.6	2.4	242.99	243.28	239.45	231.15	219	223	227	222
1200	288.5	2	243.23	243.53	239.57	230.93	219	223	226	222
1300	293.6	1	243.84	244.15	239.89	230.36	213	223	226	222
1400	297.7	1.3	243.66	243.96	239.8	230.53	215	223	226	222
1500	301.7	1.7	243.41	243.71	239.67	230.76	218	223	226	222
1600	304.1	2.1	243.17	243.46	239.54	230.98	219	223	226	222
1700	303.9	3	242.62	242.9	239.25	231.46	219	223	227	222
1800	300	3.6	242.26	242.53	239.05	231.77	218	223	227	225
1900	294.4	4.9	241.47	241.72	238.62	232.38	218	212	227	212
2000	291.8	6	240.8	241.03	238.25	232.85	216	214	227	213
2100	282.6	13	236.59	236.71	235.71	234.91	231	254	235	255
2200	268.4	29	227.11	227.03	229.04	234.25	226	227	230	226
2300	281.4	42	219.58	219.37	222.85	229.36	272	282	220	285
2400	294	58	210.52	210.18	214.46	219.53	300	290	239	307
2500	297.7	80	198.43	197.98	201.89	201.84	113	208	184	321
2600	299.4	135	170.09	169.62	168.23	154.4	330	256	207	340
2700	299.8	189	144.86	144.68	136.69	123.3	180	209	216	348
2800	301	231	127.04	127.18	115.87	113.17	176	203	198	317
2900	312.2	289	104.99	95.64	94.64	114.87	180	225	180	304

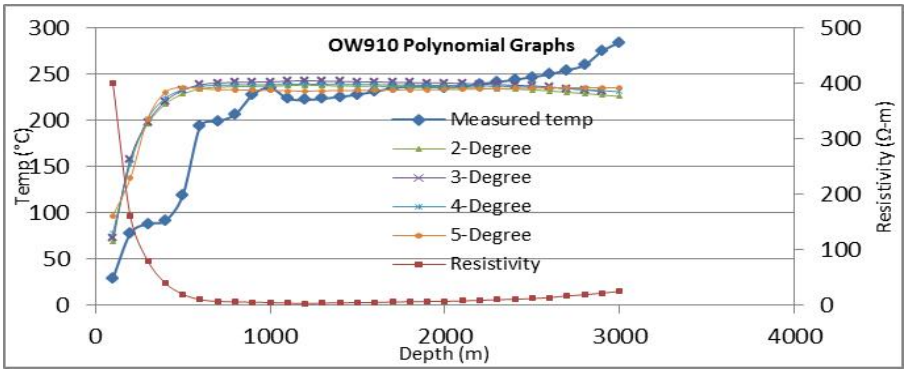
Appendix VI: Predicted Vs Actual Graphs

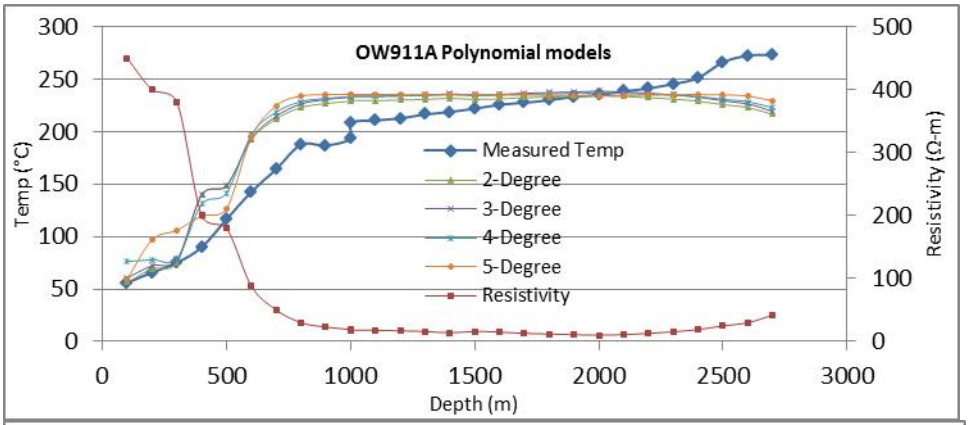


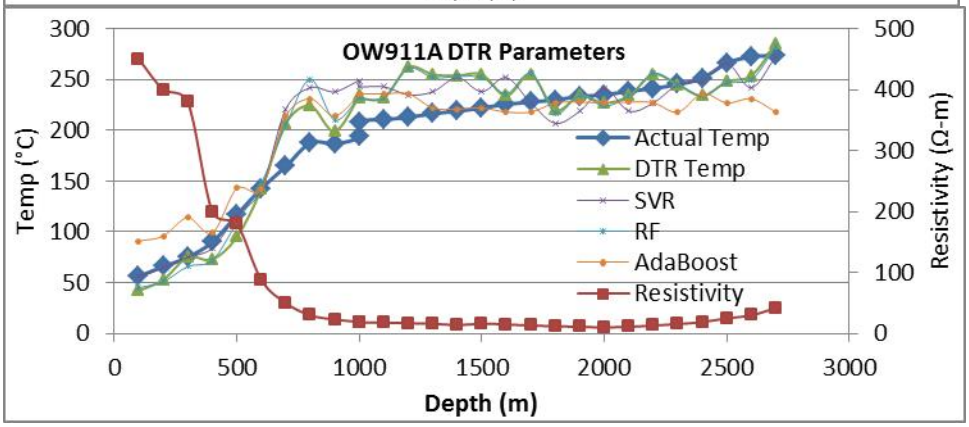

OW902B Polynomial Graphs
Temperature (°C)
Resistivity (Ω-m)
Depth (m)
measured temperature
2-Degree polynomial
3-Degree polynomial
4-Degree polynomial
5-Degree polynomial
Resistivity
OW902B SVR, RFR AdaBoost and DTR Graphs
Temp (°C)
Resistivity (Ω-m)
Depth (m)
Actual Temp
DTR Temp
SVR
Random Forest
Resistivity

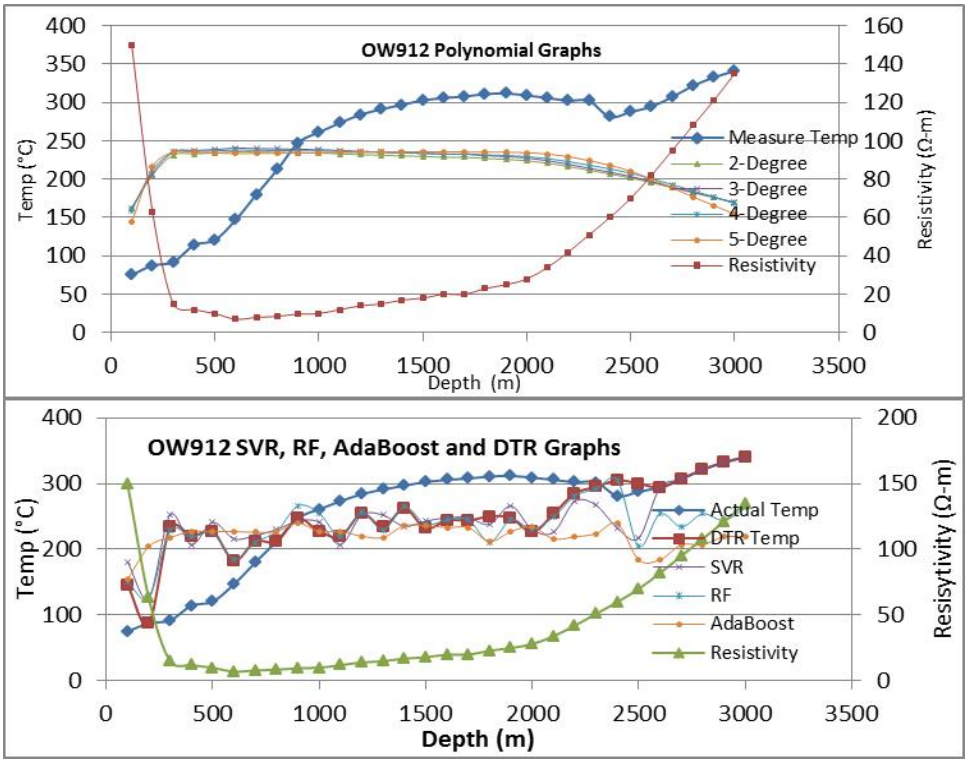

OW903 Polynomial Graphs
Temp (°C)
Resistivity (Ω-m)
Depth (m)
measured temp
2-Degreel
3-Degreel
4-Degree
5-Degreel
Resistivity
OW903 DTR Parameters
Temp (°C)
Resistivity (Ω-m)
Depth (m)
Actual Temp
DTR Temp
SVR
Random Forest
Resistivity

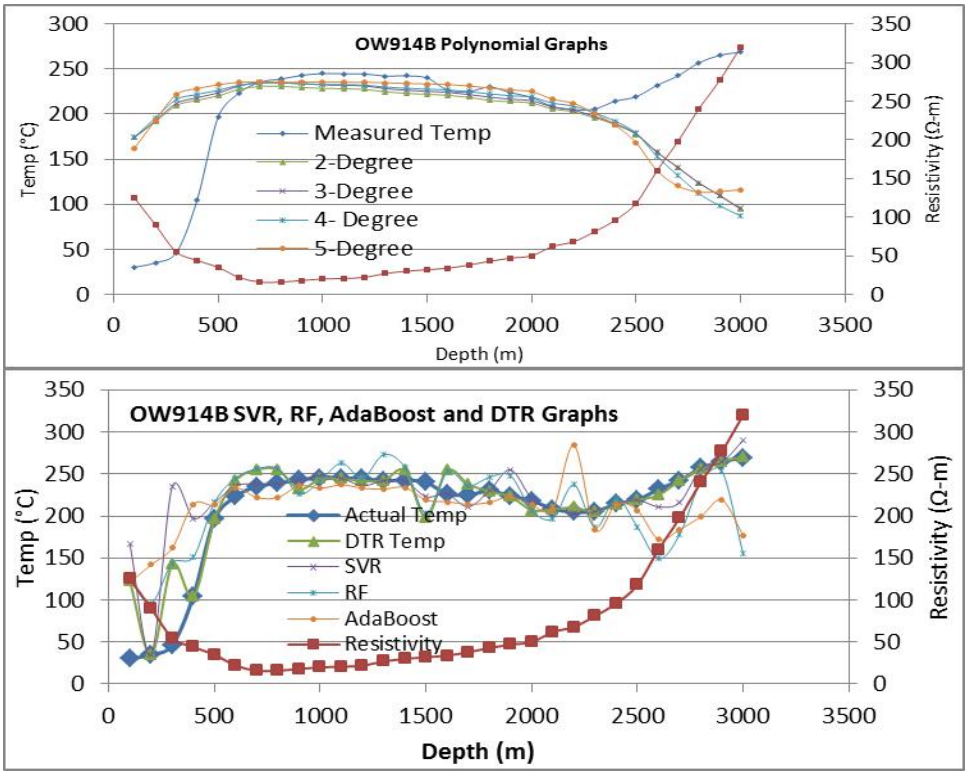

OW905A Polynomial Graphs
Temp (°C)
Resistivity (Ω-m)
Depth (m)
Measured Temp
2-Degree
3-Degree
4-Degree
5-Degree
Resistivity
OW905A SVR, RF, AdaBoost and DTR Graphs
Temp (°C)
Resistivity (Ω-m)
Depth (m)
Actual Temp
DTR
SVR
RF
Resistivity

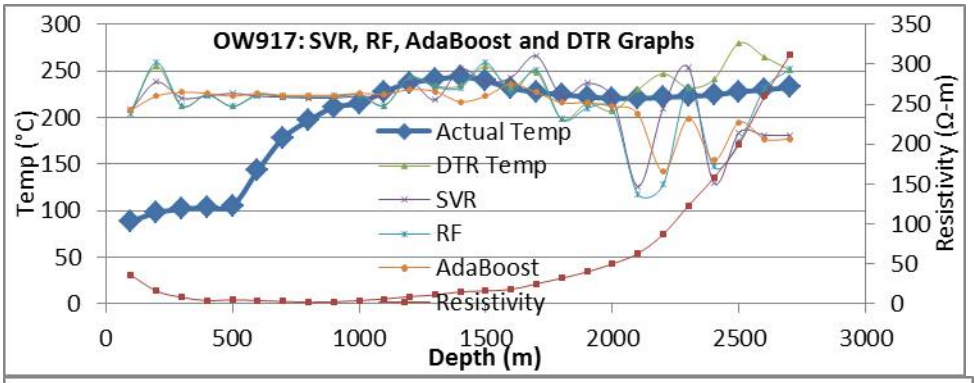

OW910 Polynomial Graphs
Temp (°C)
Resistivity (Ω-m)
Depth (m)
Measured temp
2-Degree
3-Degree
4-Degree
5-Degree
Resistivity

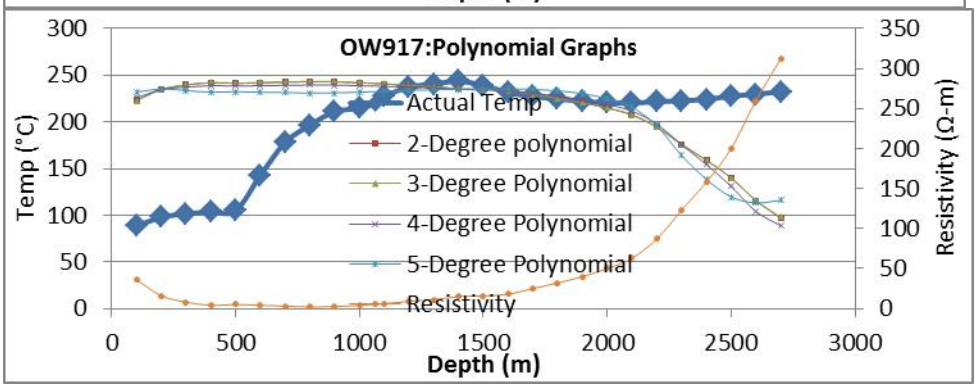

OW911A Polynomial models
Temp (°C)
Resistivity (Ω-m)
Depth (m)
Measured Temp
2-Degree
3-Degree
4-Degree
5-Degree
Resistivity

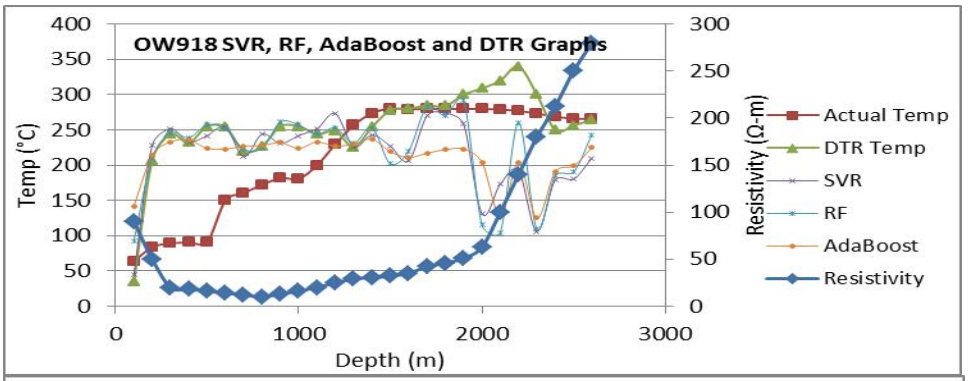

OW911A DTR Parameters
Temp (°C)
Resistivity (Ω-m)
Depth (m)
Actual Temp
DTR Temp
SVR
RF
AdaBoost
Resistivity

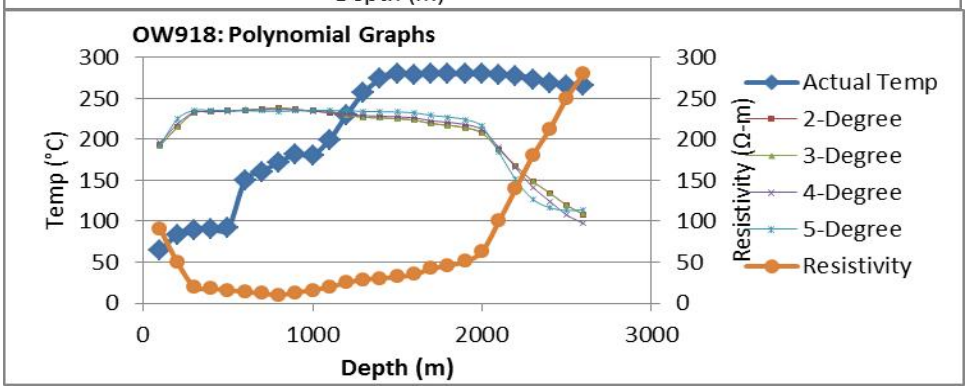

OW912 Polynomial Graphs
Temp (°C)
Resistivity (Ω-m)
Depth (m)
Measure Temp
2-Degree
3-Degree
4-Degree
5-Degree
Resistivity
OW912 SVR, RF, AdaBoost and DTR Graphs
Temp (°C)
Resisytivity (Ω-m)
Depth (m)
Actual Temp
DTR Temp
SVR
RF
AdaBoost
Resistivity

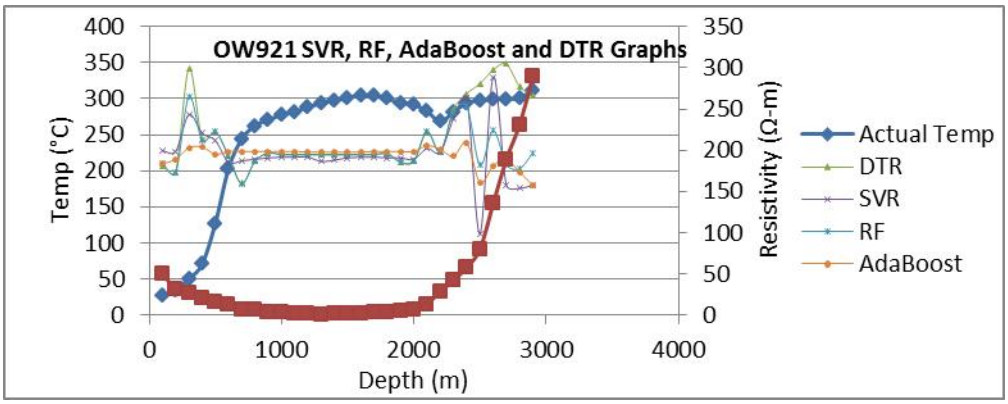

OW914B Polynomial Graphs
Temp (°C)
Resistivity (Ω-m)
Depth (m)
Measured Temp
2-Degree
3-Degree
4- Degree
5-Degree
OW914B SVR, RF, AdaBoost and DTR Graphs
Temp (°C)
Resistivity (Ω-m)
Depth (m)
Actual Temp
DTR Temp
SVR
RF
AdaBoost
Resistivity


OW917: SVR, RF, AdaBoost and DTR Graphs
Temp (°C)
Resistivity (Ω-m)
Depth (m)
Actual Temp
DTR Temp
SVR
RF
AdaBoost
Resistivity


OW917:Polynomial Graphs
Temp (°C)
Resistivity (Ω-m)
Depth (m)
Actual Temp
2-Degree polynomial
3-Degree Polynomial
4-Degree Polynomial
5-Degree Polynomial
Resistivity

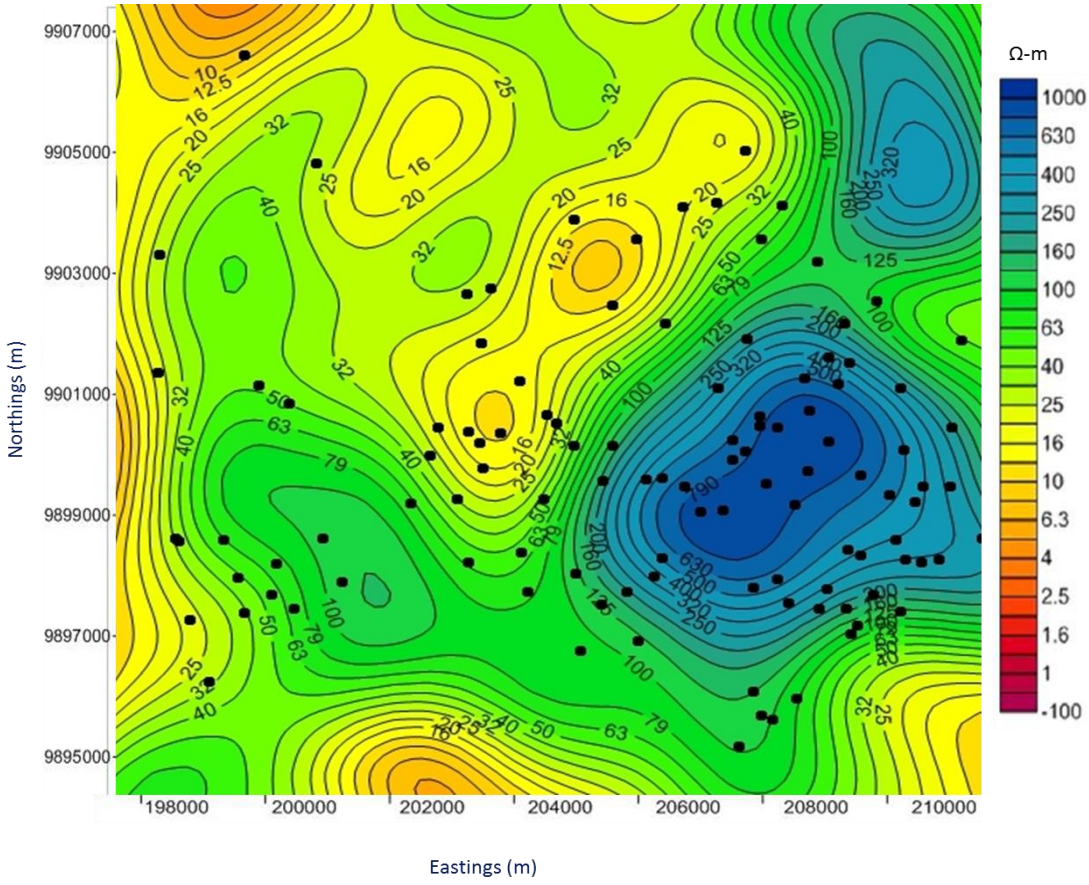

OW918 SVR, RF, AdaBoost and DTR Graphs
Temp (°C)
Resistivity (Ω-m)
Depth (m)
Actual Temp
DTR Temp
SVR
RF
AdaBoost
Resistivity


OW918: Polynomial Graphs
Temp (°C)
Resistivity (Ω-m)
Depth (m)
Actual Temp
2-Degree
3-Degree
4-Degree
5-Degree
Resistivity

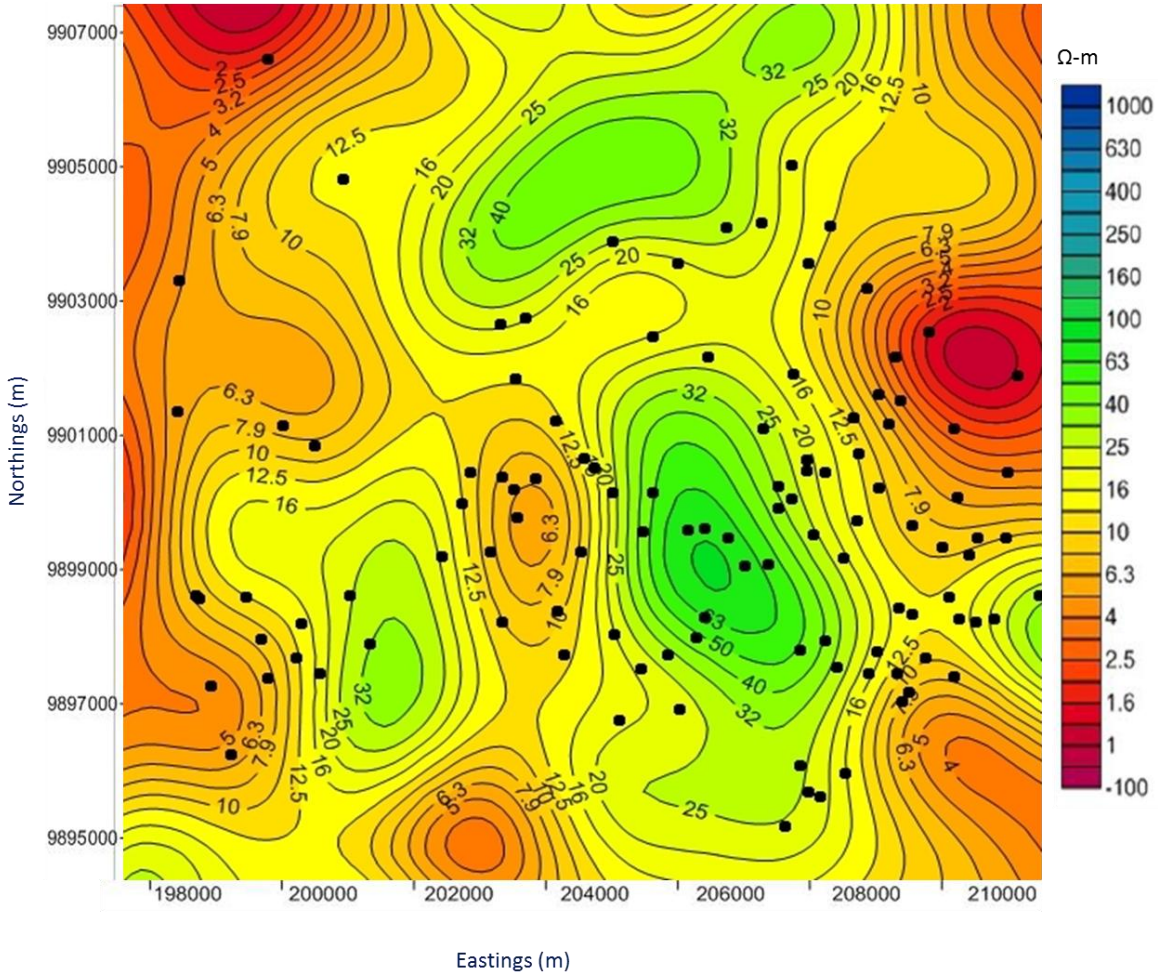

OW921 SVR, RF, AdaBoost and DTR Graphs
Temp (°C)
Resistivity (Ω-m)
Depth (m)
0
1000
2000
3000
4000
0
50
100
150
200
250
300
350
400
Actual Temp
DTR
SVR
RF
AdaBoost

Appendix VII: Resistivity Iso Maps

Appendix VIIa: Resistivity Iso Map at 1700 m a s l



Appendix VIIb: Resistivity Iso Map at 1500 m a s l



Appendix VIIc: Resistivity Iso Map at 1000 m a s l

Figure VIId: Resistivity Iso Map at 200 m b s l

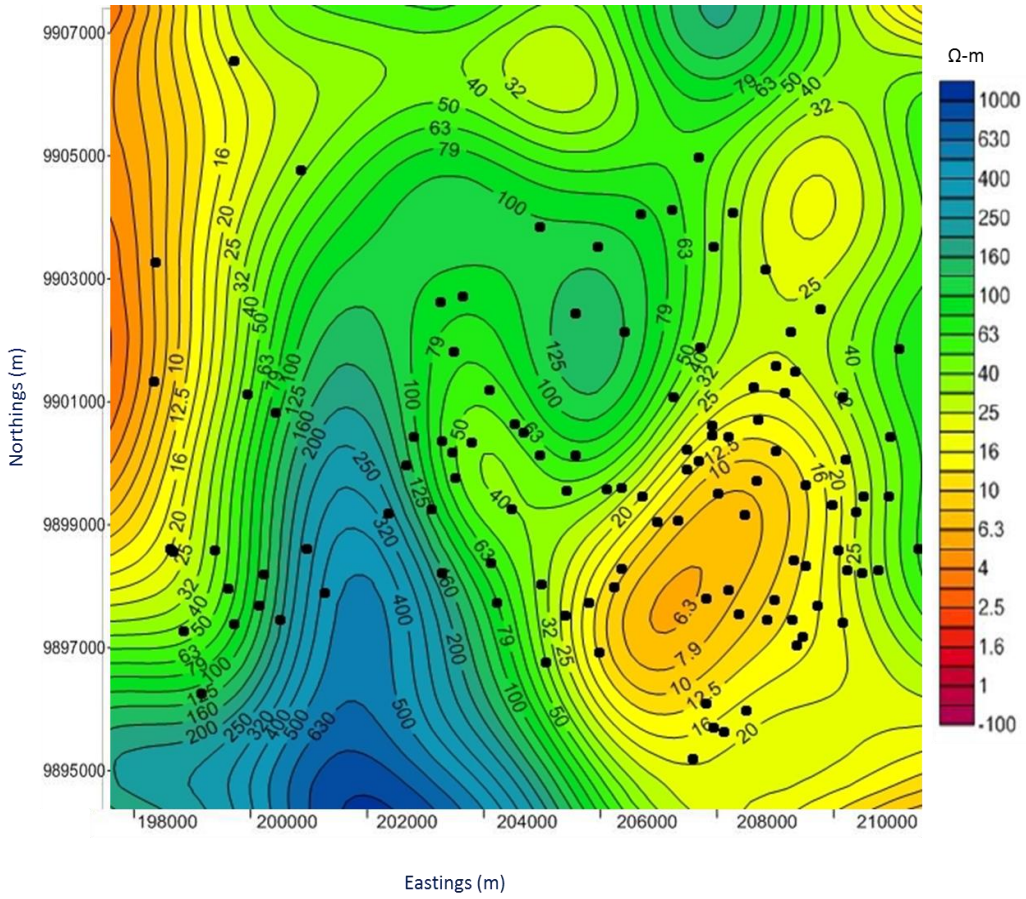


Figure VIIe: Resistivity Iso Map at 1000 m b s l