

TIME SERIES ANALYSIS OF VEGETATION DYNAMICS TO IDENTIFY POTENTIAL GEOHERMAL HOTSPOTS

ANNA ARORA

May, 2023

SUPERVISORS:

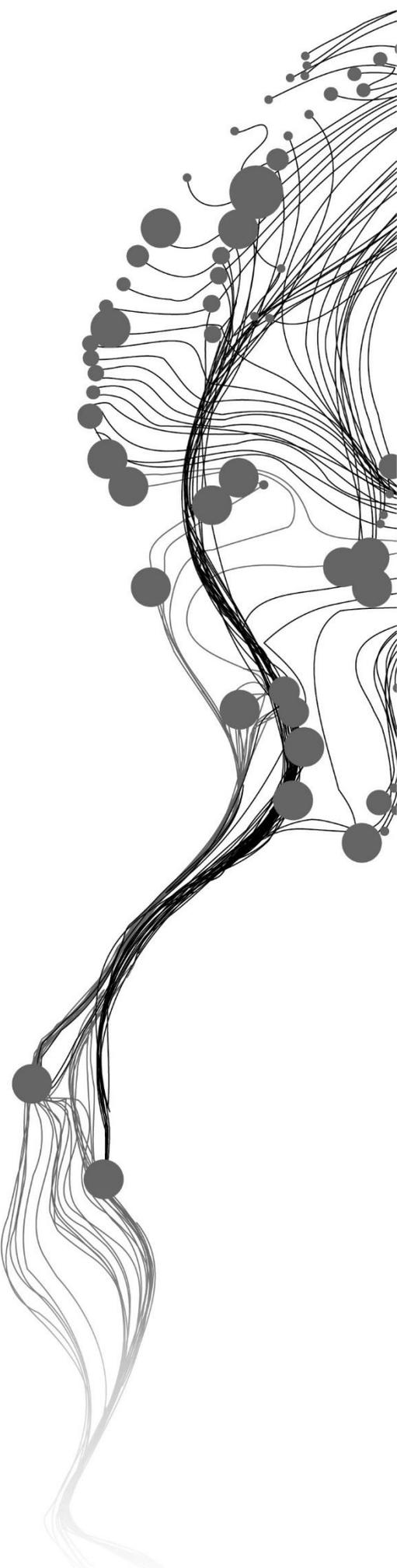
Dr.Ir. T.A. Groen (Thomas)

Dr. M. Belgiu (Mariana)

STUDY ADVISORS:

Dr. C.A. Hecker (Chris)

Dr. Rer. Nat. A. Soszynska (Agnieszka)



TIME SERIES ANALYSIS OF VEGETATION DYNAMICS TO IDENTIFY POTENTIAL GEOHERMAL HOTSPOTS

ANNA ARORA

Enschede, The Netherlands, May 2023

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.

Specialization: Geoinformatics

SUPERVISORS:

Dr.Ir. T.A. Groen (Thomas)

Dr. M. Belgiu (Mariana)

STUDY ADVISORS:

Dr. C.A. Hecker (Chris)

Dr. Rer. Nat. A. Soszynska (Agnieszka)

THESIS ASSESSMENT BOARD:

Dr. R. Darvishzadeh Varchehi (Roshanak) (Chair)

Dr. V.G. Jetten (Victor)(External Examiner, Faculty ITC (AES))

DISCLAIMER

This document describes work undertaken as part of a programme of study at the Faculty of Geo-Information Science and Earth Observation of the University of Twente. All views and opinions expressed therein remain the sole responsibility of the author, and do not necessarily represent those of the Faculty.

ABSTRACT

The reliance of the energy industry on fossil fuels makes it a prime contributor to greenhouse gas emissions, negatively impacting the goal of a sustainable world. The global energy industry is currently undergoing an upheaval as the efforts continue to materialize for its transition to renewable energy sources. Geothermal energy is one such sustainable source of energy that is being actively utilized for electricity production and as a heating source in many regions of the world, such as Iceland, Kenya, New Zealand, and some parts of the United States. These countries have highly active geothermal areas, which are usually characterized by surface manifestations such as fumaroles, hot springs, and steam emissions. While such expressions can help identify the active regions, the use of geothermal heat for electricity production requires points of a steeper subsurface geothermal gradient, often termed geothermal hotspots. Apart from geological and geophysical surveying and well exploration, many studies have employed geoinformation science and earth observation for preliminary exploration by examining the various surface characteristics such as lithology, mineralogy, soil characteristics, and topology. This study focused on studying the vegetation dynamics of the geothermally active regions and examining its use in detecting geothermal hotspots in the Olkaria region of Kenya. It was based on the hypothesis that after a rain event, the greenness of the vegetation improves in the region before returning to its original altered state; however, at geothermal hotspots, the wilting of the vegetation to its original state will be at a faster pace due to excessive ground heating. The temporal profiles of vegetation indices such as Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), and other bands, such as Red band, Near Infrared band and Shortwave Infrared band were assessed at 56 ground control points (geothermal and non-geothermal hotspots) around the Olkaria region for identifying characteristic patterns in the vegetation dynamics. The results did not show a characteristic response to verify the hypothesis. Additionally, the results were more depictive of soil dynamics due to scarce vegetation in Olkaria region. There were differences in responses of geothermal and non-geothermal hotspots but these could not be conclusively proven to be representative of the geothermal activity of the ground because of lack of enough cloud free ground control points around the rain events. Additionally these profiles were also quantitatively assessed using Dynamic Time Warping (DTW) by comparing the shapes, which showed that geothermal hotspots have a different profile because of their high dissimilarity with non-geothermal hotspots and low dissimilarity with geothermal hotspots. While these differences in the shapes could not be visually assessed, dynamic time warping showed potential in differentiating geothermal hotspots. Further, the NDVI range of geothermal hotspots over the entire time period was consistently lower than a certain limit or threshold. Combining the results of the NDVI value ranges and the dissimilarities in the shapes of the curves using DTW established the use of NDVI as one of the parameters for distinguishing geothermal hotspots from the background in a geothermally active region. This could be used as an additional condition for narrowing down the exploration areas of geothermal hotspots before a ground survey.

Keywords: Geothermal Energy, Dynamic Time Warping, NDVI, rain events, vegetation dynamics, energy

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to all those who have supported and guided me throughout the journey of completing the thesis. Without their unwavering encouragement, invaluable insights and belief in my abilities, this achievement would not have been possible.

First and foremost, I extend my deepest gratitude to my supervisors, Dr. Thomas Groen and Dr Mariana Belgiu for their guidance, expertise and continuous support in this journey. They always ensured that I receive advice and proper direction while working on the thesis along with constant motivation to excel as a critical researcher in this thesis. Their insightful feedback, constructive criticism, and willingness to engage in meaningful discussions have greatly enriched the quality of my work. Having a chat with Dr. Groen always pushed me to critically reflect on the results and work beyond my comfort zone. He ensured I was taking practical decisions throughout the whole academic year and not get overwhelmed in the process. I want to sincerely thank him for helping me gain more confidence in the way I present my work, supporting me in front of audiences, take better decisions and providing constructive feedback. Without his expertise, I would not have been able to work on vegetation dynamics, which was an entirely new field of study for me. Dr. Belgiu has been an exceptional mentor and confidante for me in this journey. Her genuine interest in my well-being and her ability to provide insightful advice and encouragement have been instrumental in keeping me motivated during challenging times. She also motivated me to work independently pushing me to explore different programming languages and frameworks, and develop a versatile skill set. I also want to thank my study advisors, Dr Chris Hecker and Dr. Agnieszka Soszynska for providing me with the information that I required for my thesis, and constantly helping with doubts, giving me a better clarity of the thesis topic.

I am indebted to my parents, and my sister, for their unconditional love, and belief in my abilities. They have always encouraged me to push my boundaries, and grow as an individual for dealing with the different challenges of personal and professional growth. I am grateful for their endless motivation and the countless opportunities they have provided me to pursue my dreams. Their unconditional love and acceptance have been a constant source of comfort during moments when I felt vulnerable or unstable.

I would also like to thank my friends here at ITC and back home in India for their companionship and willingness to lend an ear during my rants. I am fortunate to have amazing friends in my life and grateful for the laughter, encouragement and having a family away from home.

TABLE OF CONTENTS

1.	Introduction.....	7
2.	Objectives	11
2.1.	Objectives	11
2.2.	Hypothesis	11
2.3.	Research Questions.....	12
3.	Study area	13
4.	Methods.....	15
4.1.	Data Collection	15
4.2.	Data Processing	19
4.3.	Data Analysis.....	21
5.	Results.....	24
5.1.	Temporal profiles of geothermal and non-geothermal hotspots.....	24
5.2.	Temporal profile analysis using Dynamic Time Warping	31
6.	Discussions	35
6.1.	Applications of the results for potentially identifying geothermal hotspots.....	38
7.	Conclusions.....	42

LIST OF FIGURES

Figure 1: An illustration of the expected response difference between geothermal hotspots and non-geothermal hotspots around rain events	12
Figure 2: Study Area Map	13
Figure 3: Data Collection Workflow	15
Figure 4: Geoactivity of Ground Control Points	16
Figure 5: Vegetation groups associated with Ground Control Points	17
Figure 6: Monthly Rainfall in Olkaria	17
Figure 7: 5-day Rainfall Distribution in Olkaria	18
Figure 8: Data Processing Workflow	20
Figure 9: Mesh grid setup for downloading the NDVI time series data for Olkaria	21
Figure 10: Data Analysis Workflow	22
Figure 11: Illustration on how Dynamic Time Warping works (Lu et al., 2016)	23
Figure 12: Temporal profiles of some bands and indices at geothermal and non-geothermal hotspots between June 2020 – May 2022	24
Figure 13: Number of Cloud free data points per rain event after cloud masking and NDVI value filter (NDVI < 0.1)	26
Figure 14: Temporal profiles around few rain events for NDVI	26
Figure 15: Temporal profiles of NDVI around rain events which showed abrupt changes	27
Figure 16: Contaminated images with very high cloud cover	27
Figure 17: Examples of visual checks for extracting cloud free points	28
Figure 18: Temporal profiles around few rain events for Red band	28
Figure 19: Temporal profiles around few rain events for NIR band	29
Figure 20: Temporal profiles around rain events for SWIR band	30
Figure 21: Temporal profiles around few rain events for NDWI	30
Figure 22: Temporal profiles during periods of no rain for NDVI	31
Figure 23: Normalized distance differences using DTW for two years	32
Figure 24: Normalized Differences using DTW in different vegetation cover	33
Figure 25: Normalized distance distribution using DTW around few rain events	34
Figure 26: Vegetation cover in Olkaria between 2019-2022 (GeoHot Project, Faculty ITC)	35
Figure 27: Illustration for differences in expectations and observed results for the hypothesis	36
Figure 28: Confusion Matrix assessment of the ground control points	39
Figure 29: NDVI threshold map for defining exploration zones for geothermal hotspots in a geothermally active region	40
Figure 30: Classification accuracy of the ground control points using NDVI threshold	40

ABBREVIATIONS

- GIS – Geoinformation Science
- TIR – Thermal Infrared
- DTW – Dynamic Time Warping
- NDVI – Normalized Difference Vegetation Index
- NDWI – Normalized Difference Water Index
- R – Red band
- G – Green band
- B – Blue band
- EVI – Enhanced Vegetation Index
- SWIR – Short Wave InfraRed band
- NIR- Near InfraRed band

1. INTRODUCTION

The increasing demands on natural resources that the growing world population has resulted, among others, in an increase in greenhouse gas emissions and carbon footprint (Koide et al., 2021). As a consequence, the world is facing challenges of rising global temperatures and increased uncertainty in the behavior of natural processes (Chen et al., 2022). This contribution to global greenhouse gas emissions is subject to various unsustainable ways of producing and consuming goods in different sectors, such as the agricultural, utilities, manufacturing, transportation, and industrial sectors. However, these sectors' functionality relies on energy use, consequentially making the energy sector the largest contributor to greenhouse gas emissions in addition to its direct share (Wang & Zhou, 2018). The energy sector mainly relies on non-renewable sources that are expected to grow to meet the demands of the human population, causing significant environmental pollution (Gustavsson et al., 2017). Hence, in order to combat climate change and reduce greenhouse gas emissions, transition to renewable and sustainable sources of energy, such as wind, solar radiation, geothermal energy, hydropower, and biomass, is necessary to move closer to the goal of a carbon-neutral world (Güney, 2019).

Geothermal heat, the thermal energy present in the Earth's subsurface, is a reliable and sustainable energy source independent of seasonal variations (Soltani et al., 2021). The upwelling heat flux from the center of the Earth provides a continuous transfer of heat to the surface that can be utilized as an energy source in regions with a favorable geological setting for extraction (Song et al., 2018). It has direct applications for use as a heat source, such as in space heating applications using ground-sourced heat pumps for residential and industrial areas. Further, the surface manifestations of the incoming geothermal energy in the form of hot springs function as a recreational activity, often also used for bathing, swimming, and pond heating. Among these and other uses of geothermal energy for heating, it is also elemental in electricity generation (Lund & Toth, 2021). Although geothermal energy currently has a low contribution to the global energy demand, this is expected to grow in the future with technological advancements and cost reductions in the drilling process and extraction of geothermal heat (Soltani et al., 2021). It is a valuable energy source, sufficing as a feasible, economical, and high-potential solution for increasing the global share of renewable energy (van der Zwaan & Dalla Longa, 2019).

Ground-sourced heat pumps to utilize geothermal heat can mostly be used anywhere in the world; however, using geothermal heat for electricity production and large-scale heating requires setting up production plants in specific regions where geothermal hotspots are present. This makes its extraction location-specific requiring a favorable geological setting (Gkousis et al., 2022). These hotspots are locations of excess internal heat emitted from the crust and are usually indicated through the presence of surface manifestations such as fumaroles, steam emissions, or hot springs. Fumaroles are vents in the Earth's crust that emit meteoric water vapor and volcanic gases such as sulphur dioxide. Similarly, hot springs are locations wherein a spring is created due to the surfacing of the heated groundwater due to geothermal energy. Such indicators suggest high tectonic activity or the presence of active plate boundaries and act as a primary indicator for mapping out areas of localized geothermal reservoirs (Majumdar & Devi, 2021). In addition to the surface temperature anomalies, geothermal hotspots can be characterized by distinct surface characteristics, which can be studied through the presence of alteration minerals (Abubakar et al., 2018), the geology of the area, including faults, fractures, distinct drainage network, elevation data (Fahil et al., 2020) and characteristic type or density of vegetation around the hotspots (Elmarsdóttir et al., 2015). Accordingly, detecting these geothermal hotspots also involves reconnaissance

surveys, studying well log data acquired during the drilling of exploration wells to understand subsurface temperature profiles, and a thorough understanding of the geological properties such as the rock-mineral alterations, earth surface deformations and subsurface models of potential locations for extraction of geothermal energy (van der Meer et al., 2014).

The subsurface and geological indications of geothermal activity, especially the identification of thermal anomalies, have been intensively explored to detect potential geothermal hotspots with Earth Observation data and Geoinformation Science capabilities (Leibrand et al., 2019). For example, Shortwave Infrared (SWIR) and Thermal Infrared (TIR) imagery of the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) have been jointly used for geological mapping predicting the mineral alterations in rocks, and understanding the surface temperatures at a medium resolution (van der Meer et al., 2014). The previous study also explored the use of Shuttle Radar Topography Mission (SRTM) data to collect elevation information for mapping lineaments and fault lines that act as indicative parameters for high geothermal activity. In addition, the combination of other bands, such as NIR, Red, and Green bands with SWIR, showed potential for improving the mapping of the hydrothermal alteration zones for detecting geothermal hotspots while also considering the vegetation and lithology of the area (Abubakar et al., 2018). Light Detection and Ranging (LiDAR) data have also been applied to study hydrothermal rock alterations through the laser return intensity (LRI) values (Freski et al., 2021). Using LiDAR data for high-resolution elevation data and high-resolution thermal imagery through airborne sensors also shows the scope of further improvement in remote sensing applications (Hecker et al., 2017). Apart from the use of earth observation data, some studies have also used the potential of GIS for identifying suitable locations for geothermal exploration, including planning the logistics for setting up power plants (Mwaura & Kada, 2017).

These studies show how the different domains of remote sensing have been explored in identifying geothermal activity using subsurface and geological properties of the rocks. Additionally, in conjunction with the subsurface and geological data used as a guide for identifying hotspots, few studies have also inspected specific surface characteristics, such as the presence of sparse vegetation, land discoloration, and surface roughness during the exploratory stages. Geothermal activity can extensively alter the behaviour of various surface and ecological phenomena; hence, analyzing these expressions can provide useful information on the subsurface activity. These locations are sometimes highlighted with low vegetation density and the presence of specific heat-tolerant vegetation species, often termed geothermal grass (Saepuloh et al., 2021). Furthermore, a typical phenomenon observed around regions with active geothermal activity is stressed vegetation due to high ground temperatures (Mary et al., 2017). The study of such surface characteristics provides a quick and economical way of delineating areas of potential geothermal hotspots, which, combined with further detailed field analysis, are utilized for optimizing locations of production wells for extracting geothermal heat for indirect use of electricity production and direct heat use (Abuzied et al., 2020).

Most of these studies using remote sensing are restricted to static surface expressions, with limited to no analysis of the dynamics of such features, with the exception of detecting thermal anomalies subject to long-term geothermal activity. It can be argued that higher ground temperatures and the depreciation of groundwater as steam due to fumarolic activity can lead to variations in vegetation growth. This is further influenced by different weather conditions affecting the vegetation health around regions of high geothermal activity. Precipitation and ground temperatures particularly affect how the vegetation changes influencing its responses as observed through earth observation data, specifically through the vegetation indices such as NDVI (Zhu et al., 2015). The temporal dynamics of these biophysical features, such as vegetation, temperature, precipitation, and more around geothermally active regions, can further facilitate

the preliminary detection of potential locations of geothermal hotspots along with other exploration methods.

This particular research targets the study of the dynamics of the surface characteristics related to vegetation responses in a geothermally active region for detecting potential geothermal hotspots using remote sensing. It is based on the theory that the vegetation's behavior at geothermal hotspots differs from the non-geothermal hotspots due to vegetation deterioration around geothermal hotspots caused by ground heating. The higher ground temperatures at geothermal hotspots alter the condition of the vegetation as it loses moisture faster, causing changes in the growth of the vegetation, its greenness, or how it is wilting. Hence, it is expected that the vegetation will show different responses near geothermal hotspots as opposed to the other regions. To identify these differences between geothermal hotspots and non-geothermal hotspots for use as a potential indicator for identifying geothermal hotspots, a hypothesis is proposed in relation to the differences in the behaviour of vegetation, particularly around rain events. The rain events here act as an external environmental factor affecting the vegetation, around which the responses are analysed to identify any differences between geothermal hotspots and non-geothermal hotspots. The hypothesis states that the short-term change in vegetation greenness soon after a rain event is expected to increase for both geothermal hotspots and non-geothermal hotspots; however, as the immediate effects of rain wear out and it returns to its original state, the response curve of decrease in photosynthetic activity of the vegetation will be steeper in points with geothermal activity (geothermal hotspots) due to the ground heating.

This time series analysis of the vegetation dynamics will be done after the rain event for points of geothermal hotspots compared to points of no geothermal activity in the Olkaria region of Kenya. These points of interest are identified through ground data collected in 2022 as part of the GeoHot project of Faculty ITC, which studied the subsurface temperature profiles of 56 locations spread across the study area to indicate the presence of geothermal activity. The temporal dynamics of vegetation is analysed using the different indices, NDVI and EVI, which are widely used as indicators to understand the vegetation health capacity using remote sensing. NDVI is a vegetation index calculated using the Red and NIR band of the optical imagery, which represents the vegetation's greenness, with some indication about its health (Tomasella et al., 2018). Similar to NDVI, EVI provides another measure for measuring vegetation greenness, using the NIR, Red, and Blue bands, with better corrections for the background and the atmospheric effects (Fraga et al., 2014). In addition, the NDWI index was also used to study the temporal variations of the vegetation, specifically examining any differences observed in the vegetation water content between geothermal hotspots and non-geothermal hotspots (Xofis et al., 2022). While the indices provided some coherent information about vegetation dynamics, the bands used for calculating these indices were also studied to independently interpret the non-vegetation characteristics of the surface and relate to the responses observed using specific indices.

Furthermore, to understand the differences in these temporal dynamics using the indices and bands, the amount of rain is a crucial parameter, as too little rainfall may not show any immediate substantial differences in the photosynthetic activity of the vegetation. On the other hand, excessive rain can create bias with soaked land and vegetation without showing definite responses of the vegetation itself. Hence, it is essential that moderate rainfall is considered to study the temporal dynamics to avoid distortions in the vegetation responses analysed for the 56 ground control points. Based on the amount of rainfall received in Olkaria using the information from a weather monitoring station, rain events greater than 10 mm were considered optimal for the analysis, which is further explained in the Methods section.

In addition to visualizing the time series graphs of the vegetation indices and the bands for significantly identifying the differences in the vegetation responses of the geothermal hotspots compared to non-geothermal hotspots, a quantitative assessment of the shape of the temporal profiles is relevant for the study since the hypothesis draws a conclusion based on the differences in the steepness of the curve of geothermal hotspots after rain events. Hence, to examine the shape of the curves, Dynamic Time Warping is chosen as opposed to other methods for characterizing time series, such as Fourier analysis which studies patterns or trend analysis. Dynamic Time Warping is an algorithm that compares the dissimilarity between different temporal sequences creating a distance matrix (Bemdt & Clifford, 1994). Standard distance matrices such as the Euclidean distance matrix compares one point in the reference curve to its corresponding single point in the test curve ignoring time and local space shifts. However, Dynamic Time warping considers this when comparing two temporal sequences. It generates a distance matrix by matching one singular point in the reference curve to many points in the test curve and vice versa, focusing on the shape rather than just the absolute values of the curve. A normalized distance value is calculated depending on how different the responses are, indicating more the value, more dissimilarities in the shape of the curves. This will help get a generalized overview of how the temporal profiles of geothermal hotspots and non-geothermal hotspots compare in terms of their shape and deduce characteristic responses in dissimilarities to use as a parameter for identifying potential geothermal hotspots. The method will be applied to different pairs of geothermal and non-geothermal activity response curves of changes in vegetation profiles. According to the hypothesis considered, the calculated dissimilarity between different pairs will depict higher values when curves of geothermal activity and no geothermal activity are matched as opposed to little to no variations between the same geothermal activity curves.

The research aims to directly impact Sustainable Development Goal 7, Affordable and Clean Energy, influencing the target of increasing the share of sustainable energy sources and improving energy efficiency. It relies on enhancing the exploration of geothermal hotspots, reducing the dependency on field analysis, and maximizing the use of Geoinformation Science and Earth Observation in such exploration activities.

2. OBJECTIVES

The research aims to study the differences in vegetation dynamics between geothermal hotspots and non-geothermal hotspots in geothermally active regions that are detectable through remote sensing, potentially identifying an additional parameter for preliminary detecting geothermal hotspots. Through this study, temporal analysis of vegetation changes will be explored to identify characteristic response curves for geothermal hotspots. The changes are particularly studied around rain events to understand if there are any differences in the responses after an external weather influence.

2.1. Objectives

Analyze the presence of characteristic responses in vegetation change by comparing geothermal vs. non-geothermal locations through time series analysis.

Sub Objectives:

- Identify significant rain events to study temporal profiles of vegetation change.
- Generate time series profiles of different indices and bands for geothermal and non-geothermal hotspots as points of interest in different time periods, such as around rain events.
- Test for differences in the temporal profiles of the time series between geothermal hotspots and non-geothermal hotspots using visual assessment and Dynamic Time Warping.
- Assess the differences observed, if any, under different conditions, such as different rain events, and vegetation cover.

2.2. Hypothesis

The changes in vegetation in geothermal hotspots differ from the non-geothermal hotspots due to vegetation health deterioration around geothermal hotspots caused by ground heating. It is expected that the wilting of the vegetation is faster at geothermal hotspots, as the conditions become drier, compared to non-geothermal hotspot. This is analysed by studying the responses in vegetation specifically around rain events. These events cause changes in vegetation by modifying the environmental conditions, wherein we expect the vegetation to become greener after the rain event but the additional ground heat at geothermal hotspots cause faster browning of the vegetation as shown in Figure 1 using an illustration. Hence. The decline in the vegetation greenness after rain event would be steeper at geothermal hotspots and analysing this after-effect of a rain event may help determine any distinguishing responses between geothermal and non-geothermal hotspots.

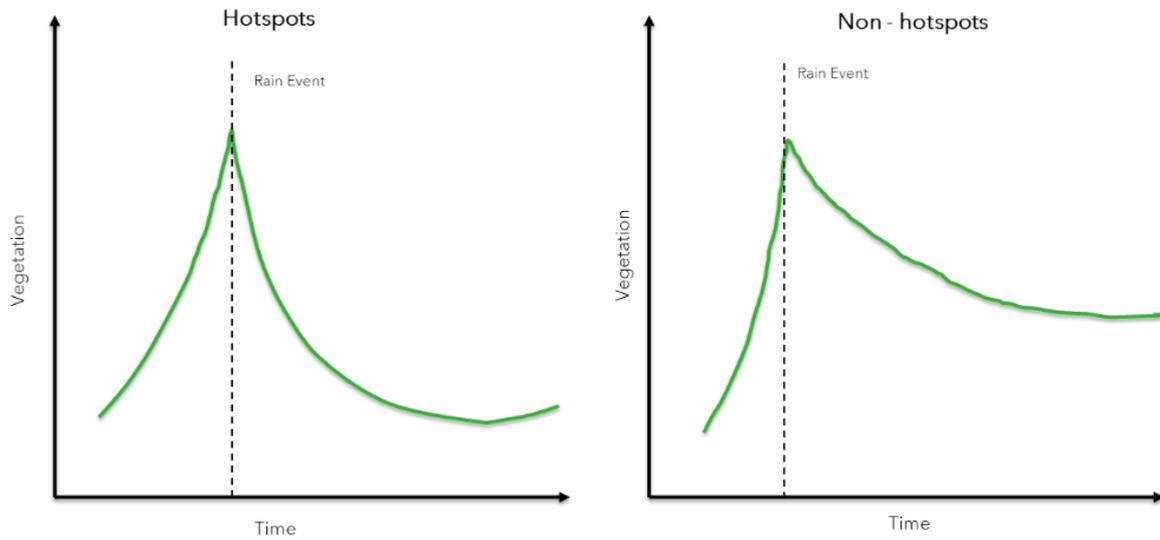


Figure 1: An illustration of the expected response difference between geothermal hotspots and non-geothermal hotspots around rain events

2.3. Research Questions

1. Do the vegetation dynamics show characteristic responses near geothermal hotspots specifically faster loss of greenness at hotspots as shown in Figure 1?
2. What remote sensing based parameters (individual bands and vegetation indices) related to vegetation dynamics show significant changes in time series profiles between geothermal and non-geothermal hotspots using Dynamic Time Warping?
3. Can the vegetation dynamics be used as an additional parameter to narrow down locations of potential geothermal hotspots?

3. STUDY AREA

This study is carried out in the Greater Olkaria Volcanic Complex in Kenya, North-West of Nairobi (Figure 2). It is a high-temperature geothermal field located in the centre of the Kenya rift valley, widely used for electricity production in Kenya, and characterized by numerous fumaroles (Omenda, 1998). The study area map (Figure 2) shows the distribution of ground control points for which temperature profiles were studied to identify potential locations of geothermal hotspots. This study will use these points to explore the temporal profiles of vegetation dynamics and make comparisons between hotspots and non-hotspots to see if there are significant differences in their dynamics around rainfall events.

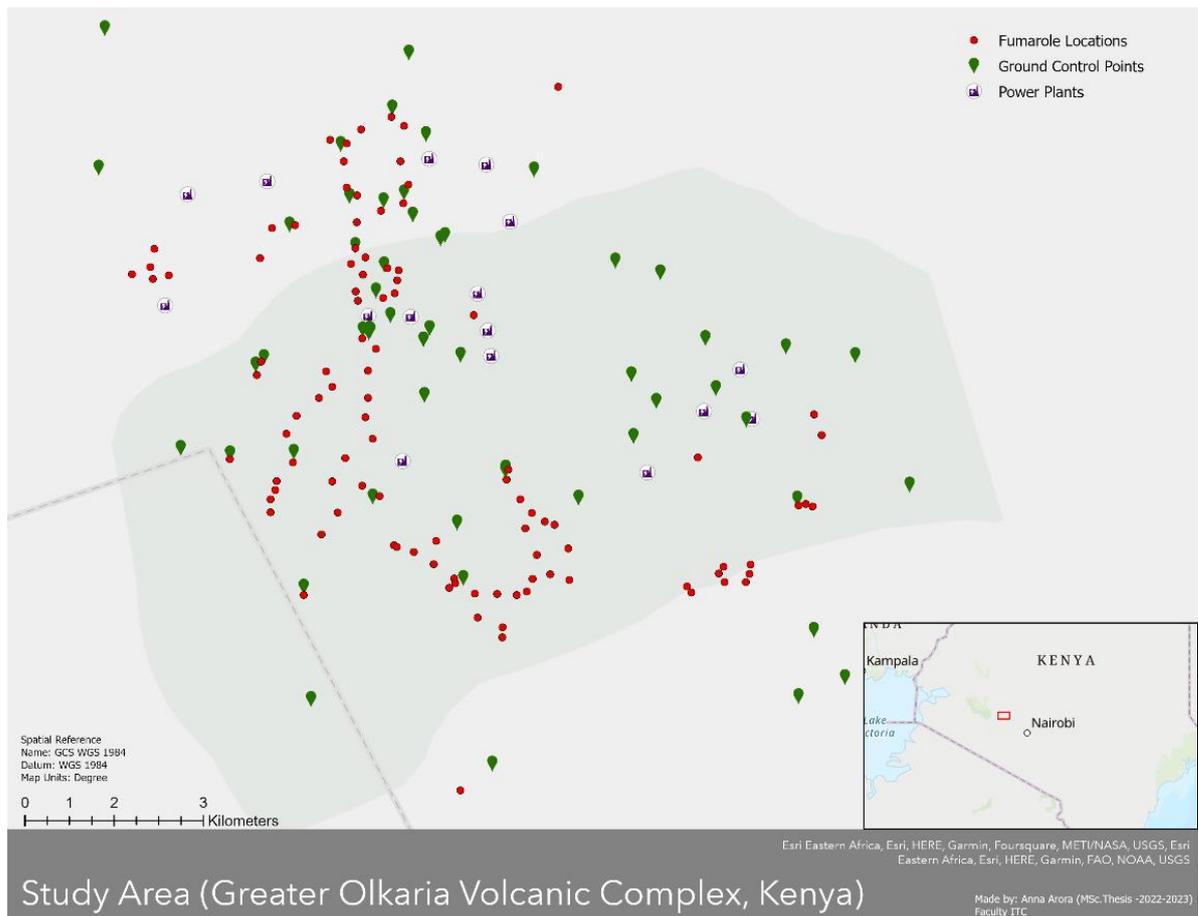


Figure 2: Study Area Map

The ground control points were collected as part of the GeoHot project of Faculty ITC, with its primary focus on geothermal hotspot detection using Ecostress data. This data is collected as part of the ECOSTRESS NASA mission launched in 2018 that captures thermal data from the international space station at a 70 x 70m resolution (NASA/JPL-Caltech, n.d.). The thermal imagery data enables the identification of temperature anomalies, using which a detection map in the geothermally active region of Olkaria was created. The Land Surface Temperature (LST) anomalies were detected through a moving window technique that distinguished points with anomalous median values. This was combined with the

presence of fumarole locations to create a detection map. Terrain evidence, such as sparse vegetation and rough, loose soil, was also used to further locate points for testing the presence of geothermal activity.

There were a total of 56 ground control points which were chosen for ground truth survey. Out of these points 24 locations showed positive geothermal activity, and the rest showed no significant geothermal activity. This was determined on the basis of the temperature gradients, where geothermal hotspots showed a recording of higher temperature gradient observed at 20cm depth. The result was also augmented with the presence of fumarole which already resulted in very high ground temperatures, that stayed such at 20cm depth as well in case of presence of a geothermal hotspot.

In lieu of this available data, these points were particularly tested to analyse the differences in the vegetation behaviour at geothermal hotspots and non-geothermal hotspots. For each of the points, the type of vegetation, and the soil conditions were also recorded, that helped in differentiating the points on the basis of the vegetation cover.

In this study, an additional parameter of vegetation dynamics is analysed to understand its suitability in identifying such locations with positive geothermal activity using Remote sensing.

4. METHODS

The response of the ground control points pertaining to the vegetation dynamics are analysed under three different time periods to test the hypothesis. The first time period covered a more extended time period of 2 years to understand prolonged differences between points of different geothermal activity (geoactivity). Following that, the temporal profiles around rain events are studied. Lastly, these dynamics are analysed during time periods with no rain (drought season). Consequently, the findings from these response functions are employed to classify the entire Olkaria region into hotspots and background points, validate the results and examine the accuracy of the proposed methods. The methodology used for this is divided into three sections: Data Collection, Data Processing, and Data Analysis, which are elaborated in the following chapters.

4.1. Data Collection

An overview of the data used for the study is shown in Figure 3. There are three sources of data that were used to test the hypothesis. These include:

1. **Ground control points** with information related to geothermal activity and other surface characteristics are available as part of the GeoHot project of Faculty ITC.
2. **Daily Rainfall data** from a weather monitoring station in Olkaria.
3. **Sentinel-2 Optical Images** accessed through Google Earth Engine for time series profiles.

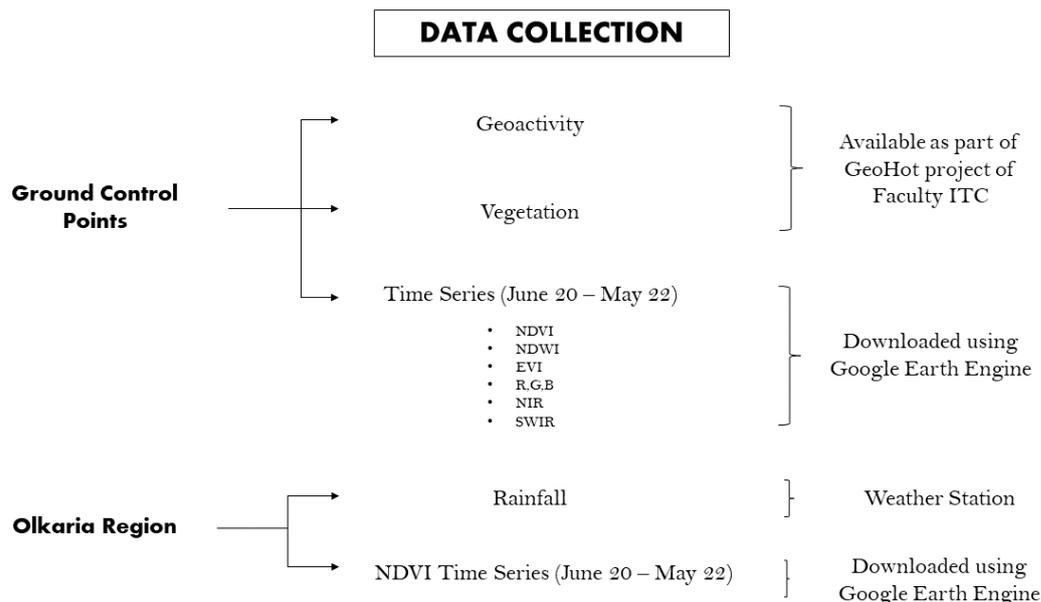


Figure 3: Data Collection Workflow

NDVI – Normalized Difference Vegetation Index, NDWI – Normalized Difference Water Index, R – Red band, G – Green band, B – Blue band, EVI – Evaporized Vegetation Index, SWIR – Short Wave InfraRed band, NIR- Near InfraRed band

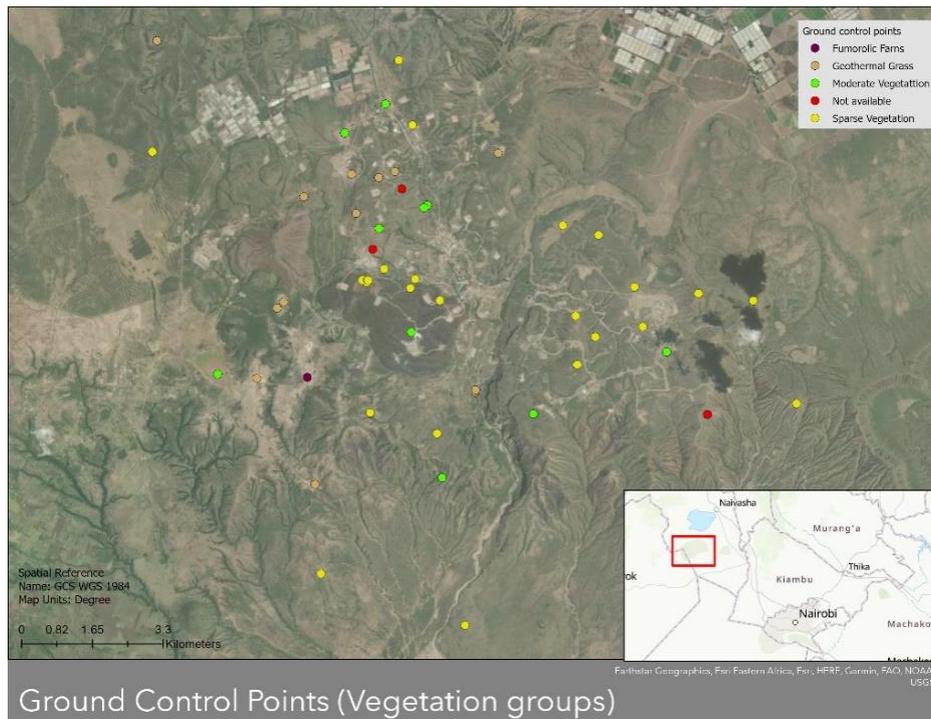


Figure 5: Vegetation groups associated with Ground Control Points

4.1.2. Daily Rainfall data

In addition to studying the change in vegetation for the entire chosen two-year period, the response curves after significant rain events were studied to understand the difference in the behaviour of the vegetation before and after rainfall.

A weather monitoring station is available in the Olkaria region, and the data from July 2019 to May 2022 was retrieved. Kenya has two rainfall seasons, termed short rains (Oct-Dec) and long rains with abundant rainfall (Mar-May) (Camberlin & Wairoto, 1997). The monthly average of rain received in Olkaria is shown in Figure 6.

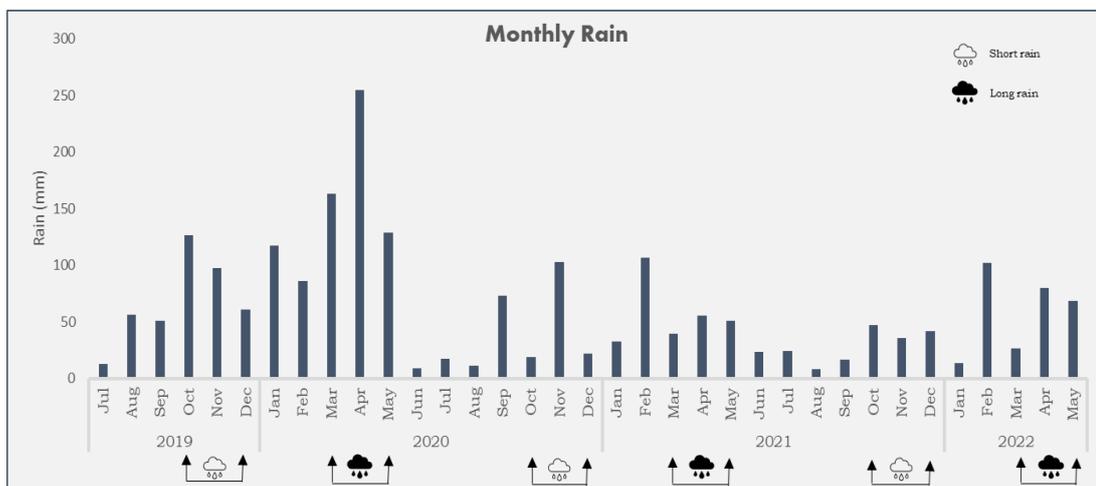


Figure 6: Monthly Rainfall in Olkaria

This figure shows that in the year 2019 until early 2020, Olkaria received heavier rainfall as compared to 2021 and 2022. Since the study focuses on vegetation dynamics, abnormally heavy rain is not considered suitable as the vegetation could be completely soaked, showing little to no variations. Heavy rainfall could lead to biased results, and hence, this study was conducted between June 2020 and May 2022.

Figure 7 shows an overview of rainfall distribution every five days representing the average rainfall for the entire area. 5 day period is chosen to account for the availability of optical images, in this case, Sentinel 2, which has a five-day temporal resolution.

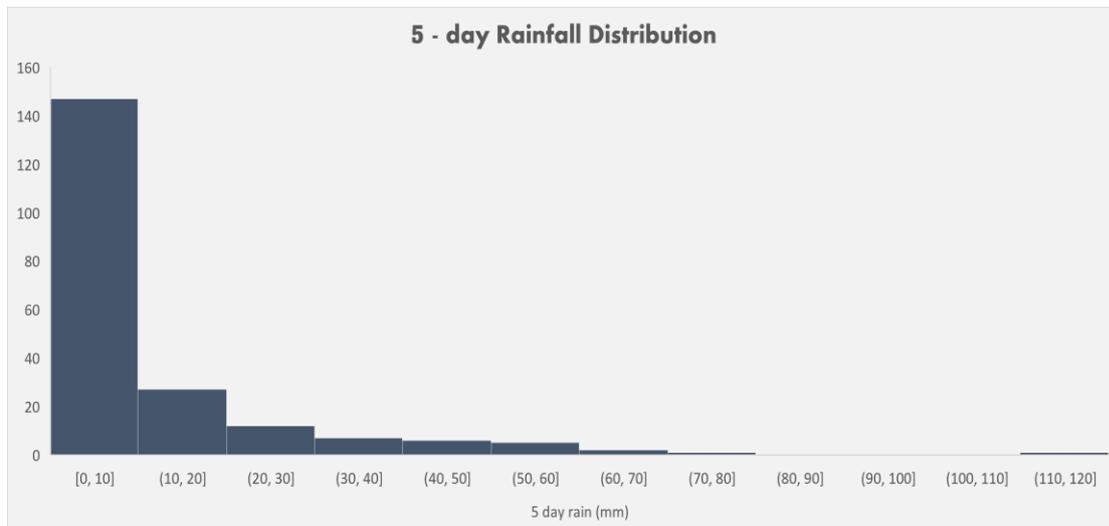


Figure 7: 5-day Rainfall Distribution in Olkaria

It can be seen that the majority of rainfall in 5 day time period was less than 10mm. Hence, rain events with rainfall above 10mm in 5 days were selected as this study focused on significant rain events that cause notable changes in the region to discern differences in vegetation. Amongst these rain events, the temporal profiles of different parameters were generated for which cloud-free data for at least 2 data points before the rain event (10-day period) and 4 data points after the rain event (20-day period) were available at the ground control point locations. The response curves around smaller to no rain periods were also analysed for comparison to ensure the observed results are specific to rain events, and not a general response, strengthening the result of the hypothesis.

4.1.3. Sentinel Optical Images for time series profiles

There were around 150 scenes available from the 5-day temporal resolution Sentinel 2 archive during the period of June 2020 – May 2022. The time series data from these images was downloaded using the capabilities of Google Earth Engine for each of the 56 ground control points for Dynamic Time Warping (DTW). In GEE, available Sentinel 2 images for the 56 Ground control points were cloud masked using the available QA60 band. This band is a quality control band which flags the clouds in the Sentinel-2 imagery using altitude thresholding with the B10 band (European Space Agency, n.d.). Although, this helps in clearing out some effect of the clouds, the band does not filter out all clouds from the image, hence to further remove cloud effect the data points which had NDVI values below 0.1 were also filtered out from all the time series to remove the effect of brightly reflective clouds before making the temporal profiles around rain events (Coluzzi et al., 2018). The temporal profiles were created using this data, though an additional check was done by manually checking the scenes visually in case any abrupt

responses in the temporal profiles was observed to ensure the clouds were not affecting the results or otherwise they were removed from interpretation.

The data was downloaded for the following parameters with their corresponding band in Sentinel-2 (Henrich, V., Krauss, G., Götze, C., Sandow, n.d.):

1. Red band (B4)
2. Green band (B3)
3. Blue band (B2)
4. Near Infrared (NIR) band (B8 – 10m)
5. Short wave infrared (SWIR) band (B11 – 20m)
6. Normalized difference vegetation index (NDVI) (Kriegler et al., 1969)

$$\frac{NIR - Red}{NIR + Red}$$

7. Normalized difference water index (NDWI) (McFeeters, 2007)

$$\frac{NIR - SWIR}{NIR + SWIR}$$

8. Enhanced vegetation index (EVI) (Huete et al., 1997)

$$2.5 * \frac{NIR - RED}{NIR + 6 * RED - 7.5 * BLUE + 1}$$

Additionally, to employ the results obtained by reviewing the time series profiles of the individual ground control points in detecting potential locations of geothermal hotspots, the NDVI time series data for the entire region of Olkaria is downloaded. Using this data, a full scene classification could be done to compare the accuracy of the proposed solutions using vegetation responses for identifying geothermal hotspots with land surface temperature study using Ecostress data (GeoHot project). The script was modified to download time series data for multiple polygons in a single run. For each run, a 10m x 10 m resolution mesh grid was provided as an input to match the resolution of Sentinel images in the format of a shapefile covering a part of the Olkaria region. This script automatically generated a CSV file of the NDVI time series for each polygon in the grid, available for download in the google drive of the account.

4.2. Data Processing

The data collected in the previous step was modified and refined to extract relevant information for further analysis. A summary of the procedure followed is shown in Figure 8.

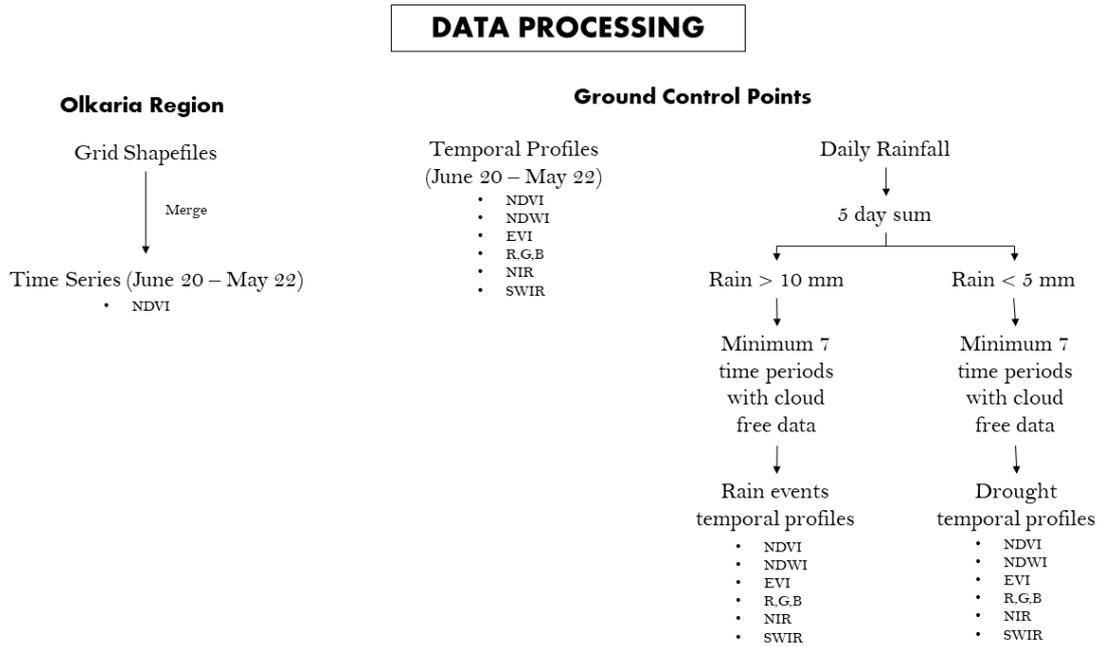


Figure 8: Data Processing Workflow

4.2.1. Ground Control Points

The temporal profiles for geothermal hotspots and non-geothermal hotspots were averaged for the entire time period of 2 years to study the differences in the vegetation responses. Further, rain events with greater than 10mm rainfall in 5 days were identified, for which enough cloud-free data was available to cover ten days before the significant rain event and 20 days after the rain event. The idea was to discern how the vegetation is affected around these events, taking into consideration two-time points before the rain event and four after it, covering a time span of about a month. The averaged-out temporal profiles of geothermal and non-geothermal hotspots were created for each event from the available time series data of the eight parameters. In addition to the rain events, to identify significant differences in response functions around rain events, these temporal profiles were also compared with those during a prolonged period of no rainfall, also termed drought temporal profile in this context. This helped in understanding if the differences between geothermal hotspots and non-geothermal hotspots were characteristic around rain events or are also seen during dry periods. The criteria for selection of dry periods was less than 5mm of precipitation every five days for at least a month (30-day time period).

4.2.2. Olkaria region

The NDVI time series data for the entire region was downloaded by dividing the area into a mesh grid of 10mx 10m resolution (Figure 9). The resolution was similar to the used Sentinel 2 optical imagery in the study and the generated shapefile was used for analysing the average vegetation response values of the selected two years for the entire region of Olkaria.

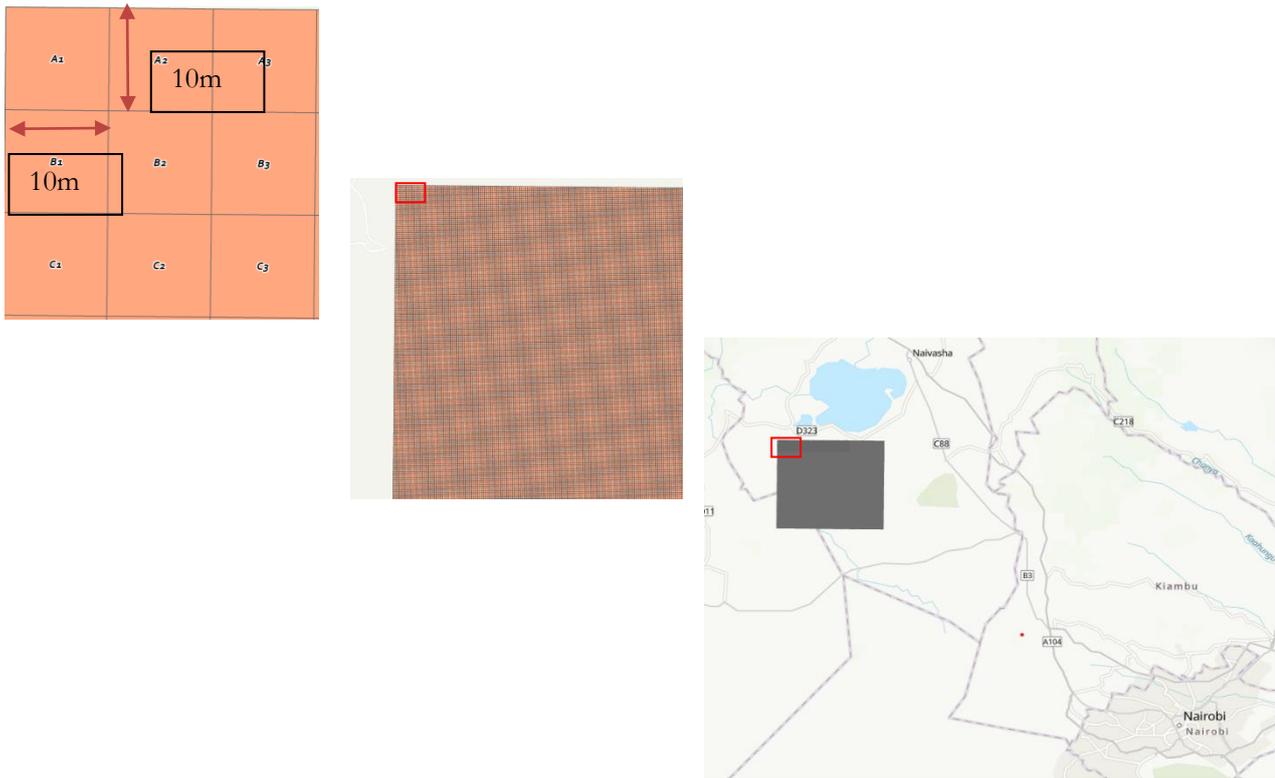


Figure 9: Mesh grid setup for downloading the NDVI time series data for Olkaria

4.3. Data Analysis

The temporal profiles generated in the previous steps are used for initial hypothesis testing. A qualitative assessment of the profiles identifying any typical response or patterns was done to understand the behaviour of the vegetation in such geothermally active regions. Initially two kinds of observations were made (Figure 10), one related to the range of values for different spectral bands and indices for the entire time period and how it relates with the vegetation dynamics. The second was to visually analysed the differences in the shapes of the temporal profiles around rain events between geothermal hotspots and non-geothermal hotspots. Through this assessment, the surface characteristics of the Olkaria region could be interpreted. This helped identify specific parameters (among Red, Blue, Green, NIR, SWIR, NDVI, NDWI and EVI) that showed varied surface characteristics between geothermal and non-geothermal hotspots for further quantitatively evaluating the proposed hypothesis. The most distinct parameters could be used as an additional specification to distinguish geothermal hotspots. However, prior to exhibiting any conclusions, the results from the visual assessment were quantitatively verified using Dynamic Time Warping (DTW). DTW compares the shapes of the temporal profiles for geothermal hotspots and non-geothermal hotspots to check the dissimilarities and therefore assess its applications in identifying potential geothermal hotspots. Additionally, the distinct parameters for differentiating geothermal hotspots are also examined for their range of values and if any considerable differences are observed to be applied in the entire region of Olkaria.

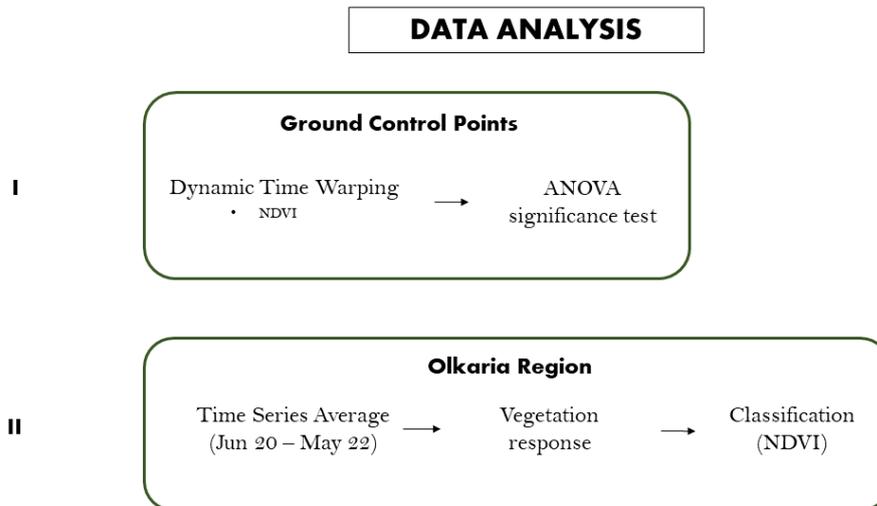


Figure 10: Data Analysis Workflow

4.3.1. Dynamic Time Warping (DTW)

The generated temporal profiles of NDVI were analysed using Dynamic Time Warping (DTW). This method assesses the shape of the variable dynamics rather than only the absolute values it represents. This step will test the hypothesis wherein a steeper slope is expected for declining vegetation presence at active geothermal points relative to vegetation on non-hotspots after an external environmental factor such as rain event.. DTW is an algorithm that can be used to compare different temporal sequences keeping in consideration the local space and time shifts. It was developed more than five decades ago for better speech recognition (Sakoe & Chiba, 1978) and eventually found application in time series matching in the 1990s (Bemdt & Clifford, 1994). DTW is now widely utilized for time series analysis in various remote sensing applications (Petitjean et al., 2012), such as for studying vegetation and agricultural changes, and is often considered a better dissimilarity measure for comparing time series data (Virnodkar et al., 2020). Fundamentally, they compare the shapes of the curves to assess the dissimilarities in the shapes of the curve rather than how the values change at each point. This helps in considering the relationship between one point in space of a sequence with one or more points in the other sequence to be comparable in the shapes of the graph. Standard matching techniques use distance matrix to understand similarity, with Euclidean distance the most commonly applied. Higher average distance shows higher dissimilarity in the compared curves. DTW is based on a similar concept; however, it does not match the curves point by point in space and time but considers its shape by comparing different points (Figure 11). It tries to superimpose the two temporal sequences by matching the different points, and find the minimum alignment cost required for this which is represented by the normalized distance parameter. It further

generates a distance matrix, along which the path of minimum alignment cost shows how dissimilar the curves are in shape (Bisht et al., 2016).

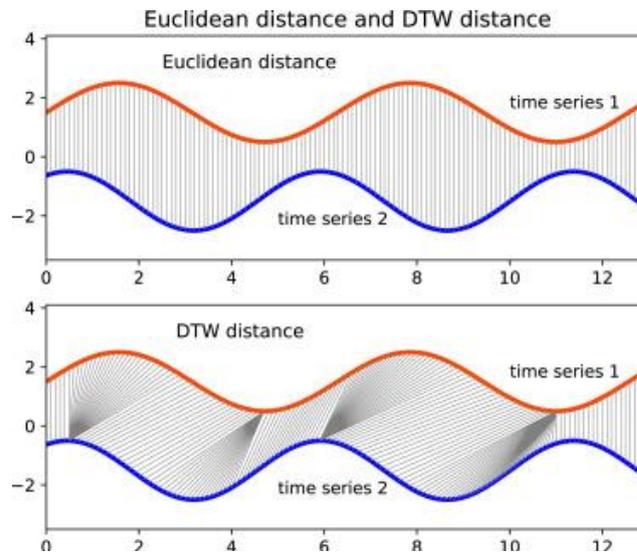


Figure 11: Illustration on how Dynamic Time Warping works (Lu et al., 2016)

The Dynamic Time Warping assessment results were tested for significance using the Analysis of Variance (ANOVA) significance test (Ohana-Levi et al., 2020). This test checks for statistical significance between different categories by analysing the variances in the means of the classes. The DTW analysis checked for curve shapes under three categories, comparing geothermal hotspots with themselves, non-geothermal hotspots with themselves, and finally, comparing shapes of curves between geothermal hotspots and non-geothermal hotspots. The ANOVA test checks if the dissimilarity distances are significantly different among each of these three categories to be implemented as a parameter for identifying geothermal hotspots.

4.3.2. Full scene classification of Olkaria region

In addition to examining the ground control points, the purpose is to find potential geothermal hotspots in a geothermally active region of Olkaria. If the temporal profiles show significant differences between geothermal hotspots and non-geothermal hotspots, the results will be used to classify the entire region into potential zones of geothermal hotspots, based on the vegetation responses around rain events.

Combined with the qualitative and quantitative assessment related to studying the temporal profiles of various parameters, the hypothesis is tested to examine vegetation dynamics as an additional parameter for distinguishing points of geothermal hotspots after exploring the region for thermal anomalies. The following chapter presents the results of the various steps and discusses the expected outcomes to validate the hypothesis.

5. RESULTS

5.1. Temporal profiles of geothermal and non-geothermal hotspots

The initial study was done to seek any typical responses as stated in the hypothesis (Figure 1) related to the temporal profiles of any spectral band or generated vegetation indices from the satellite images and establish those parameters (Red, Blue, Green, NIR, SWIR, NDVI, NDWI and EVI) for the subsequent investigation to understand the vegetation dynamics associated with geothermal hotspots. There were some trends in vegetation dynamics that visually appeared to be different between geothermal hotspots and non-geothermal hotspots, however, not all the indicators had similar responses (Figure 12). The temporal profiles for the two years, though cloud masked had some sudden dips which affected some individual data points, hence the focus here was to look at the general response of the profiles and their pattern over two years.

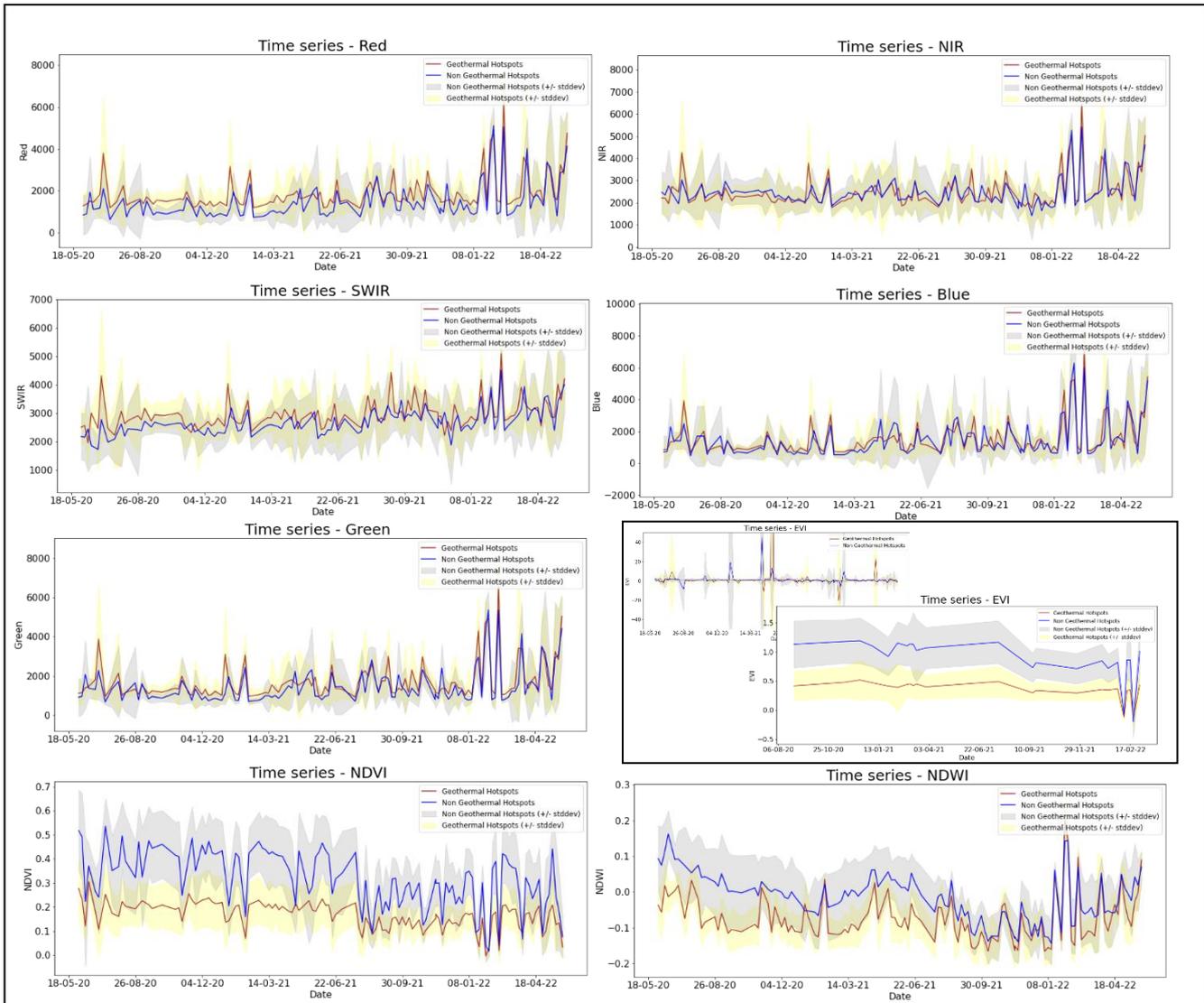


Figure 12: Temporal profiles of some bands and indices at geothermal and non-geothermal hotspots between June 2020 – May 2022

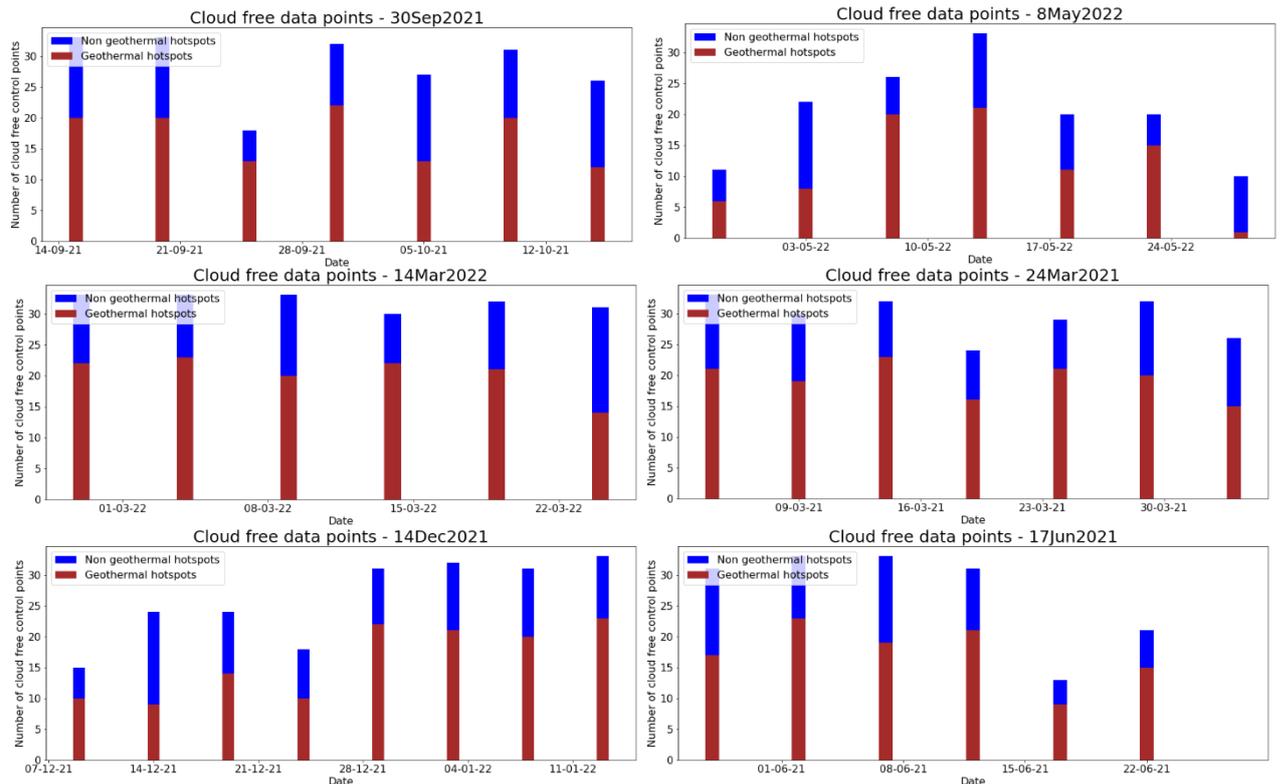
There are two graphs shown for EVI (third row on the right) as it had a lot of noise in the dataset, but still had differences in the values of geothermal and non-geothermal hotspots.

It was observed that among the bands, Red and SWIR bands had some differences in the value range as opposed to blue and green bands, which show similar values throughout the two years for geothermal hotspots and non-geothermal hotspots. Hence, the blue and green bands were not used for further analysis. Across the entire time range considered, NDVI and NDWI stand out, as consistent differences in values for hotspots and non-hotspots appear. These observations were logical as they are in line with the systematic differences we see in Red and SWIR bands which make up these indices. EVI had a lot of variations in value with extreme outlier data points as seen in the inset graph and even though removing these outliers showed quite some difference, there were not enough data points to consider for the study. Looking at the range of these indices, the geothermal hotspots consistently have a lower vegetation cover which needs to be further analysed to relate to the presence of geothermal hotspots.

While the time series data did show differences in response values for geothermal and non-geothermal hotspots for some parameters, to ascertain the presence of a distinctive temporal profile at geothermal hotspots, in other words, recognizable differences in the shape of the time series curves, further analysis had to be performed.

5.1.1. Temporal profiles around rain events

Over the period of two years, 10 rain events were identified to understand patterns in the temporal profiles of the selected parameters. These events had more than 10mm of rainfall over five days in the entire region of Olkaria. For each of these rain events, first cloud masking was done, followed by a filter based on NDVI values before the temporal profiles were made. The number of data points for each of the acquisition date out of the 33 non-geothermal points and 23 geothermal points around the rain events after these two steps is shown in Figure 13. It can be seen that after the rain event of January 28, 2022 (highlighted), there is not even one data point available at two time points (February 2, 2022 and February 7, 2022). Hence this rain event was not included for further analysis.



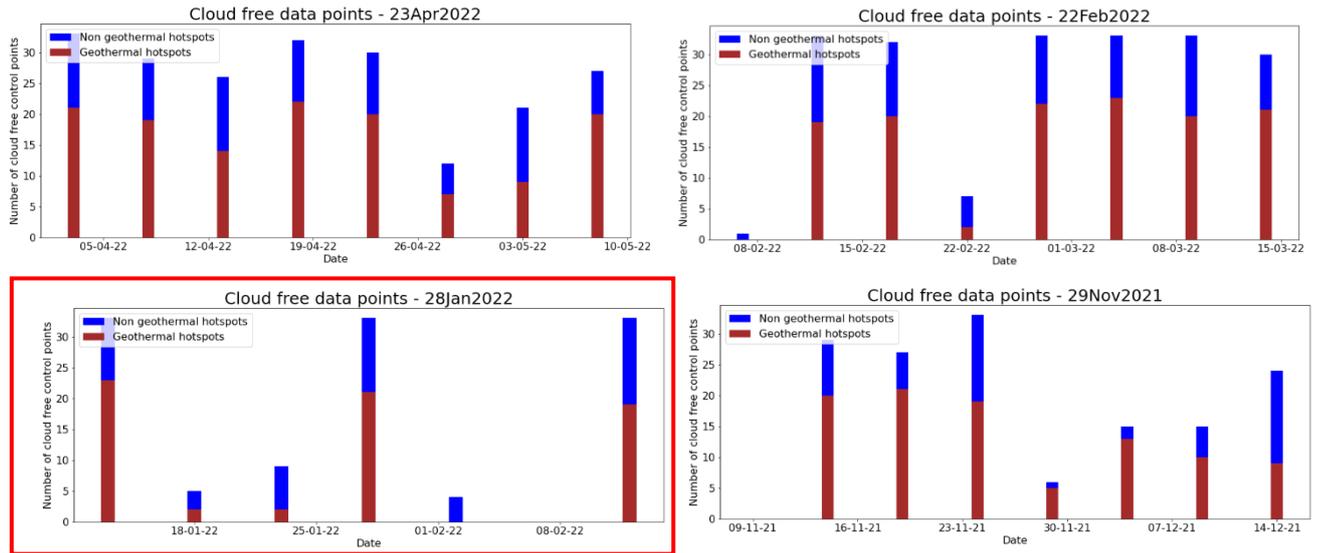


Figure 13: Number of Cloud free data points per rain event after cloud masking and NDVI value filter (NDVI < 0.1)

The temporal profiles for the other rain events for each of the chosen parameter NDVI, Red, Nir, NDWI and SWIR is shown in the following sections.

5.1.1.1. NDVI

The NDVI profiles (Figure 14) show some minor changes in the way the graphs dip around rain events, in this scenario, a larger dip at non-geothermal hotspots. Based on qualitative assessment, most of the changes were not as considerable for smaller rain events showing a small decline in the values, especially at geothermal hotspots and could not be significantly proven only by visual analysis.

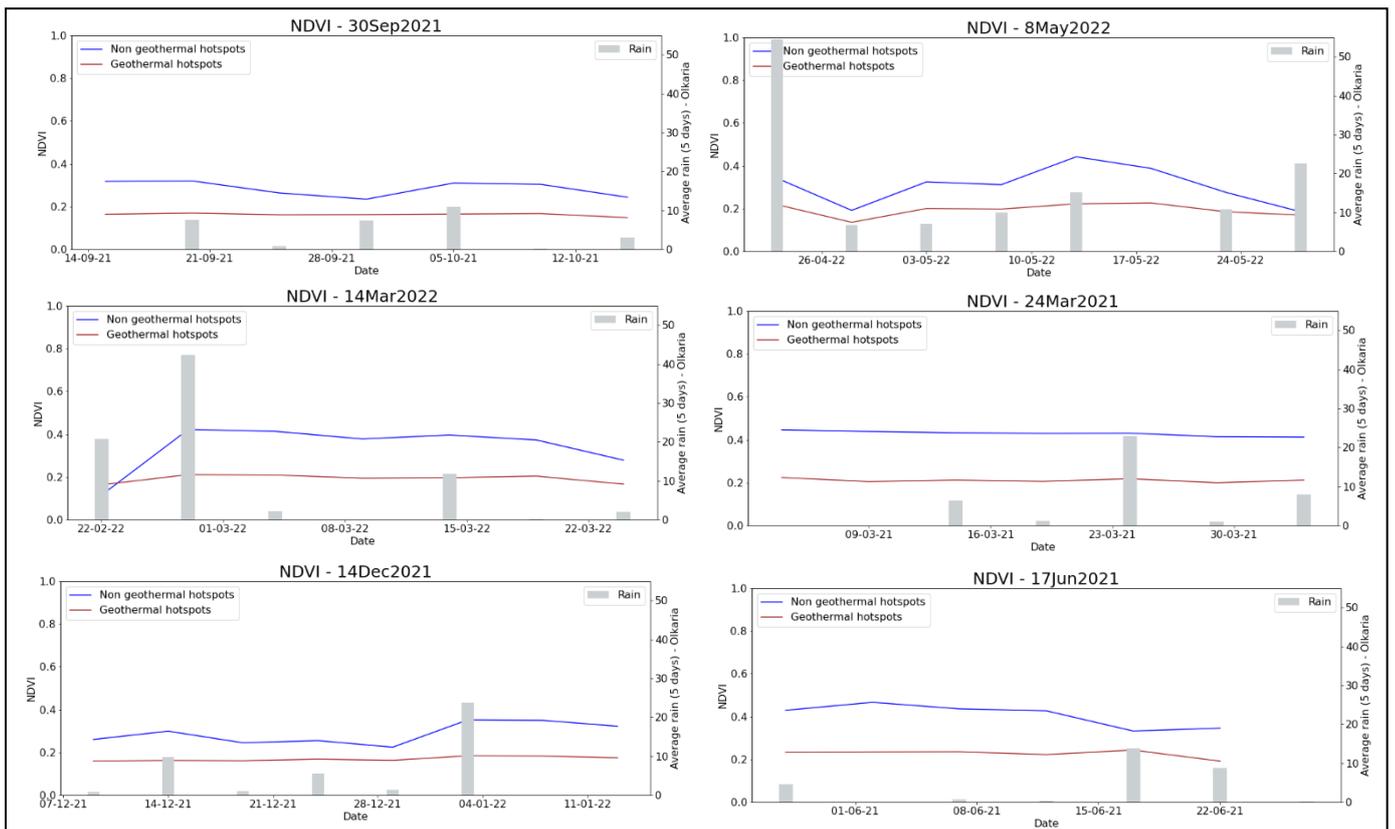


Figure 14: Temporal profiles around few rain events for NDVI

Though at some of the rain events, especially the larger ones there was a significant dip observed with NDVI values very close to 0.1 (Figure 15).

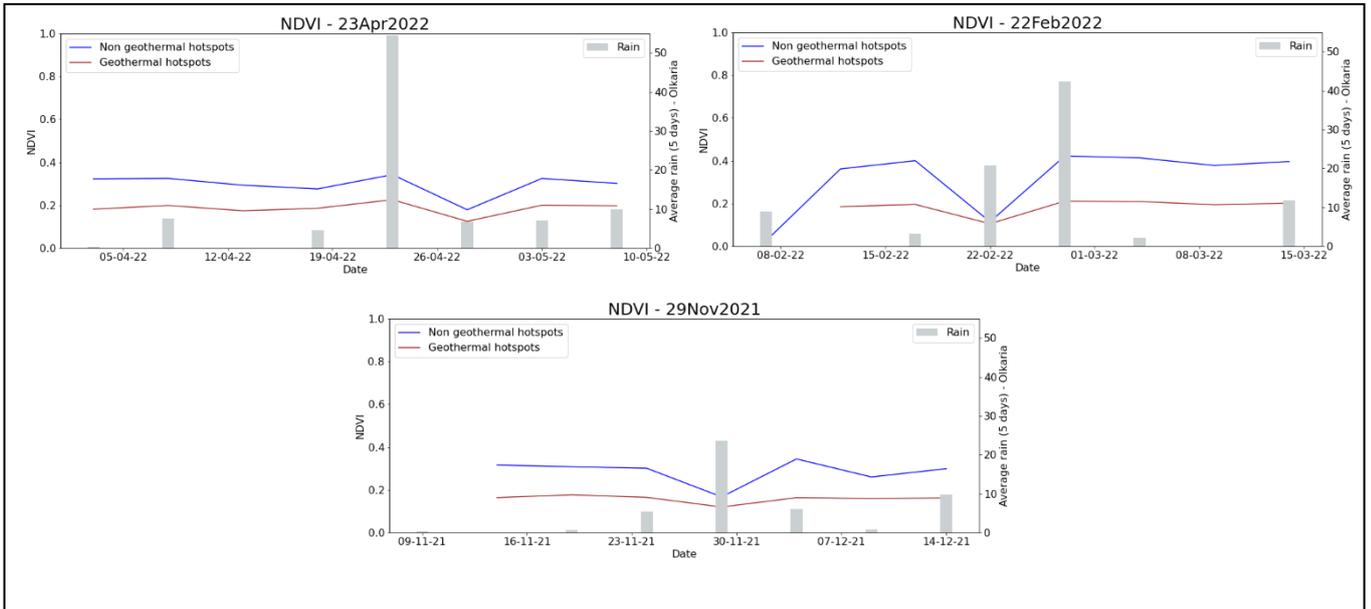


Figure 15: Temporal profiles of NDVI around rain events which showed abrupt changes

These three temporal profiles were further verified for clouds by visually checking the scenes and it could be observed that even after cloud filter, the remaining points were also highly contaminated by clouds (Figure 16). Hence, these rain events were removed from the analysis.

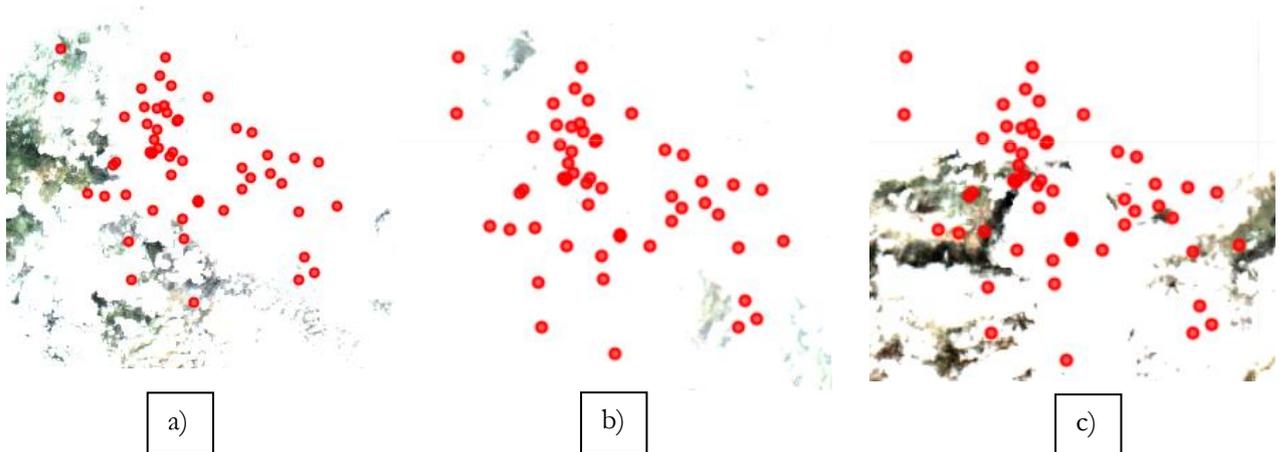


Figure 16: Contaminated images with very high cloud cover

a) April 28 2022 b) February 22, 2022 c) November 29, 2021

Looking at the cloud contamination in Figure 16, to verify the accuracy of the other temporal profiles (Figure 15), additional visual checks were done. The points used for making the profiles were manually checked to be cloud free points. Some examples are shown in figure 17.

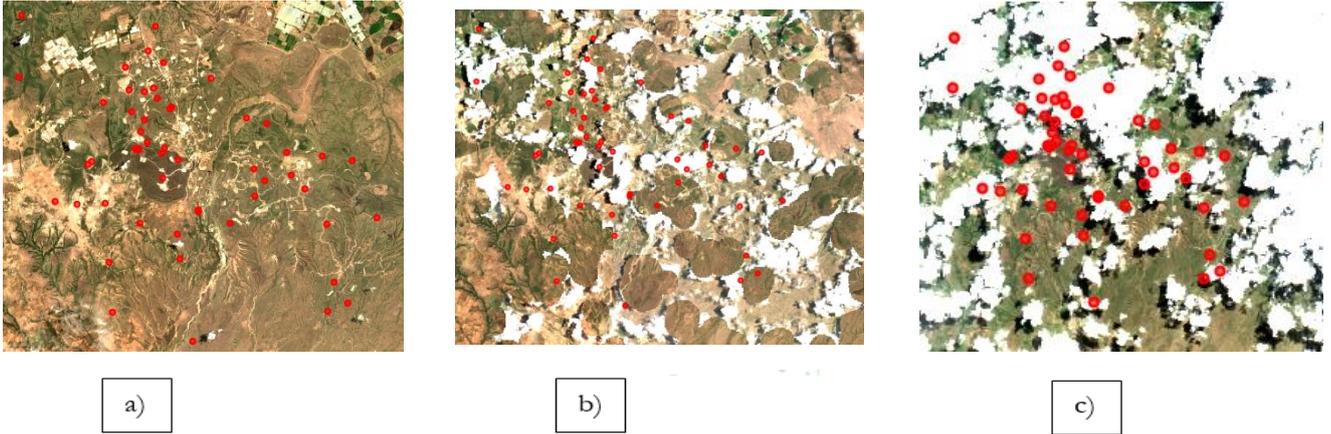


Figure 17: Examples of visual checks for extracting cloud free points
 a) September 15 2021 b) September 25, 2021 c) May 3, 2022

We previously observed that NDVI also showed a consistently significant difference in the NDVI range over the two years between geothermal and non-geothermal hotspots. This was observed around rain events as well. In order to understand the observed subtle differences in temporal profiles in combination with these value differences between geothermal and non-geothermal hotspots a quantitative assessment by comparing the shapes of these graphs was done using Dynamic Time Warping (DTW). This was carried out to identify any differences, however small in the temporal profiles of geothermal and non-geothermal hotspots and check its viability for identifying potential locations of geothermal hotspots (Section 5.2).

5.1.1.2. Red

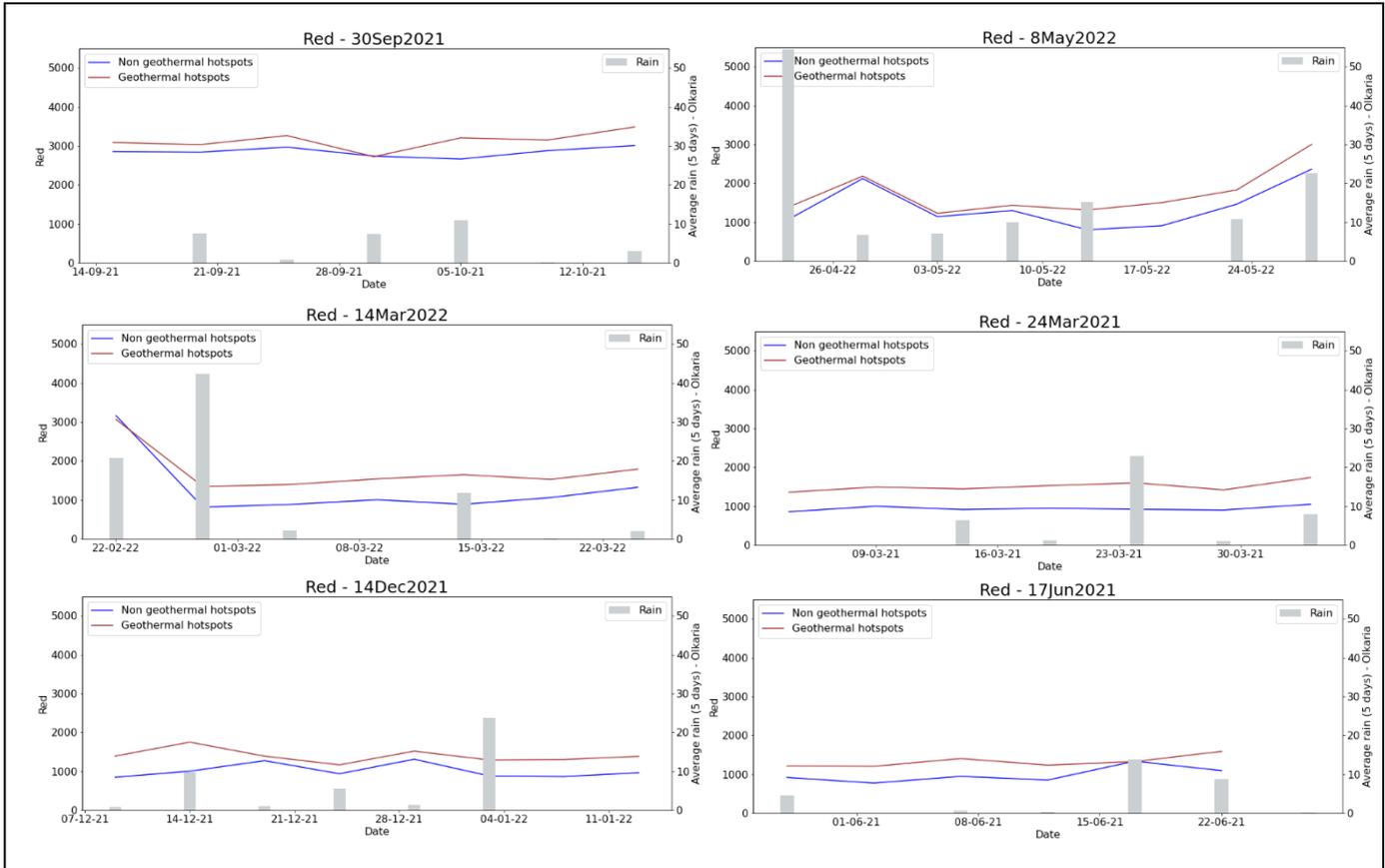


Figure 18: Temporal profiles around few rain events for Red band

A quick visual assessment of the temporal profiles of Red band (Figure 18) does not verify the presence of any specific behavioral response in the temporal profiles of geothermal hotspots and non-geothermal hotspots in the red band. The only noticeable point is the dip in the curves of the non-geothermal hotspots after rain is more rapid than for geothermal hotspots; however, it is quite negligible.

5.1.1.3. NIR band

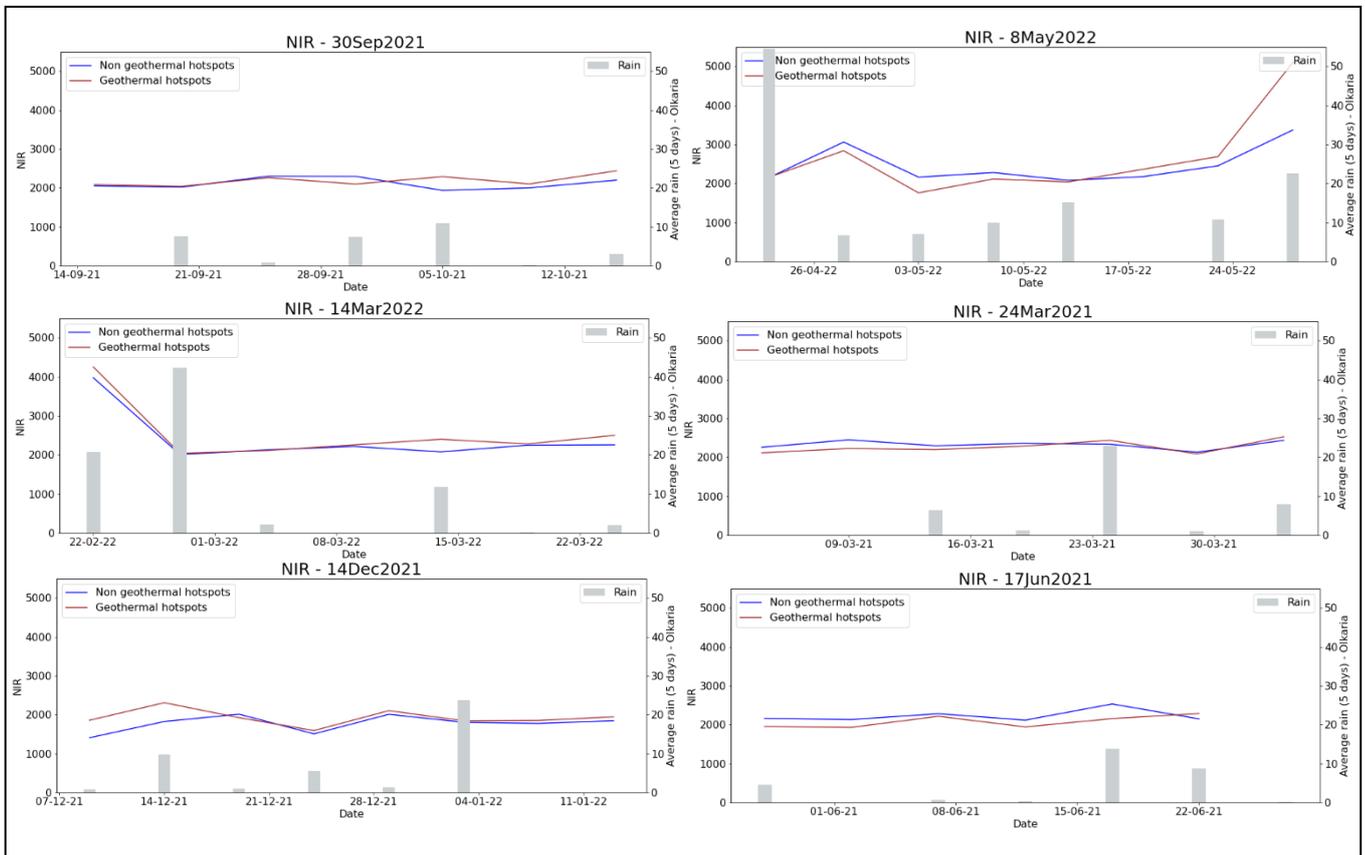


Figure 19: Temporal profiles around few rain events for NIR band

Temporal profiles of the NIR band (Figure 19) show a similar response for both geothermal and non-geothermal hotspots with no major inferable differences.

5.1.1.4. SWIR band and NDWI

Figure 20 shows a few examples of the temporal profiles around rain events for the SWIR band. It did not show any pattern that can be used to differentiate between the profiles of geothermal and non-geothermal hotspots, except at the times of very heavy rainfall. This band was used for calculating the NDWI, and as seen in Figure 21, the NDWI values did not show any changes in the temporal profiles around rain events, even though the values were different over the two years for geothermal and non-geothermal hotspots.

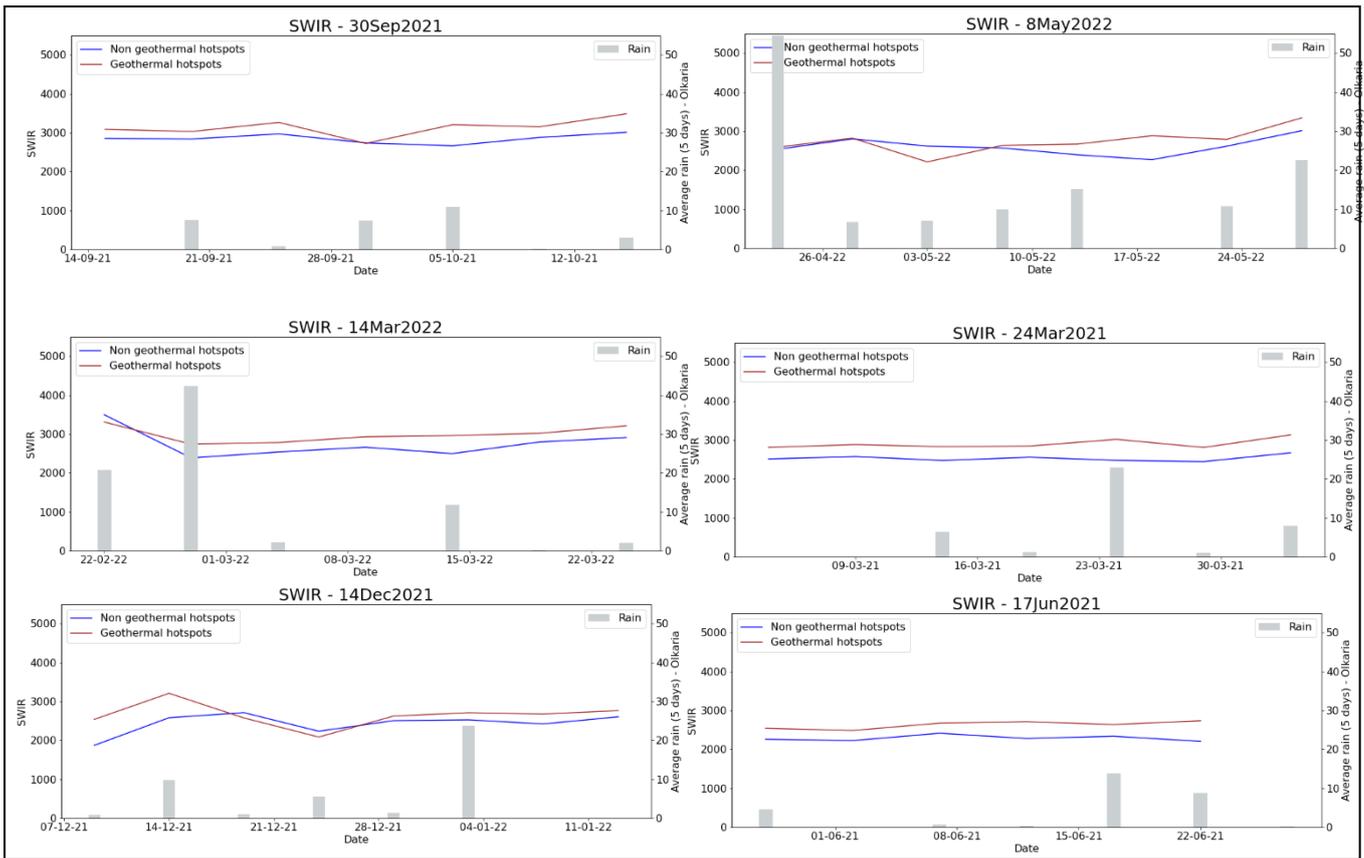


Figure 20: Temporal profiles around rain events for SWIR band

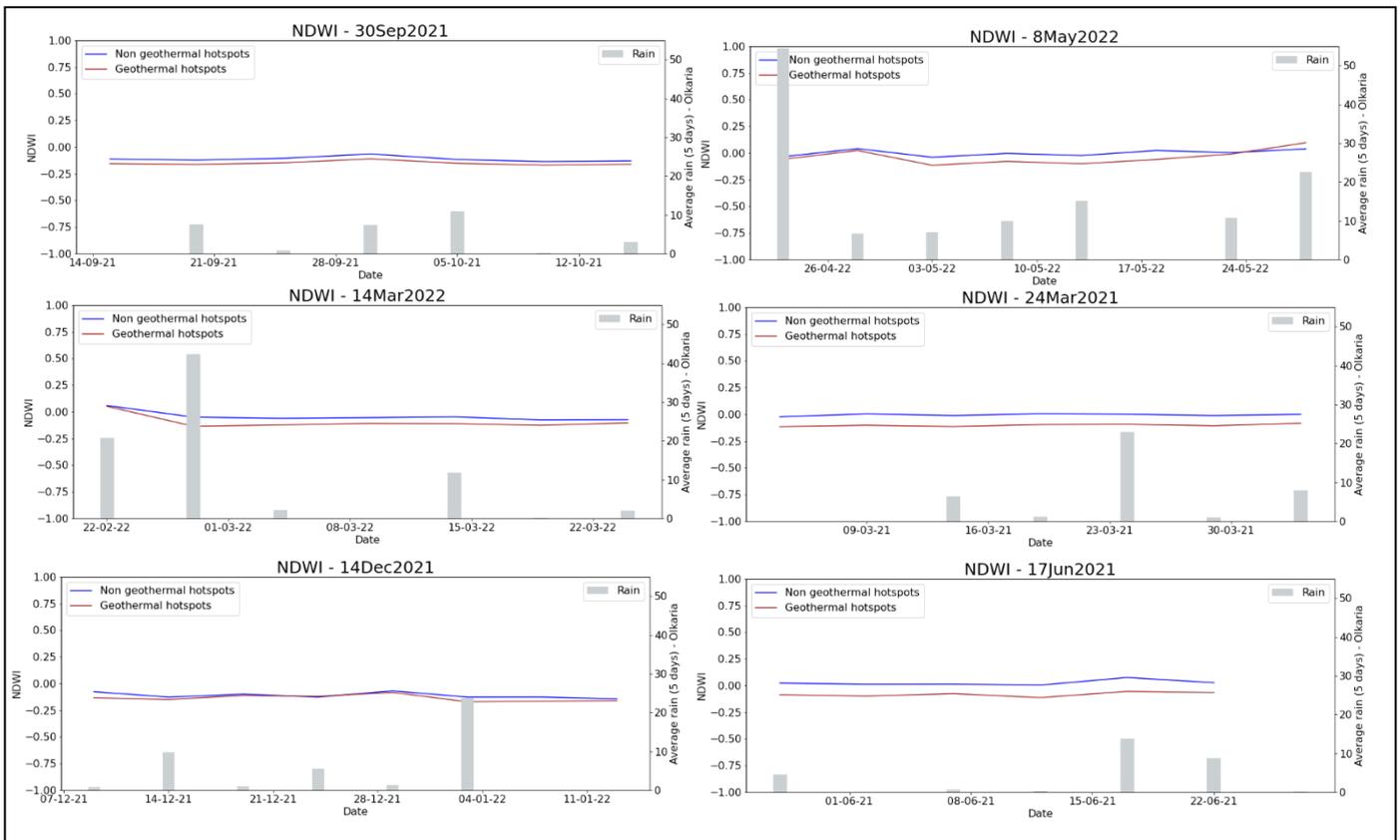


Figure 21: Temporal profiles around few rain events for NDWI

The lack of any pattern or value differences in the temporal profiles in the two suggests in this study that the moisture content in geothermally active zones cannot be easily deciphered and requires further analysis related to soil moisture content, which was out of scope for this study

In this section, the temporal profiles of some of the parameters around rain events were analysed, which indicated a visual pattern in the Red and NIR bands and hence in the NDVI profile. This was based on the concept of the effect of rainfall on vegetation and how this can change in the presence of a geothermal hotspot. In addition, to isolate this qualitative assessment during rain events, the temporal profiles of NDVI were also assessed during a monthly period when there was not any significant rainfall (Figure 22). It can be seen that the shapes here tend to remain the same over the month.

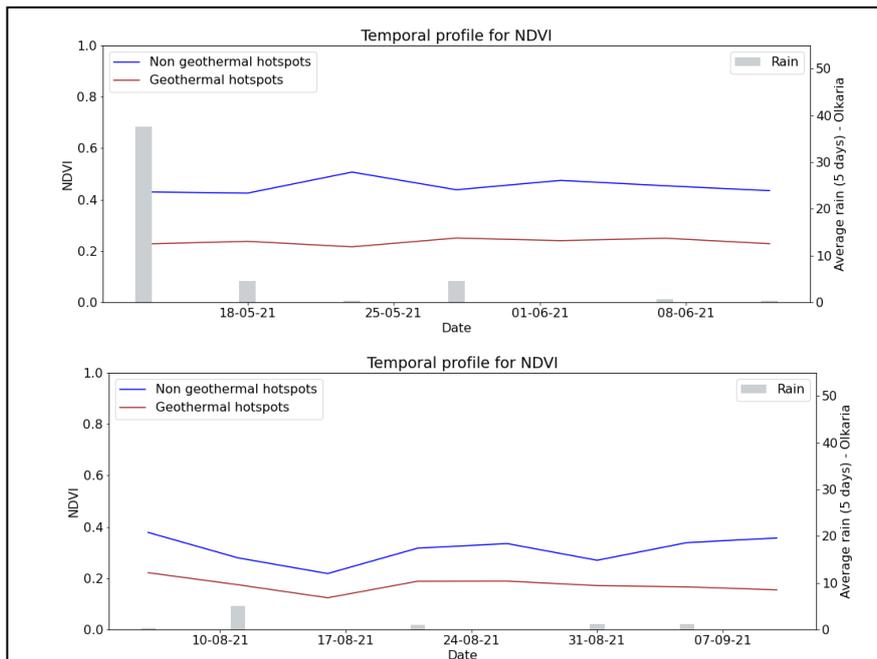


Figure 22: Temporal profiles during periods of no rain for NDVI

5.2. Temporal profile analysis using Dynamic Time Warping

The study of the temporal profiles over the time period of two years as well as around rain events show that the NDVI has a slightly different response for geothermal hotspots. This is, in effect, determined using the Dynamic Time Warping technique, which compares the shapes of the curves, aligning them together and yielding the alignment cost given by the normalized distance parameter. This way, the similarity of the curves can be assessed. In this study, the similarity of the temporal profiles or the shapes of the time series curves is determined in three categories:

- i. Two geothermal hotspots (1 – 1)
- ii. Two non-geothermal hotspots or the background points (0 – 0)
- iii. A geothermal hotspot and non-geothermal hotspot (1 – 0)

5.2.1. DTW analysis over the time period of two years

The following figure (Figure 23) shows the distribution of the normalized distance in these three comparisons, calculated by randomly choosing points from the available 56 ground control points, tested with 2000 different combinations. Here 0 indicates no geothermal activity, and 1 indicates the presence of geothermal activity (geothermal hotspot). This chart shows how the temporal profiles compare; the higher the normalized distance, the more the curve shapes differ.

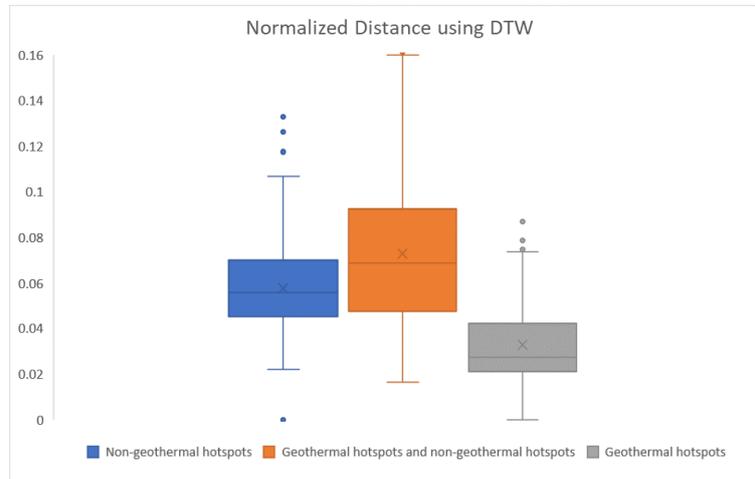


Figure 23: Normalized distance differences using DTW for two years

It illustrates that the temporal profiles of two geothermal hotspots (1-1) are more similar to each other in contrast to when the temporal profile of a geothermal hotspot is compared with a non-geothermal hotspot (1-0). In addition, comparing two non-geothermal hotspots (0-0) shows higher differences even though they are the same category. This suggests geothermal hotspots have distinctive responses that can be used to isolate these points in a geothermally active zone by comparing their similarities. These variations were established through an ANOVA significance test, indicating notable differences in the mean of these groups. The p-values was considerably less than 0.001, concluding that these box plots have different means and have significantly different values.

While this shows variances in the temporal profiles, this might also be a result of the fact that geothermal hotspots often have lower vegetation cover; hence, such could be attributed to merely the differences in the vegetation. Therefore to further verify these, the data points with similar vegetation cover were compared to find the differences between geothermal and non-geothermal hotspots. Figure 24 shows the distribution in the DTW normalized distances when comparing points with sparse and moderate vegetation cover. The results are similar to those previously observed, indicating differences in the temporal profiles of geothermal and non-geothermal hotspots.

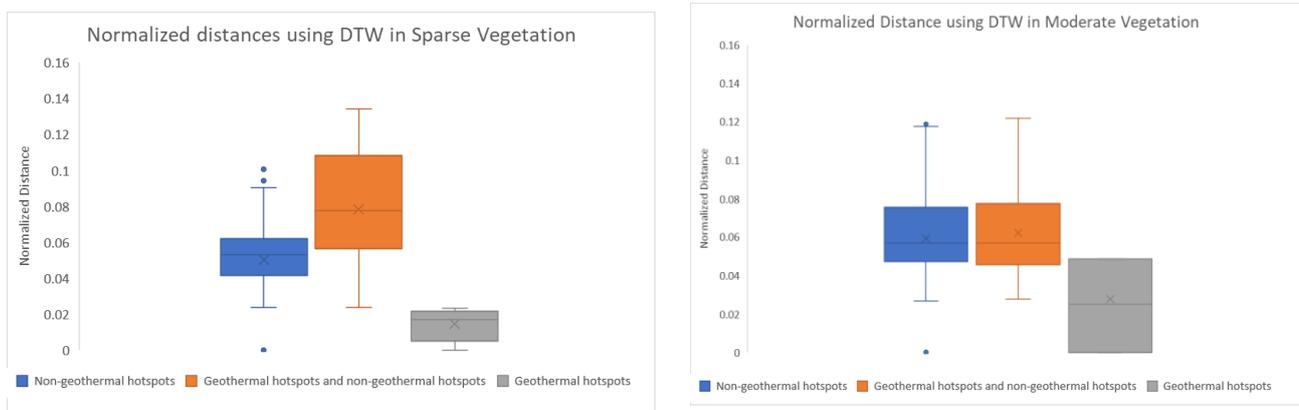


Figure 24: Normalized Differences using DTW in different vegetation cover

The difference in distances between pairs of geothermal hotspots, pairs of non-geothermal hotspots, and the combination of geothermal hotspots with non-geothermal hotspots is much more pronounced in the sparse vegetation plots compared to the moderate vegetation plots. The results (Figure 24) seem to suggest that the difference between a geothermal hotspot and a non-geothermal hotspot is actually more considerable under sparse vegetation. The ANOVA significance results show that when comparing the three plots, the results were significant ($p < 0.001$), however in moderate vegetation, the differences when comparing the variances of distances for two non-geothermal hotspots and a non-geothermal hotspots with a geothermal hotspot were not as significant. This suggests that possibly that the difference would have been even less if there was abundant vegetation further inferring that hotspots can be better differentiated on areas with sparse vegetation.

5.2.2. DTW analysis during rain events

The DTW analysis that explores the similarity in the shapes of the curves indicated in the previous section that the temporal profiles around geothermal hotspots are very similar (shown by smaller distances to align them together) as compared to when any other non-geothermal hotspot is compared with itself or with a geothermal hotspot.

The differences in the temporal profiles around rain events visually showed some pattern, but to substantiate if these shapes are actually different around these rain events, DTW was performed. This showed similar responses to as observed earlier for all the rain events, few of which are shown in Figure 25. Even if some events were biased because of less number of cloud free (or noise free) samples, the results were consistent across all the rain events independent of the rainfall received, indicating some differences in the temporal profiles of geothermal hotspots compared to non-geothermal hotspots. These results were also tested using the ANOVA significance test, and for all three categories the p value was less than 0.001.

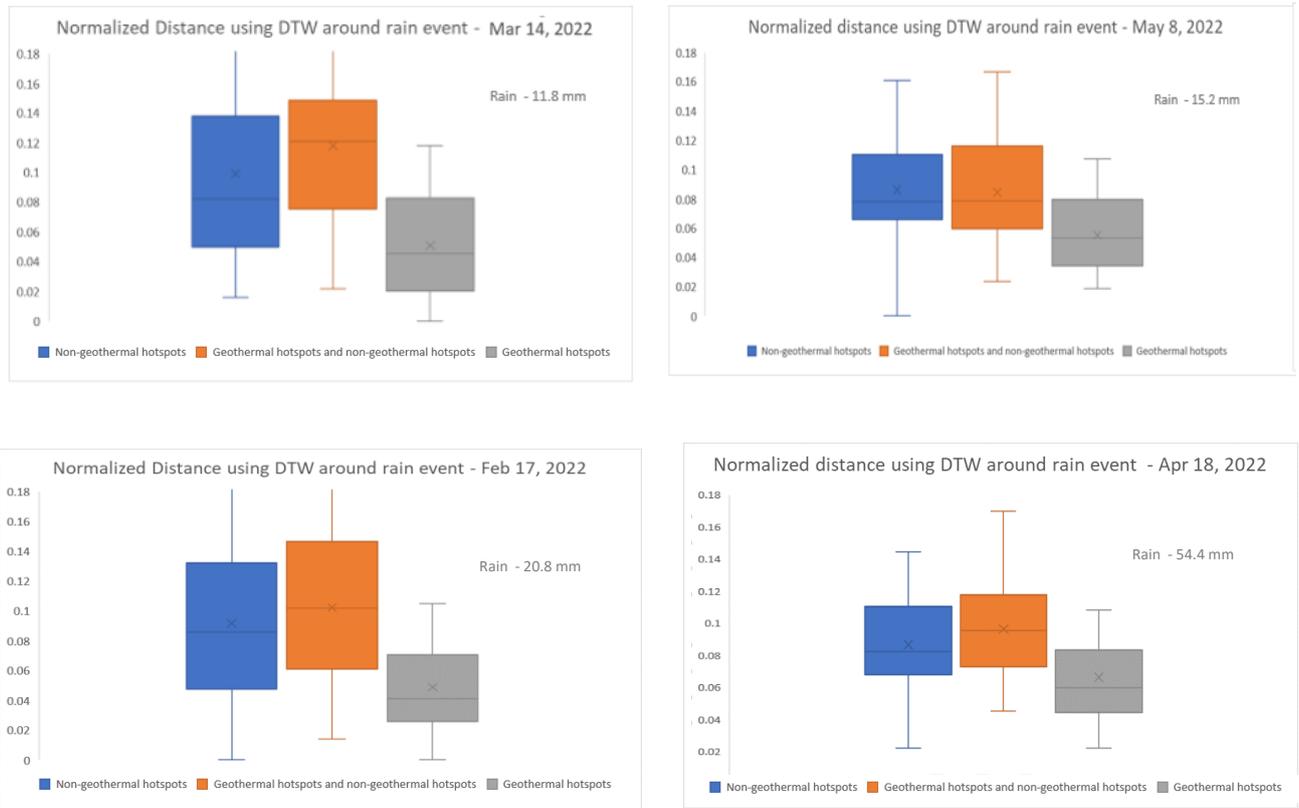


Figure 25: Normalized distance distribution using DTW around few rain events

The distribution of the distance for aligning the curves together mostly shows similar response values for the period of two years and around rain events when comparing different categories. This further verifies that the curve similarity is more around geothermal hotspots, implying distinctive response features at these points.

6. DISCUSSIONS

The Olkaria region in Kenya is characterized by high-temperature hydrothermal activity, altering the physical properties of the region. These surface characteristics are indicators of the subsurface properties and could be utilized to identify geothermal hotspots in the region which have a steeper temperature gradient. This study focused on analyzing the vegetation dynamics in these active regions and potentially assessing this as a parameter for separating out zones for locating the geothermal hotspots.

The vegetation in the Olkaria region is characterized by low to moderate vegetation cover consistent over a year, except for the year 2020 (Figure 26), during which the region also received heavier rainfall than usual (Figure 6). This data was available from the GeoHot project of Faculty ITC which combined the use

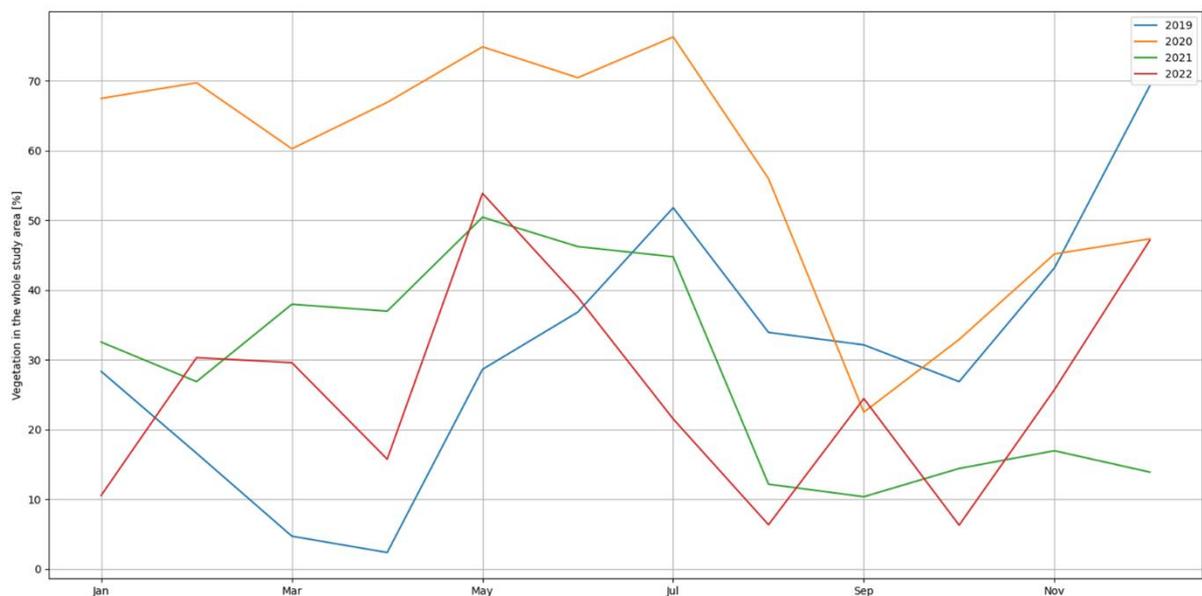


Figure 26: Vegetation cover in Olkaria between 2019-2022 (GeoHot Project, Faculty ITC)

of remote sensing and ground truth data to study the variations in the vegetation cover in Olkaria between 2019 and 2022. While the overall percentage of vegetation cover in the entire region is necessarily stagnant (about 30-40% vegetation cover), the effect of ground heating and fumaroles in the region is expected to alter the dynamics of the vegetation. The hypothesis stated in this study expected that the vegetation on geothermal hotspots would respond differently to increased soil moisture as a result of rain compared to vegetation on non-geothermal hotspots, because the soil moisture would evaporate faster on geothermal hotspots due to the raised ground temperatures. However, the scarce vegetation cover of Olkaria, limits how the vegetation response is captured by the Sentinel-2 sensor, objecting the results to be considered as vegetation dynamics or as conclusively evident of soil dynamics. Such limitations had to be considered in the results and in understanding of the vegetation dynamics of geothermal hotspots, specifically in Olkaria region.

The NDVI and EVI indices, along with the Red band, showed certain absolute differences in their values for geothermal hotspots and non-geothermal hotspots, all indicative of the differences in the presence of green vegetation between the two categories. The range of values for NDVI was usually below 0.25 for geothermal hotspots and greater than 0.35 for non-geothermal hotspots. Considering the low NDVI values at geothermal hotspots, and the indications of sparse vegetation, the observed responses are not

entirely vegetation responses but are also affected by the bare soil. For instance, the Red band has a higher reflectance for the geothermal hotspots which could be directly linked to the presence of alteration minerals, such as iron oxides in the soil in geothermal zones (Kamau et al., 2020), which reflect more in the red band wavelength (Frutuoso et al., 2021).

The SWIR band and NDWI index did show some differences in their values, but these were not prominent around rain events. Also, the blue band showed no such differences, hence limiting the understanding of the variations in the water content in the vegetation of geothermal and non-geothermal hotspots. The analysis of vegetation indices showed that the geothermal hotspots mostly have sparse vegetation and could be at bare patches of soil, indicating that the rain event affects the soil moisture and not the vegetation moisture content for such control points. Since these bands are not sensitive to the differences in soil moisture content (Hegazi et al., 2023), their use was of little importance for differentiating geothermal hotspots. This was further realized when the individual temporal profiles around rain events were studied, and negligible differences were seen in the time series of these parameters (Figures 20 and 21).

The vegetation in Olkaria region is highly altered because of the geothermal heating, affecting how the values in the NIR, SWIR and Red band show up for the geothermal hotspots and non-geothermal hotspots. The value range is indicative of bare soil rather than vegetation responses. Although based on photo evidence of presence of vegetation at geothermal hotspots made available by the GeoHot project of Faculty ITC, the lower values is mainly depicting unhealthy vegetation or bare soil at some points.

Analyzing the general dynamics {or pattern} of the reflectance data and the calculated indices over two years helped recognize the parameters that could be used for studying the temporal profiles or shapes of the curves for vegetation dynamics at geothermal hotspots. These temporal profiles were analysed around rain events to understand if there is a specific pattern in the behaviour of vegetation on geothermal hotspots. Based on the hypothesis, it was expected that greenness in the vegetation would increase in the area after a rain event; however, it will decline rapidly due to the loss of moisture caused by ground heating at geothermal hotspots compared to the non-geothermal hotspots. The observed results around rain events, however did not show any sizeable response in the temporal profiles of vegetation dynamics to be considered characteristic of geothermal hotspots. There were some differences in the time series profiles of NDVI, with non-geothermal hotspots showing a dip after the rain event, but the values at geothermal hotspots usually remained more or less the same around the rain events with only a slight drop.

The DTW analysis was able to quantitatively analyze the dissimilarities in the temporal profiles of geothermal hotspots and non-geothermal hotspots, checking through with around 2000 combinations of different points. When comparing two geothermal hotspots, the differences in the DTW distances (normalized distance) were low, indicating that it is easier to align two temporal sequences of geothermal hotspots, and the time series curves were quite similar. These values were relatively lower when two non-geothermal hotspots were compared, which in one way shows that comparing other points in an active region can show different temporal responses, but comparing geothermal hotspots show similar temporal responses. Additionally, the normalized distance value when comparing a geothermal hotspot and a non-geothermal hotspot was even higher, particularly on points with sparse vegetation. These results were consistent around all the rain events as well although no direct inference could be interpreted relating to the size of the rainfall event.

The temporal dynamics around rain events were difficult to be considered for identifying a distinctive pattern for geothermal hotspots or non-geothermal hotspots due to the effect of cloud cover around rain events. The images were cloud masked and all data was filtered based on the NDVI values ($NDVI < 0.1$) for clouds to work around this limitation, though this reduced the number of samples around some rain events which were quite low to be regarded as characteristic of the geo-activity it represents. Therefore, less number of sample points could just be outliers in the data showing abrupt responses: e.g. patches of very sparse vegetation. Further through visual analysis of the scenes, it was observed that some time points were still contaminated by clouds, and hence were removed from the observations. Nonetheless, the DTW responses were still different providing an opportunity to explore this criteria further for differentiating geothermal hotspots. The DTW distances showed significant differences between the temporal profiles of non-geothermal hotspots and geothermal hotspots at sparse vegetation. Such results could be utilized for differentiating geothermal hotspots, though requires more detailed analysis with a large dataset to be considered a reliable solution.

These observations were mainly indicative that the vegetation cover tend to be lower at geothermal hotspots, however there are certainly bare soil patches in these regions which are not geothermal hotspots. In other words, not all low NDVI value locations could indicate a geothermal hotspot. However, the differences in the temporal profiles observed around rain events between geothermal hotspots and non-geothermal hotspots when quantitatively analysed using DTW gives more insight on whether geothermal hotspots and non-geothermal hotspots can be differentiated in sparsely vegetated areas. The significant differences in the DTW distances of pairs of geothermal hotspots as compared to pairs of non-geothermal hotspots and geothermal hotspot with a non-geothermal hotspot in sparsely vegetated areas could be used to further isolate the geothermal hotspots. This requires further analysis with higher number of samples, and tests in different regions with geothermal active regions.

6.1. Applications of the results for potentially identifying geothermal hotspots

For the Olkaria region, the thermal analysis was carried out as part of the GeoHot project. The thermal study was performed using two different methods, and the maps prepared (Figure 27) were used to identify the ground control points for which the ground data exploration was done.

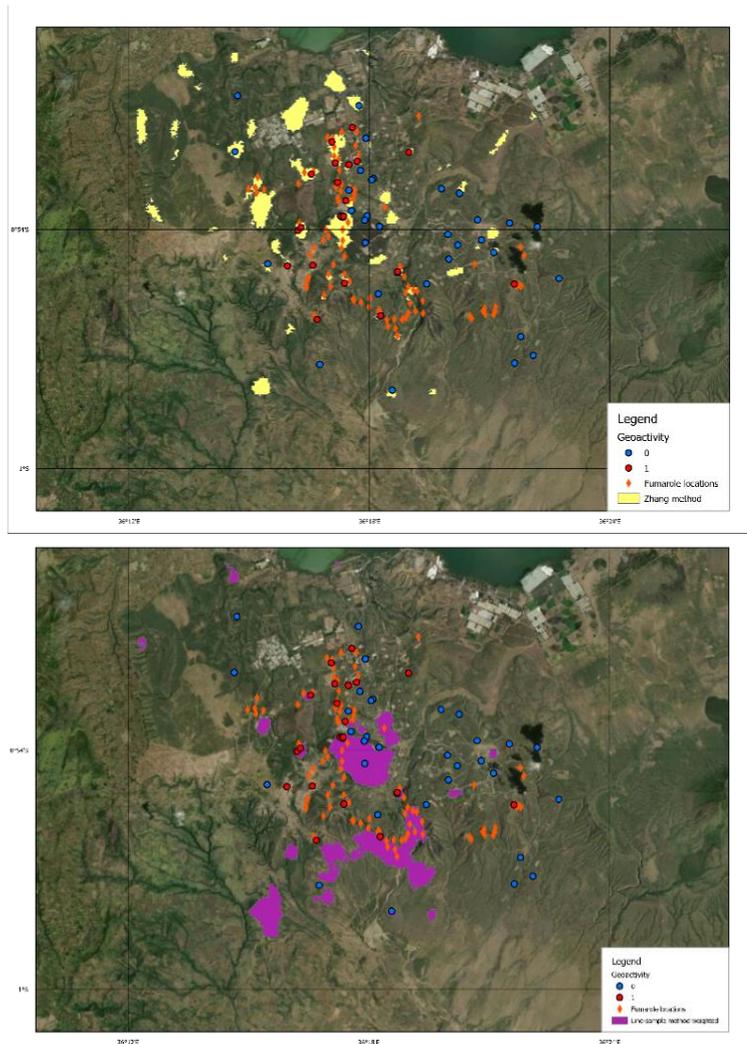


Figure 27: Thermal Infrared study to find locations for ground data exploration of geothermal hotspots (GeoHot Project of Faculty ITC)

These points were fairly accurate in determining the geothermal hotspots, with overall accuracy of about 80% (Figure 28)

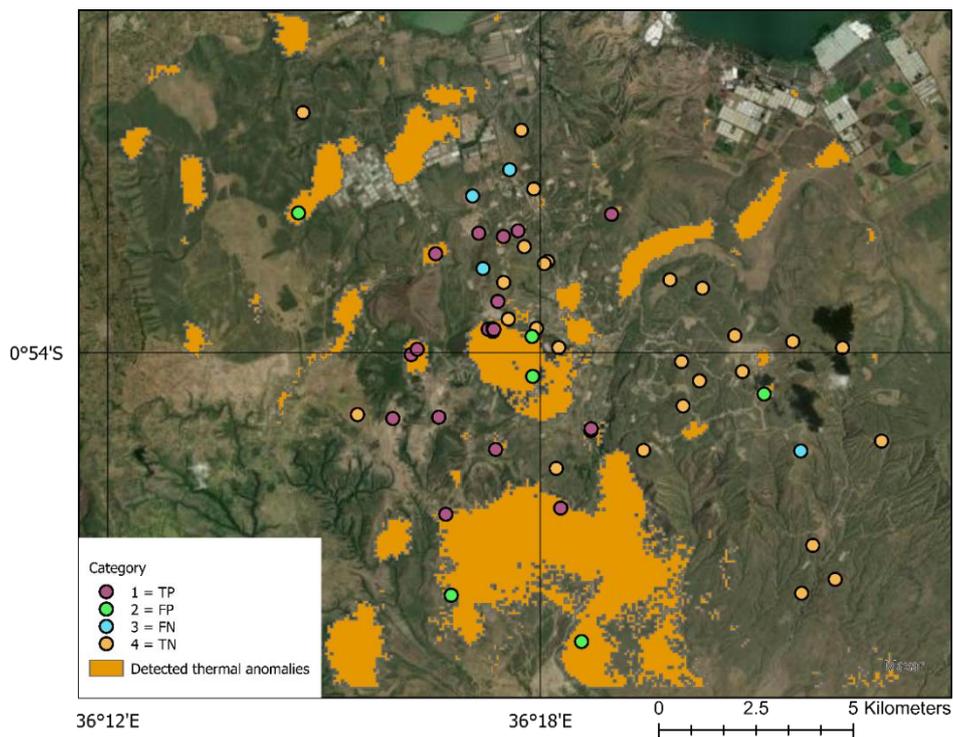


Figure 28: Confusion Matrix assessment of the ground control points

TP – True Positive, FP – False Positive, FN – False Negative, TN – True Negative

In this study, we aimed to analyse the use of vegetation dynamics for identifying these geothermal hotspots. While that didn't show any considerable differences, one of the evident observations, though elementary, was that the geothermal hotspots consistently had NDVI values below a certain threshold of 0.25. Applying this threshold in a geothermally active region can narrow down the exploration zones. Figure 29 shows the NDVI threshold zones, where the brown locations indicate zones where the hotspots could be present.

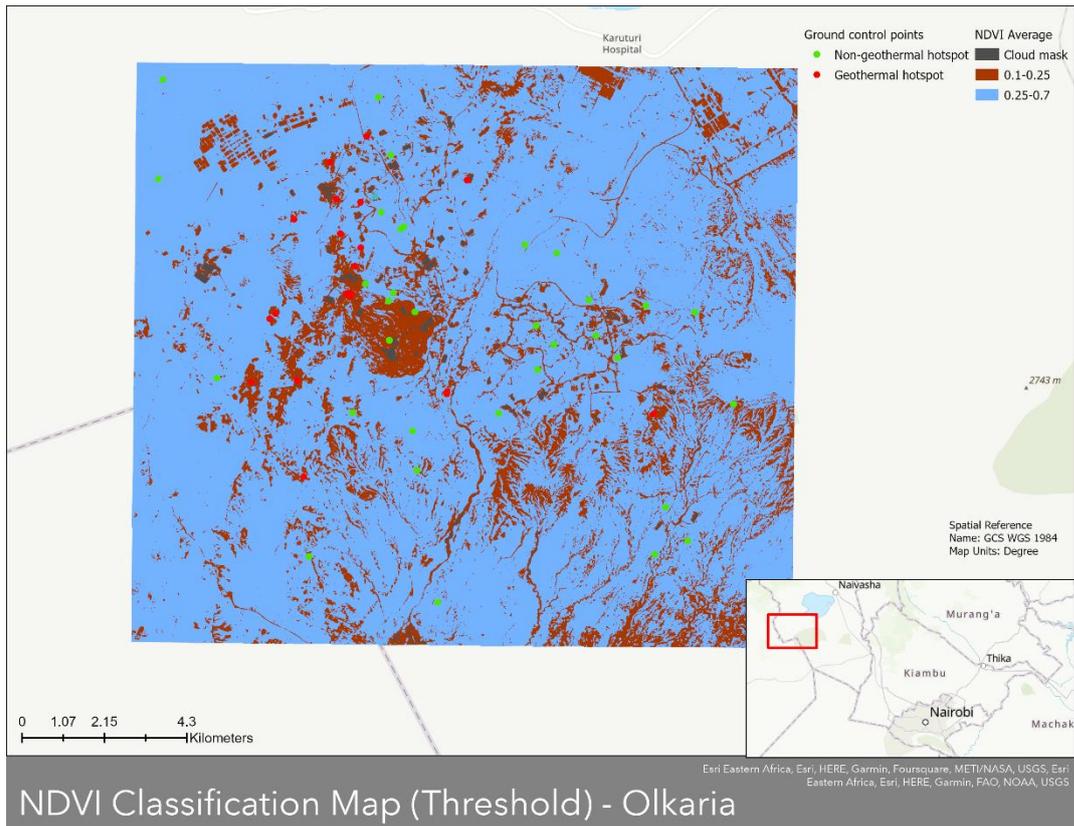


Figure 29: NDVI threshold map for defining exploration zones for geothermal hotspots in a geothermally active region

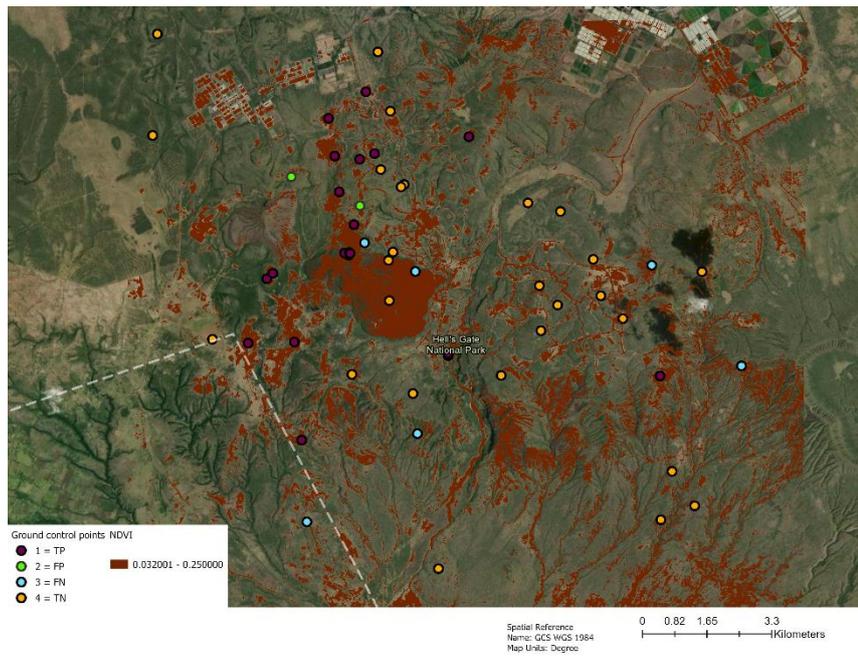


Figure 30: Classification accuracy of the ground control points using NDVI threshold

TP – True Positive, FP – False Positive, FN – False Negative, TN – True Negative

The points were classified with the accuracy of 86% using the NDVI threshold (Figure 30). Comparing the results from the two classifications, Out of the 23 hotspots, 21 (True positives) were accurately in the low NDVI zone as compared to 20 accurately identified with the land surface temperature study. Similarly, out of the 33 non-geothermal hotspots, only five were incorrectly in the low NDVI zone as compared to 6 false positives using the land surface temperature study.

In view of these observations, there are a few points to consider when utilizing vegetation as a parameter for identifying potential geothermal hotspots. In a geothermally active region, the temperature study is efficient for isolating zones where geothermal hotspots could be spotted. However, other surface characteristics, including vegetation dynamics, could be studied to further improve the results before ground exploration. While vegetation cannot be exclusively used to find geothermal hotspots, it can certainly be used as an additional parameter along with the thermal study before the ground truth surveys are carried out. Furthermore, using DTW to classify these points in low vegetation zones further employing machine learning methods like K-NN could provide more specific points and improve the accuracy, potentially saving the effort in ground truth survey.

7. CONCLUSIONS

The use of remote sensing for identifying potential locations of geothermal hotspots through the study of the surface characteristics of a geothermally active zone holds a lot of promise. Its applications in studying the soil properties, minerals, earth surface deformations, etc., and how it can be linked to the subsurface have been explored and utilized for surveying these areas (van der Meer et al., 2014). This study focused on understanding the effect of geothermal heat on vegetation dynamics, possibly assessing its use in identifying potential locations of geothermal hotspots in a geothermally active region.

The time series analysis of these points using the Sentinel-2 optical imagery, initially over the time period of 2 years, showed the differences in the reflectance values of various bands and the values of the vegetation indices. For example, higher differences in the reflectance range of the red band in geothermal hotspots and non-geothermal hotspots showed the presence of alteration minerals in the soil at geothermal hotspots. NDVI showed differences in the range for the entire two years, using which a threshold in the NDVI values was identified for geothermal hotspots.

The research question also focused on examining the presence of a characteristic response at geothermal hotspots, by observing the shapes of the time series curves around rain events. In general, a small dip in the NDVI was observed, for non-geothermal hotspots after a rain event, which eventually recovered to its original state. There were some limitations related to this observation due to the lack of completely cloud free images and hence all sample points could not be used to determine how the climate change over time. There was not much change observed at geothermal hotspots, and hence a characteristic response related to the hypothesis of the study could not be observed. Furthermore, with overall vegetation being scarce, the results could not be directly linked to be vegetation dynamics and was affected by soil dynamics.

These differences could not be indeed verified using only visual analysis, hence DTW was used. DTW is a dissimilarity test that helps identify how dissimilar two time series curves are by checking the shape of the curves combined with the differences in absolute values. This assessment showed that comparing two geothermal hotspots showed least dissimilarity as compared to when compared with non-geothermal hotspots.

These results suggest that timeseries of vegetation dynamics especially in areas with sparse vegetation, could provide information on the presence of a geothermal hotspot. This have a higher relation with the threshold of the NDVI values rather than the vegetation dynamics. This could be further assessed by testing the theory in a different region or for a longer time period to identify rain events around which higher percentage of cloud free data is available. The results from the DTW analysis requires more assessment by employing its use for automatic classification of geothermally active zones using machine learning algorithms such as K-NN, which was out of scope for this study. Additionally, these results are based on an independent dataset, and requires further verification with larger datasets in different environments.

The use of vegetation dynamics for identifying potential locations of geothermal hotspots shows limited use when utilized as an individual parameter. After assessing the thermal responses through remote sensing, NDVI values can be used to survey the region, while the vegetation dynamics only show weak indications of such differences for identifying potential geothermal hotspots.

Hence, it has limited use as a parameter to narrow down the zones for identifying geothermal hotspots in geothermally active regions and require further assessment with larger datasets.

LIST OF REFERENCES

- Abubakar, A. J., Hashim, M., & Pour, A. B. (2018). Identification of hydrothermal alteration minerals associated with geothermal system using ASTER and Hyperion satellite data: a case study from Yankari Park, NE Nigeria. *Https://Doi.Org/10.1080/10106049.2017.1421716*, 34(6), 597–625. <https://doi.org/10.1080/10106049.2017.1421716>
- Abuzied, S. M., Kaiser, M. F., Shendi, E. A. H., & Abdel-Fattah, M. I. (2020). Multi-criteria decision support for geothermal resources exploration based on remote sensing, GIS and geophysical techniques along the Gulf of Suez coastal area, Egypt. *Geothermics*, 88, 101893. <https://doi.org/10.1016/J.GEOTHERMICS.2020.101893>
- Bemdt, D. J., & Clifford, J. (1994). *Using Dynamic Time Warping to FindPatterns in Time Series*. www.aaii.org
- Bisht, A., Garg, N., Ryait, H. S., & Kumar, A. (2016). Comparative Analysis of DTW based Outlier Segregation Algorithms for Wrist Pulse Analysis. *Indian Journal of Science and Technology*, 9(47). <https://doi.org/10.17485/ijst/2016/v9i47/101746>
- Camberlin, P., & Wairoto, J. G. (1997). Theoretical and Applied Climatology Intraseasonal Wind Anomalies Related to Wet and Dry Spells During the “long” and “short” Rainy Seasons in Kenya. *Theor. Appl. Climatol*, 58, 57–69.
- Chen, L., Msigwa, G., Yang, M., Osman, A. I., Fawzy, S., Rooney, D. W., & Yap, P. S. (2022). Strategies to achieve a carbon neutral society: a review. *Environmental Chemistry Letters* 2022 20:4, 20(4), 2277–2310. <https://doi.org/10.1007/S10311-022-01435-8>
- Coluzzi, R., Imbrenda, V., Lanfredi, M., & Simoniello, T. (2018). A first assessment of the Sentinel-2 Level 1-C cloud mask product to support informed surface analyses. *Remote Sensing of Environment*, 217, 426–443. <https://doi.org/10.1016/J.RSE.2018.08.009>
- Elmarsdóttir, Á., Vilmundardóttir, O. K., & Magnússon, S. H. (2015). Vegetation of High-temperature Geothermal Areas in Iceland. *Proceedings World Geothermal Congress*, 19–25.
- European Space Agency. (n.d.). *Level-1C Cloud Masks - Sentinel-2 MSI Technical Guide - Sentinel Online - Sentinel Online*. Retrieved May 27, 2023, from <https://sentinel.esa.int/web/sentinel/technical-guides/sentinel-2-msi/level-1c/cloud-masks>
- Fahil, A. S., Ghoneim, E., Noweir, M. A., & Masoud, A. (2020). Integration of Well Logging and Remote Sensing Data for Detecting Potential Geothermal Sites along the Gulf of Suez, Egypt. *Resources* 2020, Vol. 9, Page 109, 9(9), 109. <https://doi.org/10.3390/RESOURCES9090109>
- Fraga, H., Amraoui, M., Malheiro, A. C., Moutinho-Pereira, J., Eiras-Dias, J., Silvestre, J., & Santos, J. A. (2014). Examining the relationship between the Enhanced Vegetation Index and grapevine phenology. *European Journal of Remote Sensing*, 47(1), 753–771. <https://doi.org/10.5721/EUJRS20144743>
- Freski, Y. R., Hecker, C., van der Meijde, M., & Setianto, A. (2021). The effects of alteration degree, moisture and temperature on laser return intensity for mapping geothermal manifestations. *Geothermics*, 97. <https://doi.org/10.1016/J.GEOTHERMICS.2021.102250>
- Frutuoso, R., Lima, A., & Teodoro, A. C. (2021). Application of remote sensing data in gold exploration: targeting hydrothermal alteration using Landsat 8 imagery in northern Portugal. *Arabian Journal of Geosciences*, 14(6), 1–18. <https://doi.org/10.1007/S12517-021-06786-0/FIGURES/11>

- Gkousis, S., Welkenhuysen, K., & Compernelle, T. (2022). Deep geothermal energy extraction, a review on environmental hotspots with focus on geo-technical site conditions. *Renewable and Sustainable Energy Reviews*, 162. <https://doi.org/10.1016/J.RSER.2022.112430>
- Güney, T. (2019). Renewable energy, non-renewable energy and sustainable development. *International Journal of Sustainable Development and World Ecology*, 26(5), 389–397. <https://doi.org/10.1080/13504509.2019.1595214>
- Gustavsson, L., Haus, S., Lundblad, M., Lundström, A., Ortiz, C. A., Sathre, R., Truong, N. Le, & Wikberg, P. E. (2017). Climate change effects of forestry and substitution of carbon-intensive materials and fossil fuels. *Renewable and Sustainable Energy Reviews*, 67, 612–624. <https://doi.org/10.1016/J.RSER.2016.09.056>
- Hecker, C., Hewson, R., Setianto, A., Saepuloh, A., & Van Der Meer, F. D. (2017). Multi-source remote sensing data analysis for geothermal targeting on Flores island. *Proceedings The 5th Indonesia International Geothermal Convention & Exhibition (IIGCE) 2017*.
- Hegazi, E. H. ; Samak, A. A. ; Yang, L. ; Huang, R. ; Huang, J., Alexakis, D., Niedbala, G., Hegazi, E. H., Samak, A. A., Yang, L., Huang, R., & Huang, J. (2023). Prediction of Soil Moisture Content from Sentinel-2 Images Using Convolutional Neural Network (CNN). *Agronomy* 2023, Vol. 13, Page 656, 13(3), 656. <https://doi.org/10.3390/AGRONOMY13030656>
- Henrich, V., Krauss, G., Götze, C., Sandow, C. (n.d.). *IDB - Credits*. Retrieved April 19, 2023, from <https://www.indexdatabase.de/info/credits.php>
- Huete, A. R., Liu, H. Q., Batchily, K., & Van Leeuwen, W. (1997). A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment*, 59(3), 440–451. [https://doi.org/10.1016/S0034-4257\(96\)00112-5](https://doi.org/10.1016/S0034-4257(96)00112-5)
- Kamau, M., Hecker, C., & Lievens, C. (2020). Use of Short-Wave Infrared Reflectance (SWIR) Spectroscopy to Characterize Hydrothermal alteration minerals in Olkaria Geothermal System, Kenya. *PROCEEDINGS*, 15.
- Koide, R., Lettenmeier, M., Akenji, L., Toivio, V., Amellina, A., Khodke, A., Watabe, A., & Kojima, S. (2021). Lifestyle carbon footprints and changes in lifestyles to limit global warming to 1.5 °C, and ways forward for related research. *Sustainability Science*, 16(6), 2087–2099. <https://doi.org/10.1007/s11625-021-01018-6>
- Kriegler, F., Malila, W., Nalepka, R., & Richardson, W. (1969). *Preprocessing Transformations and Their Effects on Multispectral Recognition*.
- Leibrand, A., Sadoff, N., Maslak, T., & Thomas, A. (2019). Using Earth Observations to Help Developing Countries Improve Access to Reliable, Sustainable, and Modern Energy. *Frontiers in Environmental Science*, 7, 123. <https://doi.org/10.3389/fenvs.2019.00123/full>
- Lu, B., Xu, S., Stuber, J., & Edgar, T. F. (2016). Constrained selective dynamic time warping of trajectories in three dimensional batch data. *Chemometrics and Intelligent Laboratory Systems*, 159, 138–150. <https://doi.org/10.1016/J.CHEMOLAB.2016.10.005>
- Lund, J. W., & Toth, A. N. (2021). Direct utilization of geothermal energy 2020 worldwide review. *Geothermics*, 90. <https://doi.org/10.1016/J.GEOTHERMICS.2020.101915>
- Majumdar, D., & Devi, A. (2021). Oilfield geothermal resources of the Upper Assam Petroliferous Basin, NE India. *Energy Geoscience*, 2(4), 246–253. <https://doi.org/10.1016/J.ENGEOS.2021.07.002>
- Mary, W. M., Moses, K. G., David, N. K., & Nicholas. (2017). Low cost geothermal energy

indicators and exploration methods in Kenya. *Journal of Geography and Regional Planning*, 10(9), 254–265. <https://doi.org/10.5897/JGRP2017.0643>

- McFeeters, S. K. (2007). The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. <https://doi.org/10.1080/01431169608948714>, 17(7), 1425–1432. <https://doi.org/10.1080/01431169608948714>
- Mwaura, D., & Kada, M. (2017). Developing a web-based spatial decision support system for geothermal exploration at the Olkaria geothermal field. <http://dx.doi.org/10.1080/17538947.2017.1284909>, 10(11), 1118–1145. <https://doi.org/10.1080/17538947.2017.1284909>
- NASA/JPL-Caltech. (n.d.). *Welcome to ECOSTRESS — ECOSTRESS*. Retrieved May 17, 2023, from <https://ecostress.jpl.nasa.gov/>
- Ohana-Levi, N., Knipper, K., Kustas, W. P., Anderson, M. C., Netzer, Y., Gao, F., Alsina, M. del M., Sanchez, L. A., & Karnieli, A. (2020). Using Satellite Thermal-Based Evapotranspiration Time Series for Defining Management Zones and Spatial Association to Local Attributes in a Vineyard. *Remote Sensing* 2020, Vol. 12, Page 2436, 12(15), 2436. <https://doi.org/10.3390/RS12152436>
- Omenda, P. A. (1998). The geology and structural controls of the Olkaria geothermal system, Kenya. *Geothermics*, 27(1), 55–74. [https://doi.org/10.1016/S0375-6505\(97\)00028-X](https://doi.org/10.1016/S0375-6505(97)00028-X)
- Petitjean, F., Inglada, J., & Gançarski, P. (2012). Satellite image time series analysis under time warping. *IEEE Transactions on Geoscience and Remote Sensing*, 50(8), 3081–3095. <https://doi.org/10.1109/TGRS.2011.2179050>
- Saepuloh, A., Dewi, W. C., Harto, A. B., & Agustan. (2021). Occurrence of Geothermal Features Based on Surface Roughness and Geobotanical Analyses derived by ALOS-2 PALSAR-2 and Sentinel-2 Images. *Proceedings - 2021 7th Asia-Pacific Conference on Synthetic Aperture Radar, APSAR 2021*. <https://doi.org/10.1109/APSAR52370.2021.9688361>
- Sakoe, H., & Chiba, S. (1978). Dynamic Programming Algorithm Optimization for Spoken Word Recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1), 43–49. <https://doi.org/10.1109/TASSP.1978.1163055>
- Soltani, M., Moradi Kashkooli, F., Souri, M., Rafiei, B., Jabarifar, M., Gharali, K., & Nathwani, J. S. (2021). Environmental, economic, and social impacts of geothermal energy systems. *Renewable and Sustainable Energy Reviews*, 140. <https://doi.org/10.1016/J.RSER.2021.110750>
- Song, X., Wang, G., Shi, Y., Li, R., Xu, Z., Zheng, R., Wang, Y., & Li, J. (2018). Numerical analysis of heat extraction performance of a deep coaxial borehole heat exchanger geothermal system. *Energy*, 164, 1298–1310. <https://doi.org/10.1016/J.ENERGY.2018.08.056>
- Tomasella, J., Silva Pinto Vieira, R. M., Barbosa, A. A., Rodriguez, D. A., de Oliveira Santana, M., & Sestini, M. F. (2018). Desertification trends in the Northeast of Brazil over the period 2000–2016. *International Journal of Applied Earth Observation and Geoinformation*, 73, 197–206. <https://doi.org/10.1016/J.JAG.2018.06.012>
- van der Meer, F., Hecker, C., van Ruitenbeek, F., van der Werff, H., de Wijkerslooth, C., & Wechsler, C. (2014). Geologic remote sensing for geothermal exploration: A review. *International Journal of Applied Earth Observation and Geoinformation*, 33(1), 255–269. <https://doi.org/10.1016/J.JAG.2014.05.007>
- van der Zwaan, B., & Dalla Longa, F. (2019). Integrated assessment projections for global geothermal energy use. *Geothermics*, 82, 203–211. <https://doi.org/10.1016/J.GEOTHERMICS.2019.06.008>

- Virnodkar, S. S., Pachghare, V. K., Patil, V. C., & Jha, S. K. (2020). Application of Machine Learning on Remote Sensing Data for Sugarcane Crop Classification: A Review. *Lecture Notes in Networks and Systems*, 93, 539–555. https://doi.org/10.1007/978-981-15-0630-7_55/FIGURES/3
- Wang, H., & Zhou, P. (2018). Assessing Global CO2 Emission Inequality From Consumption Perspective: An Index Decomposition Analysis. *Ecological Economics*, 154, 257–271. <https://doi.org/10.1016/J.ECOLECON.2018.08.008>
- Xofis, P., Spiliotis, J. A., Chatzigiouvanakis, S., & Chrysomalidou, A. S. (2022). *Long-Term Monitoring of Vegetation Dynamics in the Rhodopi Mountain Range National Park-Greece*. <https://doi.org/10.3390/f13030377>
- Zhu, L., Gong, H., Dai, Z., Xu, T., & Su, X. (2015). An integrated assessment of the impact of precipitation and groundwater on vegetation growth in arid and semiarid areas. *Environmental Earth Sciences*, 74(6), 5009–5021. <https://doi.org/10.1007/S12665-015-4513-5/FIGURES/7>