ASHESI UNIVERSITY

REMOTE SENSING OF QUALITY PARAMETERS OF MAJOR WATER BODIES IN GHANA USING SATELLITE IMAGERY.

CAPSTONE PROJECT

B.Sc. Computer Engineering

Yaw Owusu Brown

2022

ASHESI UNIVERSITY

REMOTE SENSING OF QUALITY PARAMETERS OF MAJOR WATER BODIES IN GHANA USING SATELLITE IMAGERY.

CAPSTONE PROJECT

Capstone Project submitted to the Department of Engineering, Ashesi University College in partial fulfilment of the requirements for the award of Bachelor of Science degree in Computer Engineering.

Yaw Owusu Brown

2022

DECLARATION

I hereby declare that this capstone is the result of my own original work and that no part of it has been presented for another degree at this university or elsewhere.

Candidate's Signature:

………………………………………………………………………………………………

Candidate's Name:

………………………………………………………………………………………………

Date:

………………………………………………………………………………………………

I hereby declare that preparation and presentation of this capstone were supervised in accordance with the guidelines on supervision of capstone laid down by Ashesi University College.

Supervisor's Signature:

………………………………………………………………………………………………

Supervisor's Name:

………………………………………………………………………………………………

Date:

………………………………………………………………………………………………

Acknowledgements

To my family, whose words of encouragement and support helped me undertake this project, I am extremely grateful. Special gratitude goes to God for bringing me this far. This project wouldn't have been possible without the support of my supervisor, Kofi Adu-Labi. His words of encouragement and technical guidance were priceless to the execution of the project.

Abstract

The concept of remote sensing to assess environmental conditions is a relatively new concept. Historically, information about the environment has been attained through physical investigation and more recently through sensors in IoT (Internet of Things) systems. Currently, in Ghana, there are a lot of activities such as illegal mining that tend to pollute water bodies. A significant proportion of Ghanaians depend on surface water for their domestic and commercial needs and so these water bodies must be protected and monitored. This project seeks to capitalize on this concept of remote sensing by monitoring of quality parameters of major water bodies in Ghana as a more cost-effective and sustainable approach.

Table of Content

DECLARATION i
Acknowledgements ii
Abstract iii
List of Tables vi
List of Figures vii
Chapter 1: Introduction 1
1.1 Background 1
1.2 Problem Definition 2
1.3 Objectives 2
1.4 Proposed Solution 2
1.5 Justification/Motivation 3
Chapter 2: Literature Review 5
2.1 Review of Existing Solutions 5
2.2 Related work 5
2.3 Remote Sensing Technologies 6
2.4 Image Requirements for Water Quality Analysis 9
2.5 Spectral Indices 10
Chapter 3: Design 11
3.1 Design Requirements 11
3.2 Design Choices 11
3.3 Block Diagram for Design 13
Chapter 4: Methodology 16
4.1 Required Setup for Implementation 16
4.2 First Approach with Sentinel-2 data. 16
4.2.1 NDVI Analysis of Sentinel-2 data for Detection of Harmful Algal Blooms 17
4.2.2 FAI Analysis of Sentinel-2 data for Detection of Harmful Algal Blooms 19
4.2.3 NDTI Analysis of Sentinel-2 data for Turbidity 22
4.2.4 GNDVI Analysis of Sentinel-2 data for Nitrogen Content 24
4.3 Second Approach with PlanetScope Data 26
4.3.1 NDVI Analysis of PlanetScope for Detection of Harmful Algal Blooms 27

4.3.2 NDTI Analysis of PlanetScope for Turbidity ... 29
4.3.3 GNDVI Analysis of PlanetScope for Nitrogen Content ... 31
4.4 Sending Information to Database ... 33
4.5 Alerting Algorithm ... 33
4.6 User Display Interface ... 34
4.7 Third Approach using a CNN with PlanetScope Data ... 35
Chapter 5: Interpretation of Results ... 39
5.1 Interpretation of NDVI results. ... 39
5.2 Interpretation of FAI results. ... 40
5.3 Interpretation of NDTI results ... 40
5.4 Interpretation of GNDVI results ... 41
Chapter 6: Conclusion and Future Works ... 43
6.1 Discussion ... 43
6.2 Limitations ... 43
6.3 Future Work ... 44
References ... 45
Appendix ... 48
Appendix A: Snippets of Python Code in Jupyter Notebooks ... 48
Appendix B: Snippets of MATLAB code ... 50
Appendix C: Source Code ... 51

List of Tables

Table 3. 1: Pugh chart for selection of remote sensing technology 12

Table 4. 1: Kpong Reservoir's NDVI Analysis results for the 3rd of January 2022……………18

Table 4. 2: Kpong Reservoir's FAI Analysis results for the 3rd of January 2022......................... 21

Table 4. 3: Kpong Reservoir's NDTI Analysis results for the 3rd of January 2022 23

Table 4. 4: Kpong Reservoir's GNDVI Analysis results for the 3rd of January 2022 25

Table 4. 5: Kpong Reservoir's NDVI Analysis results for the 4th of January 2022 27

Table 4. 6: Kpong Reservoir's NDTI Analysis results for the 4th of January 2022...................... 29

Table 4. 7: Kpong Reservoir's GNDVI Analysis results for the 4th of January 2022 31

Table 4. 8: Percentage of Water Cover across the Month 37

List of Figures

Figure 2.1: Image acquired from MODIS 7

Figure 2.2: Image acquired from Landsat-8 7

Figure 2.3: Image acquired from Sentinel-2 8

Figure 2.4: Image acquired from PlanetScope 9

Figure 3.1: Design Block Diagram 13

Figure 3.2: Entity Relationship Diagram 14

Figure 4.1: Kpong Reservoir's NDVI colourmap for the 3rd of January 2022 19

Figure 4.2: Kpong Reservoir's NDVI histogram for the 3rd of January 2022 19

Figure 4.3: Kpong Reservoir's FAI colourmap for the 3rd of January 2022 21

Figure 4.4: Kpong Reservoir's FAI histogram for the 3rd of January 2022 22

Figure 4.5: Kpong Reservoir's NDTI colourmap for the 3rd of January 2022 23

Figure 4.6: Kpong Reservoir's NDTI histogram for the 3rd of January 2022 24

Figure 4.7: Kpong Reservoir's GNDVI colourmap for the 3rd of January 2022 25

Figure 4. 8: Kpong Reservoir's GNDVI histogram for the 3rd of January 2022 26

Figure 4.9: Kpong Reservoir's NDVI colourmap for the 4th of January 2022 28

Figure 4. 10: Kpong Reservoir's NDVI histogram for the 4th of January 2022 28

Figure 4.11: Kpong Reservoir's NDTI colourmap for the 4th of January 2022 30

Figure 4.12: Kpong Reservoir's NDTI histogram for the 4th of January 2022 30

Figure 4.13: Kpong Reservoir's GNDVI colourmap for the 4th of January 2022 32

Figure 4.14: Kpong Reservoir's GNDVI histogram for the 4th of January 2022 32

Figure 4.15:Some tables in the MySQL database 33

Figure 4.16: Sample Email sent by Alerting Algorithm 34
Figure 4.17: User Display Interface. 35
Figure 4.18: Typical CNN Architecture 36
Figure 4.19: Segmented Image 37
Figure 4.20: Outcome of Pearson Correlation Analysis 38
Figure 5.1: Pixel Value Range for NDVI ……………..……………..………………………39
Figure 5.2: Pixel Value Range for FAI 40
Figure 5.3: Graph of TSS Concentration against NDTI values 41

Chapter 1: Introduction

1.1 Background

Water is a very important aspect of our lives and hence the very common adage "Water is life". It is of essence that we ensure that the major water bodies people in local communities are heavily reliant on are always kept at a high quality. Unfortunately, here in Ghana, many activities such as agriculture and illegal mining can potentially pollute these water bodies and so there is a need for regular monitoring to ensure that these water bodies are protected.

Currently, in Ghana, many people rely on surface water, and in some regions of the country, the water has significant water discolouration and contains dangerous minerals [1]. About 11% of the population, still drink from surface and other unsafe water sources [2]. The information above points to the fact that it is pivotal that we monitor and protect certain water bodies. A huge number of Ghanaians heavily rely on them, and irresponsible actions of people could easily lead to the pollution of these water bodies.

This can be done the traditional way by regularly collecting samples from the water body and running tests on them to ensure it is safe enough to use. This, however, is not so sustainable because the water body could be very far away. In addition, in this era of Covid-19, moving around regularly is not safe. A system could also be developed near the water body which would make use of sensors to provide information about the state of the water body. Although such a system could be very effective, maintenance issues would arise [3]. The system could get damaged and if the water body is in an area with unstable internet, it would affect the reliability.

1.2 Problem Definition

There is a need for a system that can be fully remote and provide consistent data on the quality of the water body. This can be achieved through satellite imaging. This approach will not require any movement or any maintenance, which would be ideal in this Covid Era. It can be used in conjunction with any of the above methods by an Environment Protection Agency (EPA) and so when it is not possible to make the journey to collect water samples or when the physical system at the water body develops some faults, the EPA can still be receiving information about the water body in question. This information received by an EPA or any governmental institution for that matter would enable them to react quickly to any activity that might be polluting the water body.

1.3 Objectives

- Develop an image processing algorithm to deduce data about water quality parameters
- Make use of remote sensing technology to attain images of water bodies of interest.
- Design a user display interface that would present the information in an organized manner.
- Develop an alerting algorithm to inform authorities when data attained is alarming.

1.4 Proposed Solution

The proposed solution is the development of a system that uses images of water bodies from remote sensing technologies, such as Modis and Landsat. The system deduces data on key water quality parameters that properly characterize the state of the water body and whether it is safe to be used. This data would be presented on a user interface where it would be accessed by personnel from some EPA and can be easily understood by the layman. The system would report information to the interface at least once every 5 days. This proposed system would allow for consistent and reliable monitoring of the water body. In addition, it is a sustainable approach that would require very little maintenance. It would enable for quick detection of problems with the water bodies and hence quick resolution of the problem. This would be done by comparing the water quality parameters to a set standard. In summation, the system would simply advance the protection of these water bodies and in essence protect the interest of people who rely so crucially on them.

1.5 Justification/Motivation

The motivation stems from the interest in the safety of the people residing in communities where they depend on surface water such as streams, rivers, and lakes. Usually, people in such communities are not educated and know very little about water quality standards and so just make do with whatever state the water is in. The authorities that can make a difference usually find out about these issues very late or they do not even find out at all. There is a need for this information to be brought to them in the most convenient manner possible so that they can act on it.

Another motivation is the problems that arise from the employment of the old approaches. The traditional method of collecting samples is not sustainable and it is also not a cost-effective approach. The improved method of having an in-situ monitoring system, which makes use of sensors and communication networks to acquire and report data respectively, would also require heavy maintenance and could get damaged. If the water body is in a location with an unstable internet connection, it would affect the reliability of the communication network. Both approaches could result in days where you don’t have any information about the water quality and hence the pivotal need for a reliable system that would provide you with the information regularly.

Chapter 2: Literature Review

2.1 Review of Existing Solutions

The employment of the already existing solutions which I have already touched on, being the traditional method of collecting water samples and the in-situ water quality monitoring system, all have their perks but also some undeniable disadvantages. The collection of water samples allows for more tests to be conducted on the water body as well as getting more detailed information. However, it would require regular visiting of the water body and the data retrieved could be skewed if the samples are not properly collected. The regular visits to the water body do not make this a cost-effective approach and the information could be unreliable if the samples are not properly collected. An in-situ water quality monitoring system would take out the regular visits to the water body because it is a remote approach. However, the issue of reliability would still come up because of the need for maintenance of the components of such a system. In addition, a system like this which is most likely to be an Internet of Things system could be disadvantaged in areas with poor connectivity.

2.2 Related work

The monitoring of Taihu water quality involved the use of high-resolution satellite imaging. In this approach, MODIS and Landsat data were used in conjunction and the algorithm developed based on the images attained was used to derive the concentration of nitrogen, concentration of phosphorus, conductivity, dissolved oxygen, and temperature [4]. In a description of an approach to monitor turbidity and suspended sediment concentration of coastal and inland waters using satellite data, there was the use of a quasi-analytical algorithm. This involves the use of the backscattering coefficient of water in the visible to near-infrared spectral regions from the satellite measured radiance and then converting it into nephelometric turbidity units [5]. These two approaches give perspective on some key decisions to be made in the execution of the project. Lastly, in the use of remote sensing satellites for water quality monitoring in the UAE, MODIS data is also employed but this time in conjunction with high-resolution data from DubaiSat-1[6]. This implementation of remote sensing has been employed to detect red tide also known as algal bloom and oil spills. For red tide detection, an algorithm is employed which uses the reflectance measured in the three red channels of MODIS to detect the presence and intensity of red tide events [6].

2.3 Remote Sensing Technologies

To do any form of remote monitoring, there is the need for remote sensing technology to provide the satellite data needed to acquire the information of interest. MODIS, which is the abbreviated version of Moderate-resolution Imaging Spectro-radiometer, is a very commonly used remote sensing technology. MODIS operates on both the Terra and Aqua spacecraft. Terra Satellite was launched on December 18, 1999, and the Aqua Satellite was launched on May 4th, 2002. MODIS has a temporal resolution of one to two days and the data acquired is at three spatial resolutions of 250metres, 500metres, and 1000metres per pixel. Its detectors measure 36 spectral bands. MODIS data is ideal for monitoring quickly changing environmental features because of its high temporal resolution but the extent to which analysis can be done on it is limited because of the not so good spatial resolution.

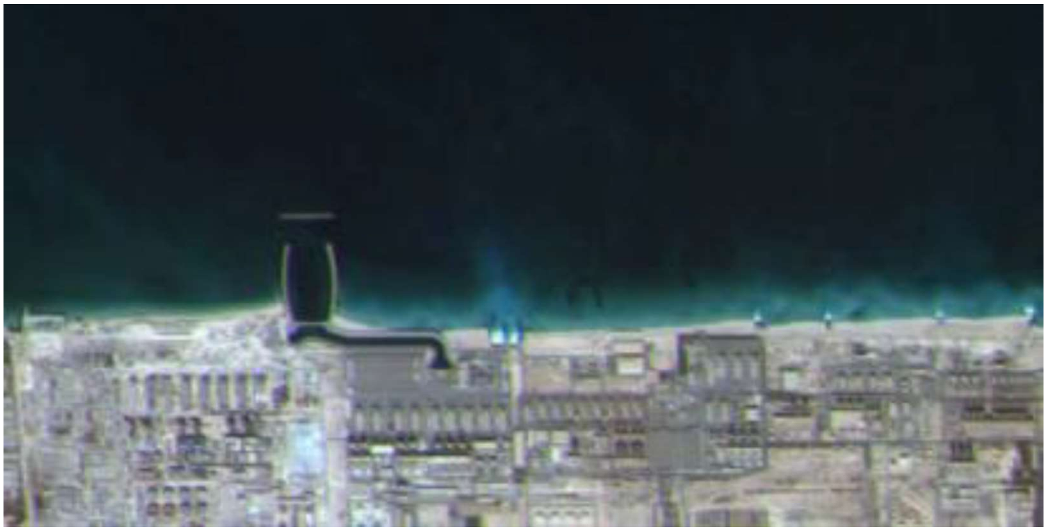


Figure 2.1: Image acquired from MODIS [6]

Landsat -8 was launched on an Atlas-V rocket from Vandenberg Air Force Base, California on February 11, 2013. It carries the Operational Land Imager and the Thermal Infrared Sensor instruments. Landsat 8 has a temporal resolution of 16 days and a spatial resolution of 30metres per pixel. Landsat 8 instruments data set includes 11 spectral bands. Landsat 8 data has a good spatial resolution which allows for more detailed analysis to be done but has an unimpressive temporal resolution.

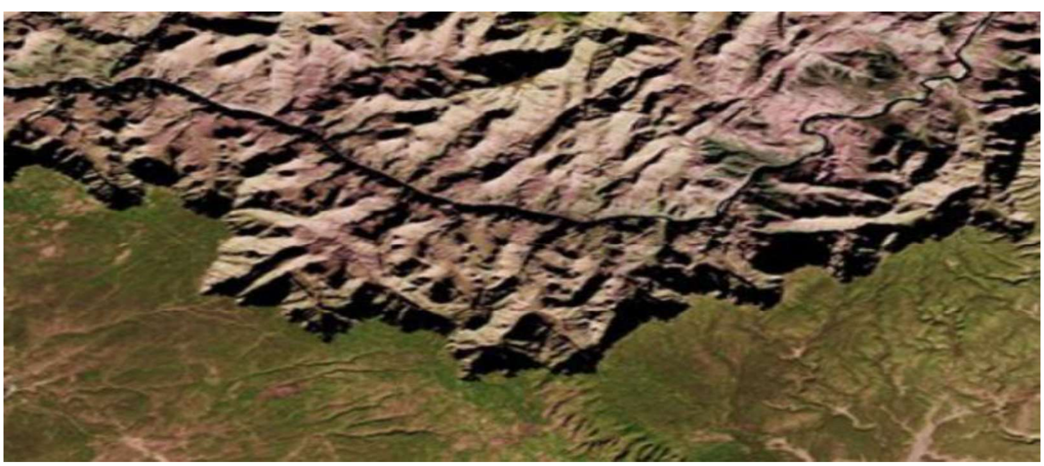


Figure 2.2: Image acquired from Landsat-8[10]

Sentinel-2 was launched as part of the European Commission's Copernicus program on June 23, 2015. It was designed to provide high-quality data and imagery. This satellite has an optoelectronic multispectral sensor, and this enables spatial resolution of as high as 10metres per pixel. Sentinel-2 has a temporal resolution of 5 days around the equator and every 2-3 days at middle latitudes. In addition, Sentinel-2 data is multispectral with 13 bands in the visible, near-infrared, and shortwave infrared spectrum. Its very high spatial resolution and revisit time make it an ideal option to be employed to study environmental challenges [7]. Consequently, Sentinel-2 data provides Global Monitoring for Environment and Security program services related to land management, agricultural production, and forestry, and monitoring of natural disasters and humanitarian operations [8].

Figure 2.3: Image acquired from Sentinel-2[7]

PlanetScope is also a source of earth observation data. It is operated by a private organization named Planet. It is a constellation of about 130 satellites that have the capacity of imaging the entire land surface of the Earth within a day. PlanetScope satellite constellation consists of multiple launches of Dove satellites. The images attained are of a spatial resolution of 3metres per pixel. Both the spatial and temporal resolution of PlanetScope data is highly impressive, however, it is not open-source data. This harms the ability of any remote-sensing activity to be continuous or sustainable.

Figure 2.4: Image acquired from PlanetScope [11]

2.4 Image Requirements for Water Quality Analysis

From any Earth Observation data, it is possible to retrieve information about a myriad of water quality parameters such as particle load, turbidity, presence of harmful algal blooms, and a host of others. However, to do good water quality retrieval there are some important things needed. The Earth Observation data should have the right radiometry. This means the VIS-NIR bands should resolve absorption and scattering of the water column. In addition, the NIR-SWIR bands for atmospheric correction should be present. Lastly, the radiometry has precise calibration[9]. Furthermore, the Earth Observation data should be characterized by frequent observations to deal with the issues such as cloud coverage. If there are multiple sets of data, the one with the least cloud coverage would be used and this would be more effective. In addition, there should be a spatial resolution of 100m or better for coastal areas and 30m or better for inland waters [9].

2.5 Spectral Indices

Spectral indices are combinations of the pixel values from two or more spectral bands in a multispectral image. Spectral indices are designed to highlight pixels showing the relative abundance or lack of a land-cover type of interest in an image [15]. Four indices would be used in this project namely the Normalized Difference Vegetation Index, Floating Algae Index, Normalized Difference Turbidity Index, and the Green Normalized Difference Vegetation Index. Normalized Difference Vegetation Index (NDVI) quantifies vegetation by measuring the difference between near-infrared (which vegetation strongly reflects) and red light (which vegetation absorbs) [16]. The floating Algae Index (FAI) is a simple ocean colour index used to detect floating algae. FAI is defined as the difference between reflectance at 859 nm (vegetation “red edge”) and a linear baseline between the red band (645 nm) and short-wave infrared band (1240 or 1640 nm) [17]. These two indices based on the characteristics can be used well in conjunction to determine the presence of harmful algal blooms. The Normalized Difference Turbidity Index (NDTI) is an index for qualitative analysis of turbidity [18]. Lastly, The Green Normalized Difference Vegetation Index (GNDVI) method is a vegetation index for estimating photosynthetic activity and is a commonly used vegetation index to determine water and nitrogen uptake into the plant canopy [19].

Chapter 3: Design

3.1 Design Requirements

The system that is to be designed in this project reports on the state of the water in regards to the following water quality parameters: Turbidity, Presence of Harmful Algal Blooms, and Nitrogen Content.

Below are the expected functionalities and requirements of the system

1. The data attained would be of high accuracy.
2. The data would be presented in a way that could be understood by the layman, on a user interface.
3. The data would be presented at least once every 5 days.
4. The data would be stored in a database for future references
5. Authorities would be informed when the data attained suggests a problem, through an alerting algorithm.

3.2 Design Choices

The idea behind this whole project is to bring information about these water bodies to the attention of the authorities in the quickest, simplest, most accurate, and most convenient way. Therefore, these are the things I look at when deciding on any aspect of the project. In essence, I consider all these factors when making any decisions. I compare all the alternatives and see which would help attain these criteria.

The decision of what remote sensing technology to use is a very crucial decision to be made in this project and so a PUGH chart was employed to determine the best one to use amongst the available options.

Table 3.1: Pugh chart for selection of remote sensing technology

Remote Sensing Technology	Cost	Frequency of images	Quality of images	Total
PlanetScope	**-20**	**+12**	**+14**	**+6**
MODIS	**0**	**+3**	**-4**	**-1**
Landsat-8	**0**	**0**	**+2**	**+2**
Sentinel-2	**0**	**+2**	**+5**	**+7**

Based on the above criteria, I have decided to use Sentinel-2 and PlanetScope Data remote sensing technology for the execution of the project.

3.3 Block Diagram for Design

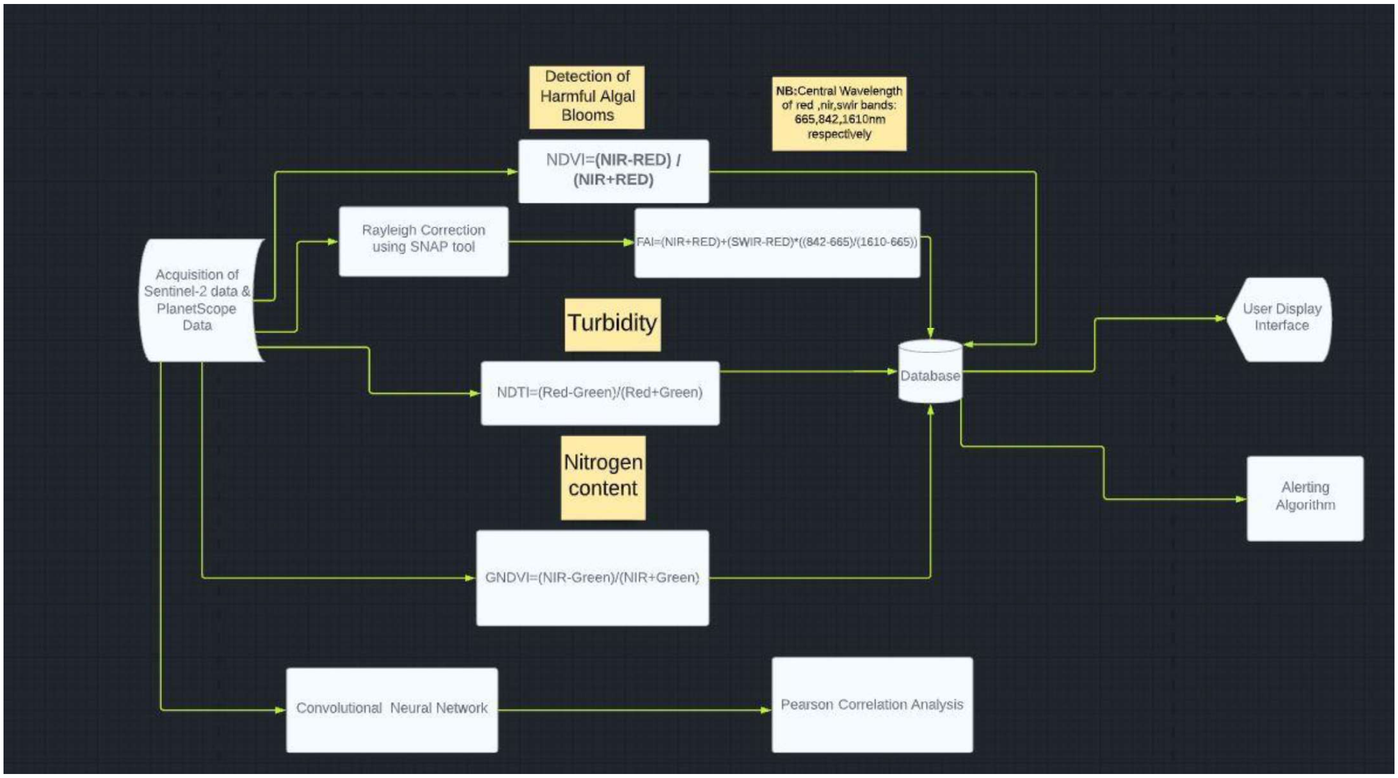


Figure 3.1: Design Block Diagram

SYSTEM FLOW:

1. Remote Sensing Technology is the genesis of the whole project. It is also a very crucial part of the project because the remote sensing technology used would influence the form of analysis that can be done and consequently the data that can be attained. The acquisition of Sentinel-2 data and PlanetScope Data is the first and most pivotal step.
2. Once the data is acquired, the Normalized Difference Vegetation Index (NDVI) and Floating Algae Index (FAI) can be computed as the indexes to determine the presence of harmful algal blooms in the water body.
3. The Normalized Difference Turbidity Index (NDTI) is computed with the data acquired and this index is used to describe the turbidity levels of the water body.

4. The Green Normalized Difference Turbidity Index (GNDVI) is also computed with the data acquired and this is the index that describes the nitrogen content in the water body.
5. This data deduced from the analysis is stored in a database.

<u>Database Design (Entity Relationship Diagram)</u>

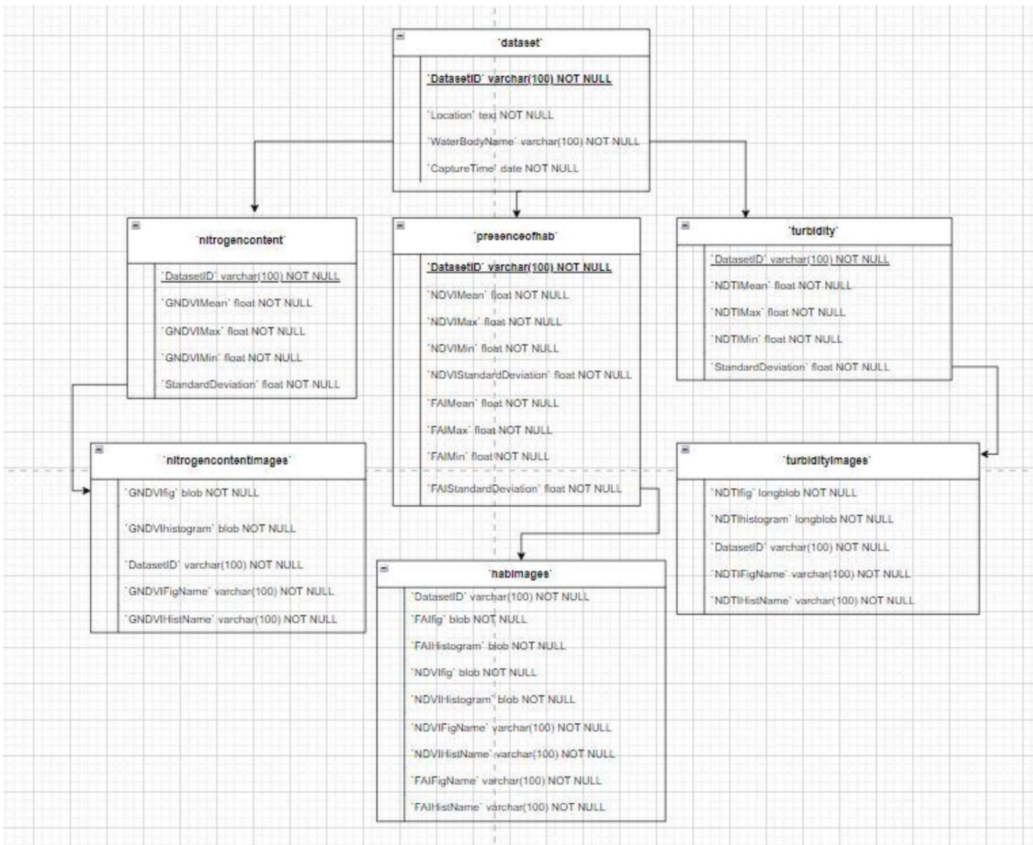


Figure 3.2: Entity Relationship Diagram

6. The data is displayed from a database onto a User Display Interface. User Display Interface would be executed using HTML, CSS, JavaScript, and PHP.
7. The alerting algorithm checks the mean values of the four indices and if the values indicate a problem sends an email to the authorities.
8. A convolutional neural network (CNN) is applied to a month's worth of data.

9. Pearson correlation analysis is applied to the data obtained from the CNN to see how it changes with respect to time.

Chapter 4: Methodology

4.1 Required Setup for Implementation

To begin executing this project, there were some technologies attained before commencing. Anaconda navigator was the first software and this has Jupyter notebook (which uses the Python programming language), in which the majority of the project was done. A virtual environment had to be created and activated for the required libraries to be downloaded and utilized. To prevent compatibility issues with the other parts of the anaconda program, a new environment was created where libraries such as rasterio could be downloaded. In addition, some accounts were needed to be set up to access and download the earth observation data. For the attainment of Sentinel-2 data, a user account had to be created on the Copernicus open access hub website. On the other hand, with regards to PlanetScope Data, there was a need for an application to their Education and Research program which is tailored for students doing research projects. Individuals accepted to this program can download up to 5,000 square kilometres of data per month for non-commercial purposes.

4.2 First Approach with Sentinel-2 data.

The Area of Interest, for this implementation, is the Kpong Reservoir, a water body located in the Eastern Region, which is drawn and used in the query for all the Sentinel-2 data. The focus for this implementation was the first 5 days in the year 2022 and this was also specified in the query. In addition, the processing level was also specified in the query for the data. The processing level selected was Level-2A: this level of processing represents surface reflectance data which is data that has undergone atmospheric correction. The atmospheric correction processing that the data has undergone means that the data is providing information about the surface and this makes it an ideal processing level to be selected for the query. The query was then executed and the sample of Earth Observation data that met the criteria described, which was data captured on the 3rd of January 2022, was displayed and then downloaded. The Sentinel-2 data came in the form of a folder that contained the different bands at the different spatial resolutions (10m,20m, and 60m per pixel) and some metadata in addition to other key information.

4.2.1 NDVI Analysis of Sentinel-2 data for Detection of Harmful Algal Blooms

The Red and Near-Infrared Bands like the other bands are stored in a JPEG 2000 file format. The rasterio library was used to open and read the bands. The bands are now mathematically combined using Equation 1 to attain the NDVI raster data.

$$\mathbf{NDVI}=\frac{NIR-RED}{NIR+RED} \qquad \textbf{(1)}$$

This is then stored in a TIF file format. The NDVI TIF file is then open and clipped because the initial Sentinel-2 data is of the Kpong Reservoir and its environs and so to get the NDVI data to be of just the Kpong Reservoir it needed to be clipped to fit that Area of Interest alone. After the NDVI TIF is clipped it is stored and then opened and read. The NumPy library was then employed to attain minimum, maximum, and mean NDVI values as well as the Standard deviation of the data. The outcome of this can be seen below in Table 4.1:

Table 4. 1: Kpong Reservoir's NDVI Analysis results for the 3rd of January 2022

Parameter	Value
MEAN NDVI	-0.016163939591295844
MINIMUM NDVI	-0.3991507430997877
MAXIMUM NDVI	0.6495185694635488
STANDARD DEVIATION	0.041831073692121384

Furthermore, after being read, it was visualized by functions in the matplotlib library. A colour bar was added to make the plot easier to interpret. The plot can be seen below in Figure 4.1:

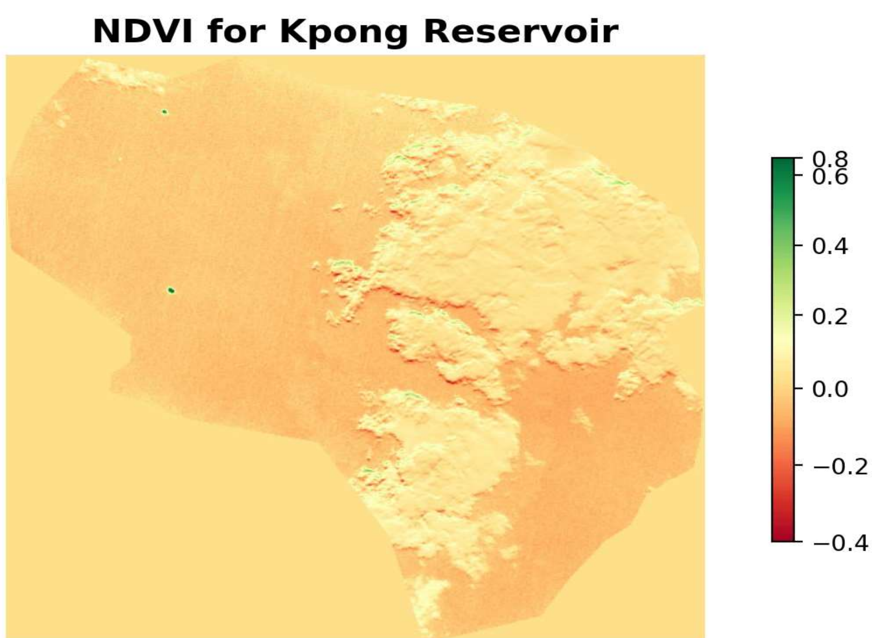


Figure 4.1: Kpong Reservoir's NDVI colourmap for the 3rd of January 2022

Lastly, the data is shown on a histogram which shows the Pixel distribution across the NDVI values. This inclusion was made as another way for the NDVI to be visualized. The histogram is seen below in Figure 4.2:

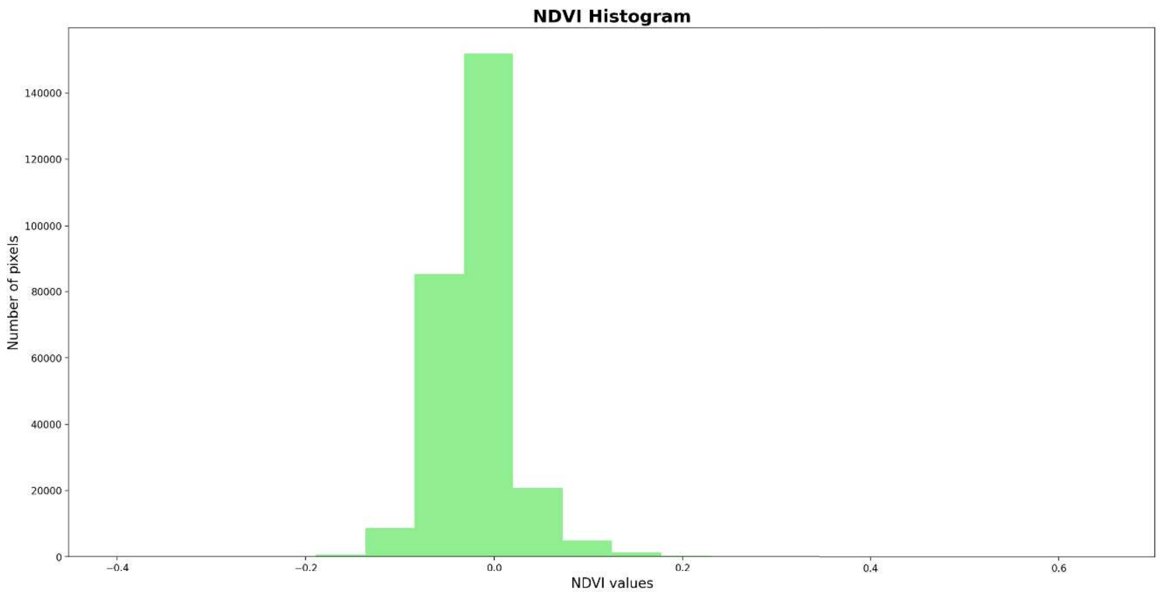


Figure 4.2: Kpong Reservoir's NDVI histogram for the 3rd of January 2022

4.2.2 FAI Analysis of Sentinel-2 data for Detection of Harmful Algal Blooms

The implementation of the FAI Analysis was unique because it required Sentinel-2 data of processing level 1C (top-of-atmosphere reflectance) that has undergone Rayleigh Correction. This was so because to compute the FAI, the Shortwave Infrared band is needed and it has a 20m per pixel spatial resolution. Hence, it became problematic to do any form of mathematical combination with the Red and Near-Infrared Band which have a 10m per pixel spatial resolution.

To address the issue, Sentinel-2 data of Level 1C processing was downloaded and Rayleigh correction was done using the SNAP software. In the parameters of the Rayleigh correction, a spatial resolution of 10m per pixel was specified for the Shortwave Infrared band.

The Rayleigh correction was done on the Red, Near-Infrared, and Shortwave Infrared bands, and the output bands were stored in TIF file formats. The rasterio library was used to open and read the bands. The bands were then mathematically combined using Equations 2 and 3 to attain the FAI raster data.

$$\mathbf{FAI = RRC(NIR) - R'RC(NIR)} \qquad \mathbf{(2)}$$

$$\mathbf{R'RC(NIR) = RRC(RED) + \frac{(RRC(SWIR) * (\lambda RED - \lambda NIR))}{(\lambda RED - \lambda SWIR)}} \qquad \mathbf{(3)}$$

λNIR=842nm
λSWIR= 1610nm
λRED= 665nm

Where RRC is the Rayleigh top-of-atmosphere reflectance and λband is the central wavelength (Red, Near-Infrared, and Shortwave Infrared)

After the FAI raster data was attained, an identical process (used in the NDVI analysis) was executed to attain minimum, maximum, and mean FAI values as well as the Standard deviation and also visualize the data in a colourmap plot and histogram. These can be seen below in Table 4.2, Figures 4.3 and 4.4:

Table 4. 2: Kpong Reservoir's FAI Analysis results for the 3rd of January 2022

Parameter	Value
MEAN FAI	-0.0037624047
MINIMUM FAI	-0.20629427
MAXIMUM FAI	0.17996715
STANDARD DEVIATION	0.014394339

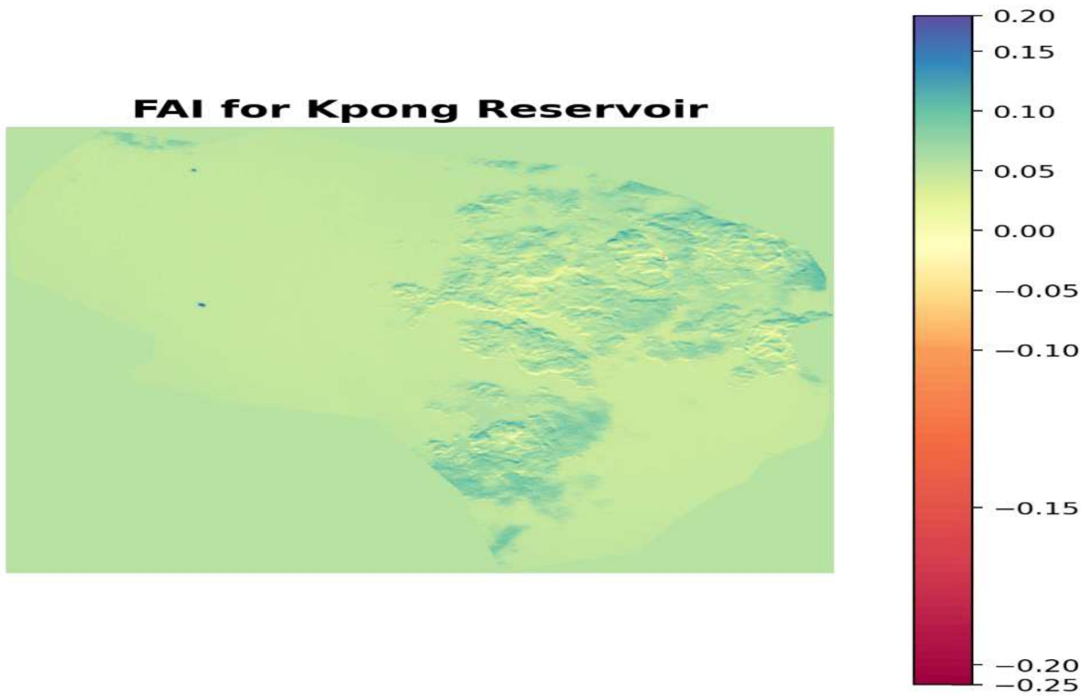


Figure 4.3: Kpong Reservoir's FAI colourmap for the 3rd of January 2022

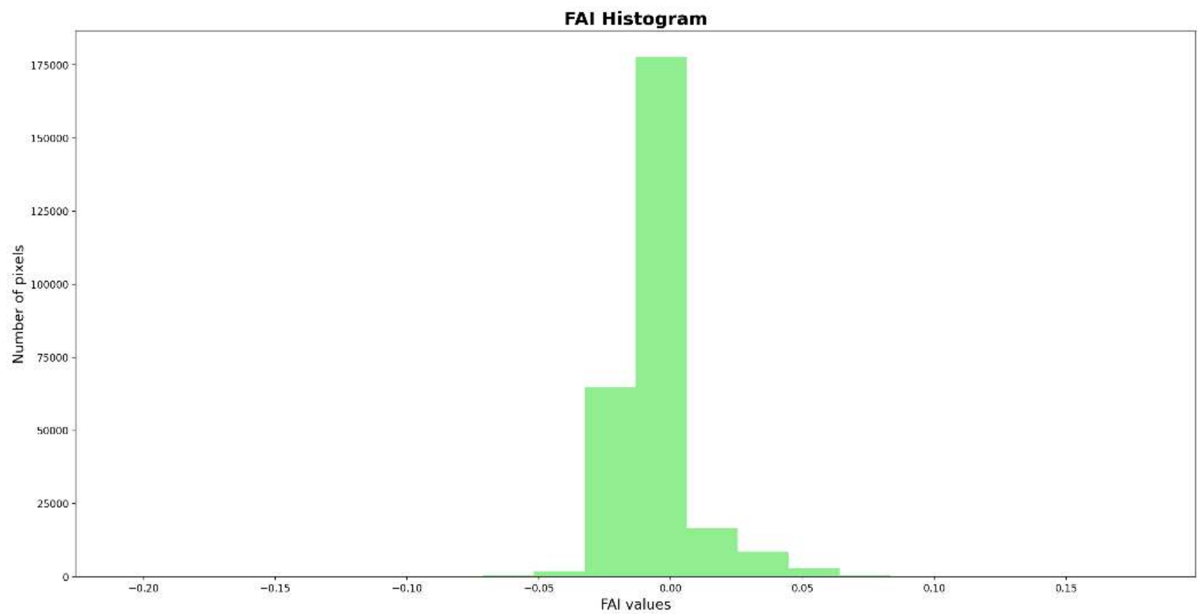


Figure 4.4: Kpong Reservoir's FAI histogram for the 3rd of January 2022

4.2.3 NDTI Analysis of Sentinel-2 data for Turbidity

The Red and Green bands were used in this analysis. Just like all the other bands they were stored in JPEG 2000 file format. These bands are mathematically combined in Equation 4 to attain the NDTI raster data.

$$\mathbf{NDTI} = \frac{\boldsymbol{RED - GREEN}}{\boldsymbol{RED + GREEN}} \qquad \textbf{(4)}$$

After the NDTI raster data was attained, this same process used in the NDVI analysis was used to attain minimum, maximum, and mean NDTI values as well as the Standard deviation. In addition, to visualize the data in a colourmap plot and histogram. These can be seen below in Table 4.3, Figures 4.5 and 4.6:

Table 4. 3: Kpong Reservoir’s NDTI Analysis results for the 3rd of January 2022

Parameter	Value
MEAN NDTI	-0.04293288341883818
MINIMUM NDTI	-0.3132295719844358
MAXIMUM NDTI	0.21522921522921523
STANDARD DEVIATION	0.050673314359819054

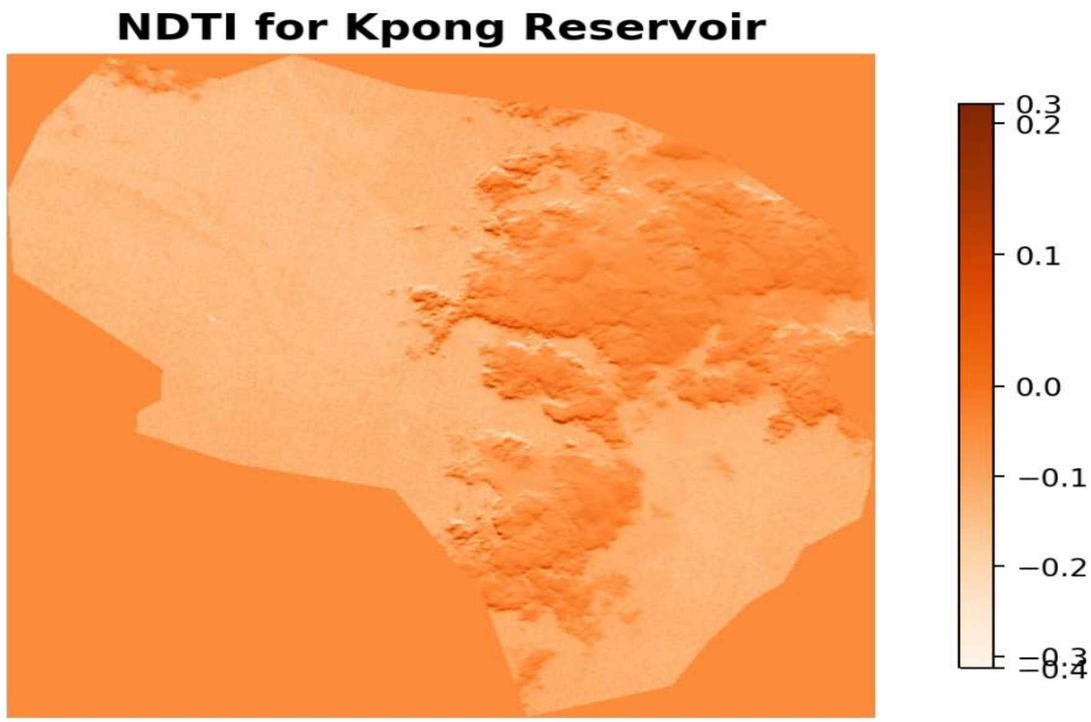


Figure 4.5: Kpong Reservoir's NDTI colourmap for the 3rd of January 2022

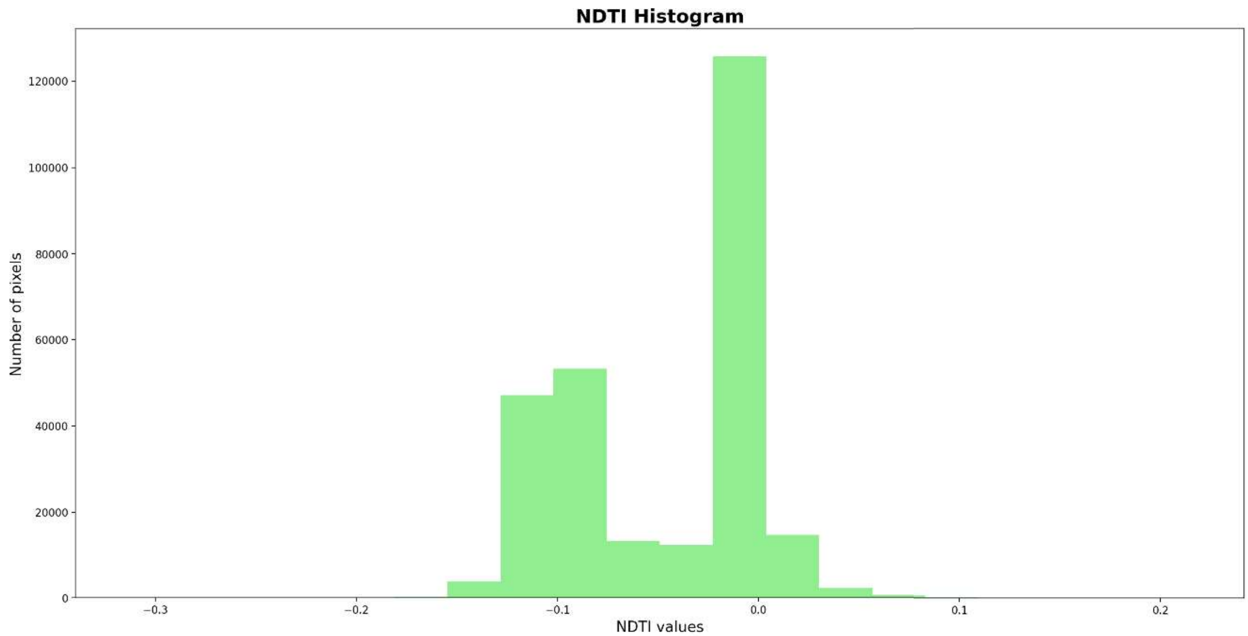


Figure 4.6: Kpong Reservoir's NDTI histogram for the 3rd of January 2022

4.2.4 GNDVI Analysis of Sentinel-2 data for Nitrogen Content

The Near-Infrared and Green bands were employed in this analysis. Similar to all the other bands they were stored in JPEG 2000 file format. These bands are mathematically combined in Equation 5 to attain the GNDVI raster data.

$$\mathbf{GNDVI} = \frac{\boldsymbol{NIR - GREEN}}{\boldsymbol{NIR + GREEN}} \quad \textbf{(5)}$$

After the GNDVI raster data was attained, the process used in the NDVI analysis was also used to attain minimum, maximum, and mean GNDVI values as well as the Standard deviation and also visualize the data in a colourmap plot and histogram. These can be seen below in Table 4.4, Figures 4.7 and 4.8:

Table 4. 4: Kpong Reservoir's GNDVI Analysis results for the 3rd of January 2022

Parameter	Value
MEAN GNDVI	-0.0587754876915526
MINIMUM GNDVI	-0.22939560439560439
MAXIMUM GNDVI	0.5288118306986231
STANDARD DEVIATION	0.07703910779153567

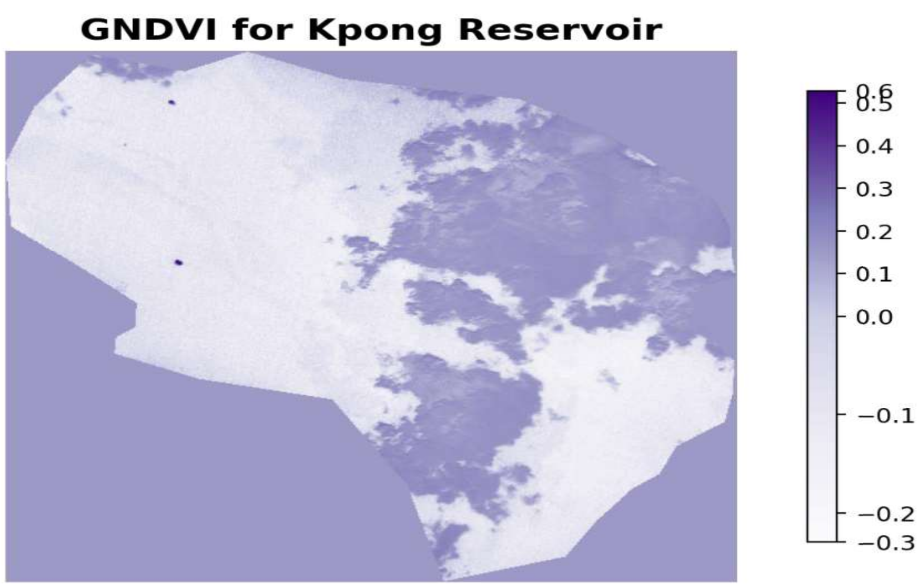


Figure 4.7: Kpong Reservoir's GNDVI colourmap for the 3rd of January 2022

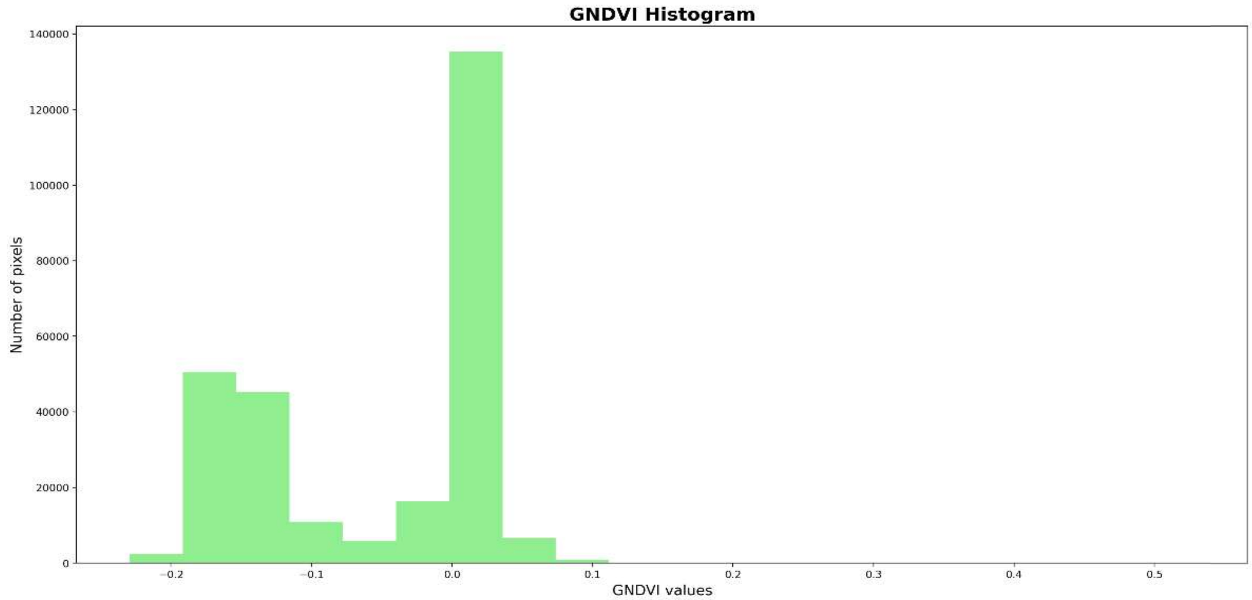


Figure 4. 8: Kpong Reservoir's GNDVI histogram for the 3rd of January 2022

4.3 Second Approach with PlanetScope Data

Unlike the Sentinel-2 data, PlanetScope data was attained straight from the website. This was done because there is a limit to the amount of data that can be downloaded and if it is attempted to download the data using an API, the earth observation data would include the Area of Interest (that would be demarcated) and its environs. This is problematic because some of the data downloaded won't be useful and it would prevent as much useful data from being downloaded since there is a limit. To curb this problem, the website was visited and the Area of Interest was demarcated and only data on this Area of Interest is downloaded. In addition, the date range was also specified and for this implementation, the data attained was captured on 4th January 2022. The product used was average surface reflectance and this is the equivalent of Sentinel-2 data of Level 2A processing in that it has also undergone an atmospheric correction and hence gives information about the surface. The Area of Interest for this implementation was also the Kpong Reservoir. The PlanetScope data came in the form of a TIF file which constitutes the Red, Green, Blue, and Near-Infrared bands all at 3m per pixel spatial resolution. Since the PlanetScope data does not include the Shortwave Infrared band, the FAI analysis could not be executed.

4.3.1 NDVI Analysis of PlanetScope for Detection of Harmful Algal Blooms

Rasterio library is used to open the data and the Red and Near-Infrared bands are read from it. The Red and Near-Infrared bands were mathematically combined using the NDVI formula (the same formula used in the first approach) to attain the NDVI raster data. This is stored in a TIF file. The NumPy library was used to attain minimum, maximum, and mean NDVI values as well as the Standard deviation of the data. The data was also visualized in a colourmap plot using the matplotlib functions and, in addition, a histogram that shows the Pixel distribution across the NDVI values. All these can be seen below in Table 4.5, Figures 4.9 and Figure 4.10:

Table 4. 5: Kpong Reservoir's NDVI Analysis results for the 4th of January 2022

Parameter	Value
MEAN NDVI	0.11458065150662002
MINIMUM NDVI	-0.0684326710816777
MAXIMUM NDVI	0.5683563748079877
STANDARD DEVIATION	0.028273112912141997

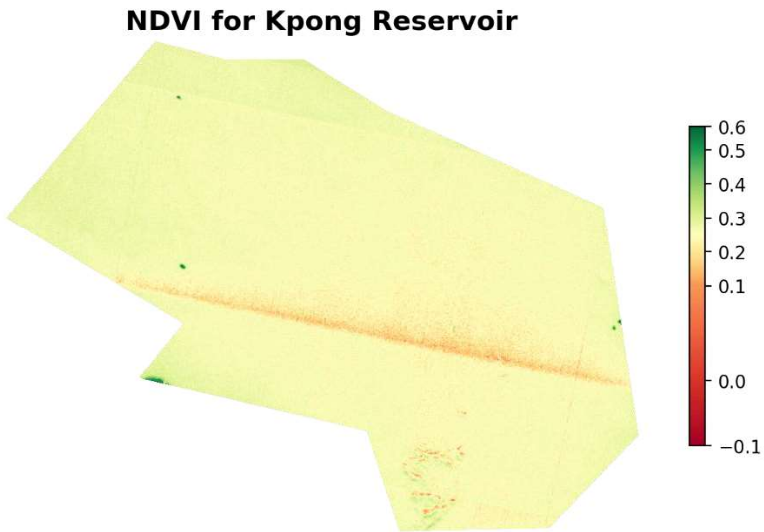


Figure 4.9: Kpong Reservoir's NDVI colourmap for the 4th of January 2022

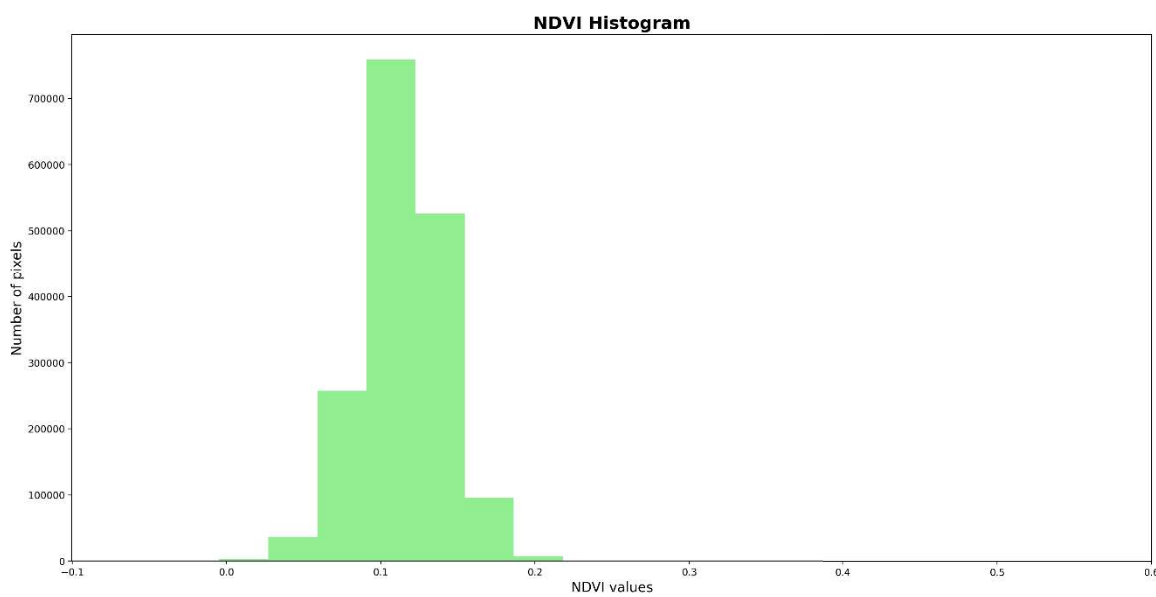


Figure 4. 10: Kpong Reservoir's NDVI histogram for the 4th of January 2022

4.3.2 NDTI Analysis of PlanetScope for Turbidity

Rasterio library is employed to open the PlanetScope data and the red bands and green bands are read from it. The bands are mathematically combined using the NDTI formula (the same formula in the first approach) to attain the NDTI raster data, which is stored in a TIF file. The NumPy library was employed to attain minimum, maximum, and mean NDTI values, as well as the Standard deviation of the data. This NDTI raster data was visualized in a colourmap and the pixel distribution across the NDTI values is shown on a histogram. These can be seen below in Table 4.6, Figures 4.11 and 4.12:

Table 4. 6: Kpong Reservoir's NDTI Analysis results for the 4^{th} of January 2022

Parameter	**Value**
MEAN NDTI	-0.04024432642780117
MINIMUM NDTI	-0.1847879083373964
MAXIMUM NDTI	0.17348256564786912
STANDARD DEVIATION	0.016968928305666247

Figure 4.11: Kpong Reservoir's NDTI colourmap for the 4th of January 2022

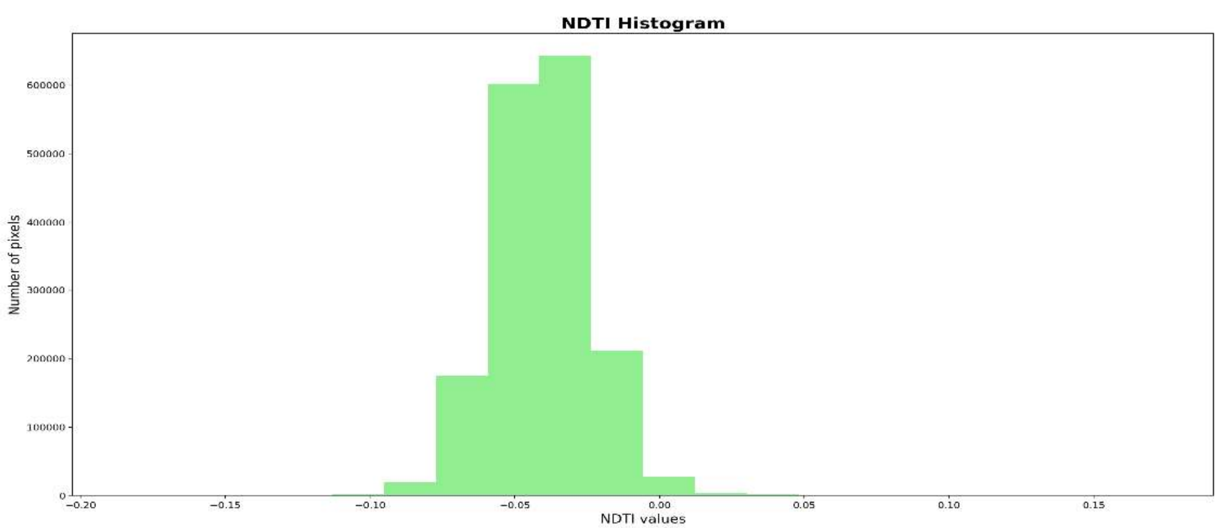


Figure 4.12: Kpong Reservoir's NDTI histogram for the 4th of January 2022

4.3.3 GNDVI Analysis of PlanetScope for Nitrogen Content

Rasterio library is employed to open the PlanetScope data and the Green bands and Near-Infrared bands are read from it. The bands are mathematically combined using the GNDVI formula (the same formula in the first approach) to attain the GNDVI raster data, which is stored in a TIF file. This GNDVI raster data is visualized in a colourmap and the pixel distribution across the GNDVI values is shown on a histogram. These can be seen below in Table 4.7, Figures 4.13 and 4.14:

Table 4. 7: Kpong Reservoir's GNDVI Analysis results for the 4th of January 2022

Parameter	Value
MEAN GNDVI	0.07469973687434434
MINIMUM GNDVI	-0.11153358681875793
MAXIMUM GNDVI	0.5135673388100572
STANDARD DEVIATION	0.029047927852804422

Figure 4.13: Kpong Reservoir's GNDVI colourmap for the 4th of January 2022

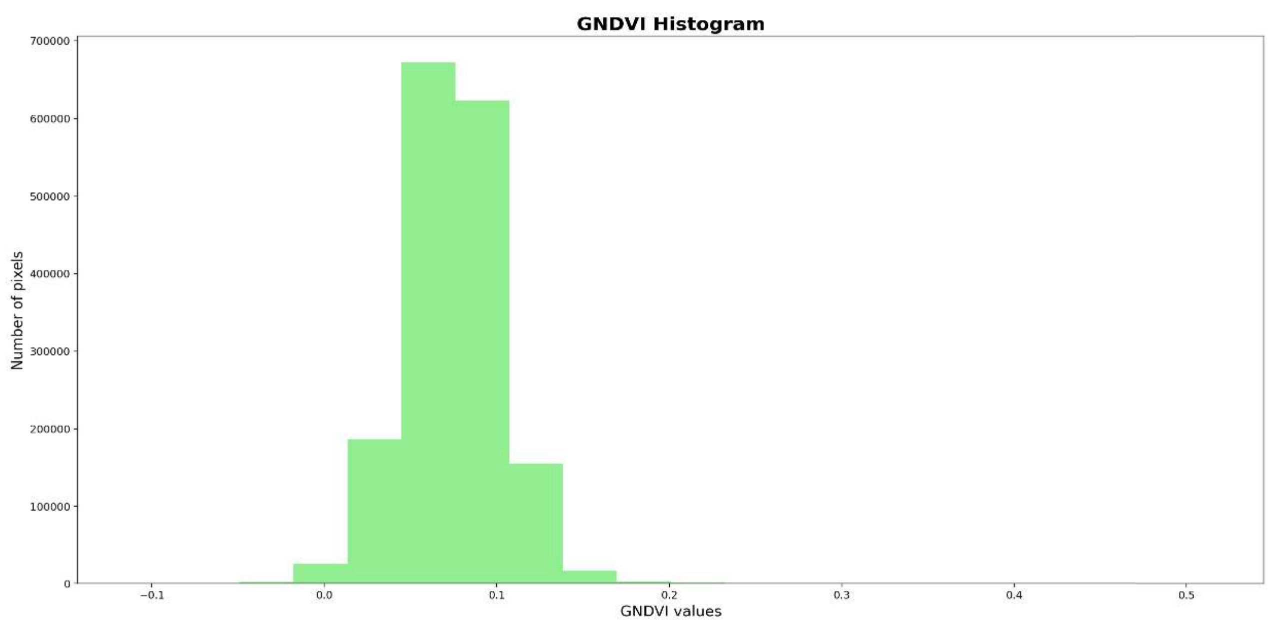


Figure 4.14: Kpong Reservoir's GNDVI histogram for the 4th of January 2022

4.4 Sending Information to Database

A MySQL database was designed and executed to store the output data of the NDVI, NDTI, FAI, and GNDVI analyses. This includes the mean, minimum, maximum, and standard deviation values as well as the figures and histograms. After the analysis of the Earth observation data, all this data is stored in the database. Functions from the mysql-connector library are employed here. Firstly, the connection is made with the host and the database of interest is selected. The information is then sent to the database and finally the connection is closed.

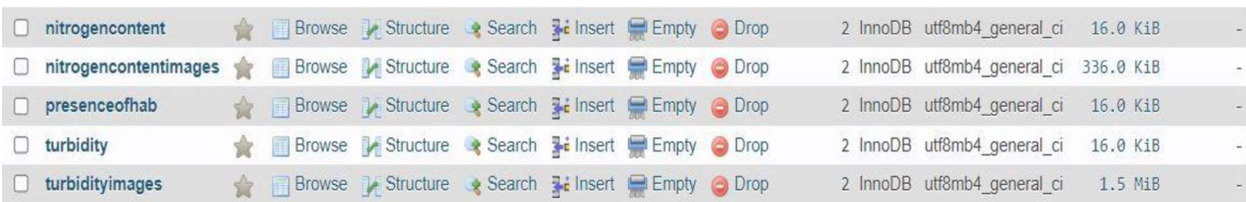


Figure 4.15:Some tables in the MySQL database

The execution of the database is a very key aspect of the project because it allows for the data which is gotten from the analysis to be stored persistently and accessed remotely. In addition, it allows for some other key aspects of the projects such as the user display interface and the alerting algorithm.

4.5 Alerting Algorithm

The Alerting algorithm is designed to operate independently from the Earth Observation data analysis algorithms (NDVI analysis, NDTI analysis, etc.). Hence the algorithm collects data from the last entries of the tables in the database. This is done using functions from the mysql-connector library. Once the data is attained, some checks are done to determine if there are significant issues with the quality of the water body in regards to the three water quality parameters. If the NDTI mean is greater than 0, the GNDVI mean is greater than 0, or the NDVI and FAI means are greater than -0.044 and 0.008 respectively, the alerting algorithm would send an email to an address of a recipient. More light would be shed on the values chosen in the subsequent chapter. This recipient would represent some form of Environmental Protection Agency or any Governmental entity that can take action in treating the waterbody. This email is sent using functions in the SMTP library. The email would include information about the issue whether it has to do with the turbidity, nitrogen content, or the presence of harmful algal blooms. Also, it would include the capture time of the Earth Observation data which was used for the analysis and the location and name of the waterbody. Lastly, there would be a direction to go and get more details from the User Display Interface.

Figure 4.16: Sample Email sent by Alerting Algorithm

4.6 User Display Interface

After the information is stored in the database, it is then displayed on a user interface. This interface is designed to display all the information that is stored in the database in a well-organized manner that can be easily understood. To be able to view the data displayed on the interface, one must have an account and be logged in with the account details. Upon logging in, there would be a home page that would show all the datasets present and these represent Earth observation data that have been analyzed. There would be a menu bar where one can access a dashboard that would show information about these indices which are used to characterize the three water quality parameters.

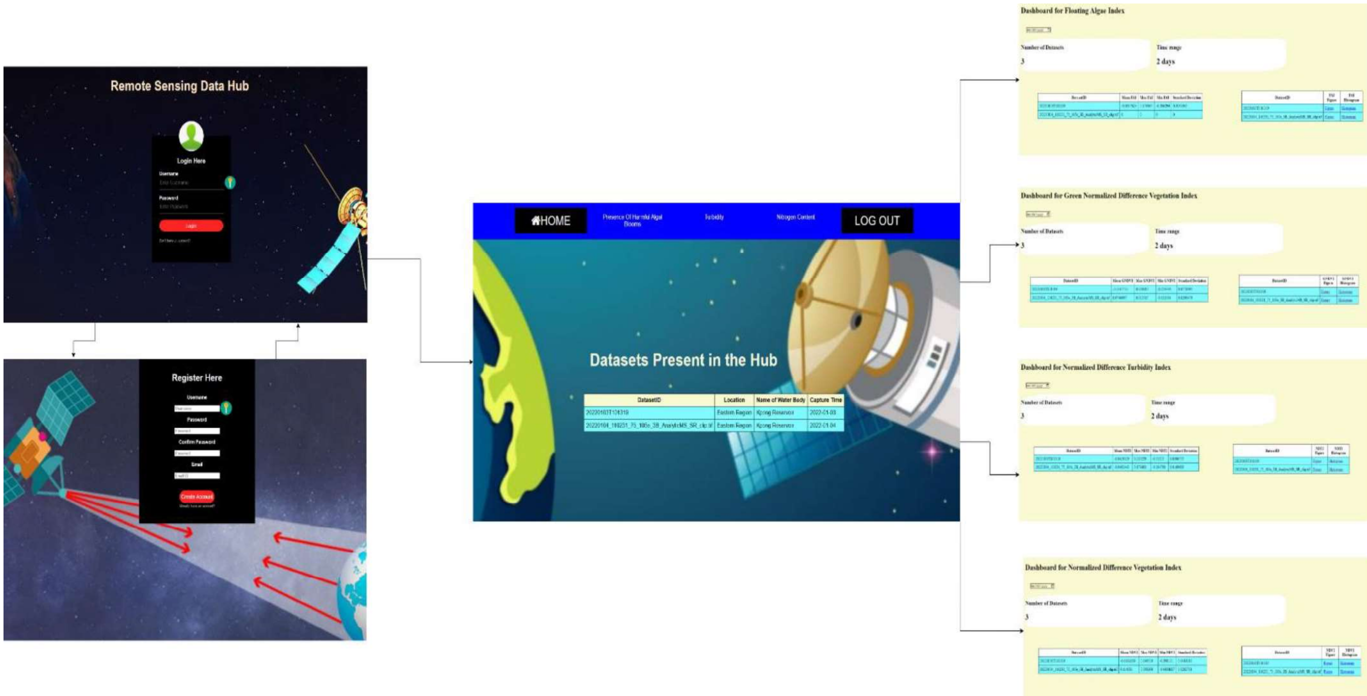


Figure 4.17: User Display Interface.

4.7 Third Approach using a CNN with PlanetScope Data

The third approach involves the use of a Convolutional Neural Network (CNN) to perform semantic segmentation on the satellite image in MATLAB. Once each pixel of the satellite image has been classified by the model, the computation of the total number of pixels that belong to the 'water' classes divided by the total number of pixels of the satellite image is performed and the outcome is expressed as a percentage.

Convolutional Neural Networks are a type of artificial neural network commonly used for image processing. It is specifically designed to process pixel data. Below in Figure 4.18, we can see a typical architecture of a Convolutional Neural Network.

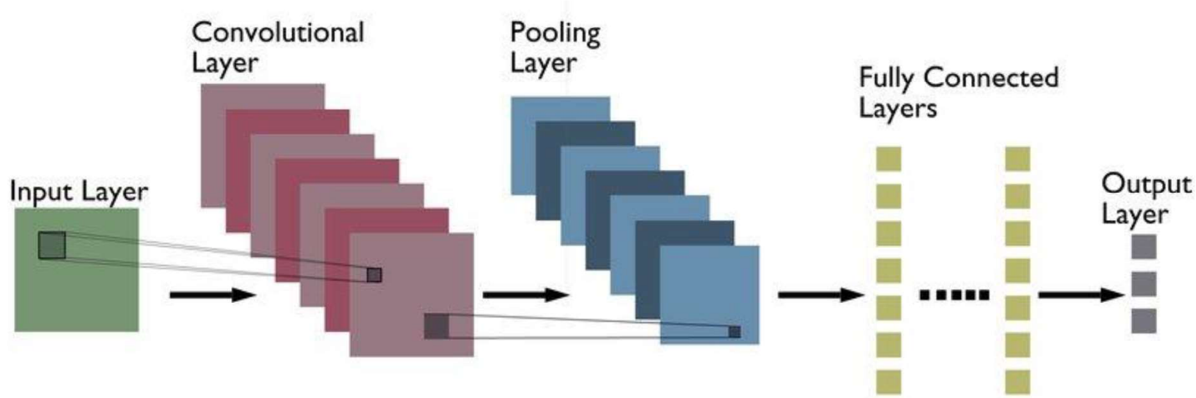


Figure 4.18: Typical CNN Architecture

The particular CNN used was a pre-trained U-Net called multispectralUnet. This network was designed to perform semantic segmentation of multispectral images such as satellite images. The network has 58 layers and requires an Image input. When the multispectralUnet was applied to a multispectral image of Hamlin Beach State Park, it displayed a global accuracy of 0.90698 when compared to the ground truth [20]. There are two 'water' classes present in the model namely 'Water_Lake' and 'Water_Pond'. The number of pixels recognized by the model to be a part of these classes is divided by the total number of valid pixels of the image and expressed as a percentage in Equation 6.

$$\textbf{Percentage of WaterCover} = \frac{\textit{Total number of water pixels}}{\textit{Total number of valid pixels}} * \mathbf{100}\% \qquad \textbf{(6)}$$

The same PlanetScope Data of the Kpong Reservoir captured on the 4th of January 2022 used in the second approach was loaded into the MATLAB workspace. The data was then ensured to be of the appropriate dimensions to be used as an image input for the multispectralUnet model. Once this was done, the multispectralUnet was applied and the segmented image was produced. This can be seen below in Figure 4.19:

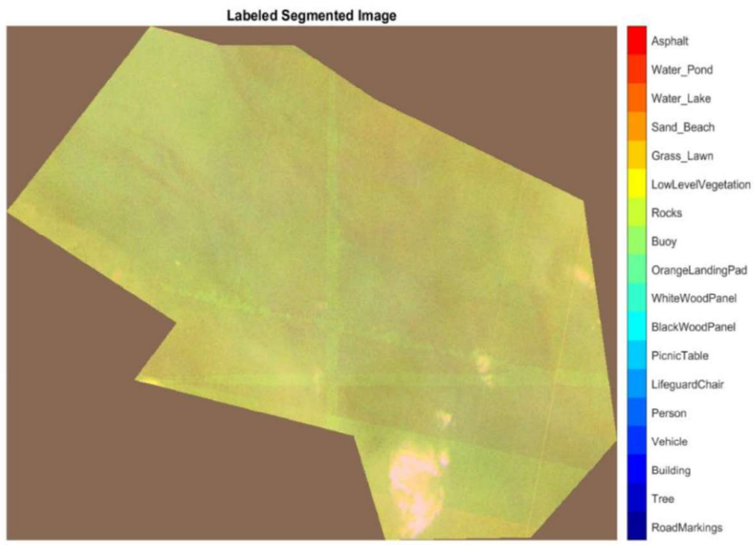


Figure 4.19: Segmented Image

The Percentage of Water Cover attained for this PlanetScope Data was 99.98287%.

For this approach, the Percentage of Water Cover was attained from PlanetScope Data across a month (from the 8th of March to the 7th of April 2022) of the Kpong Reservoir. The outcome can be seen below in Table 4.8:

Table 4. 8: Percentage of Water Cover across the Month

Time(days)	Date	Percentage of Water Cover (%)
0	8th March 2022	99.94559
1	9th March 2022	99.93652
11	20th March 2022	99.95300
18	27th March 2022	99.98112
19	28th March 2022	99.94623
29	7th April 2022	99.96271

Pearson correlation analysis was executed in Excel to study the relationship between Time and the Percentage of Water Cover and in essence, see how the Percentage of Water Cover changes over some time. The outcome of this can be seen below in Figure 4.20:

Time(days)	WaterPercentageCover(%)				
0	99.94559				
1	99.93652			**Coefficient:**	0.617606
11	99.953			**N**	6
18	99.98112			**T statistic:**	1.570543
19	99.94623			**DF:**	4
29	99.96272			**p value:**	0.19138

Figure 4.20: Outcome of Pearson Correlation Analysis

Chapter 5: Interpretation of Results

Now that the data about the different water quality parameters have been obtained and also shown on the User Display Interface, it is key that conclusions are now made from the data. To ensure that the data obtained is actionable, information must be extracted from it.

5.1 Interpretation of NDVI results.

The mean value, standard deviation, and histogram give a fair idea of how the pixels of the satellite image (Earth observation data) are distributed across the NDVI values. The colourmap shows the NDVI value at different points in the satellite image. In addition, it aids with easy visualization of the data. The mean value tells us on average what the NDVI value of a pixel in the satellite image would be. These apply to all the other indexes. The interpretation is realized in the meaning behind the NDVI values and this can be seen below in Figure 5.1:

Class	Pixel Value Range
Water	−0.145 to −0.044
Seaweed	−0.044 to 0.121
Mangrove	0.159 to 0.372

Figure 5.1: Pixel Value Range for NDVI [12]

For the first approach, as can be seen in Table 4.1, the mean NDVI value was -0.016163939591295844 and for the second approach, as can be seen in Table 4.5, the mean NDVI value was 0.11458065150662002. The mean from the first implementation is slightly above the upper limit of the water class and this suggests that the majority of the satellite image is representing water alone and minimal vegetation. The mean from the second approach (which is in the seaweed class) suggests a high presence of vegetation in the water body in the form of seaweed and harmful algal blooms. The NDVI is not ideal to be used alone in the detection of the presence of harmful algal blooms and so that is why it is used in conjunction with the FAI.

5.2 Interpretation of FAI results.

Class	Pixel Value Range
Water	−0.510 to −0.008
Seaweed	0.008 to 0.130
Mangrove	0.130 to 0.530

Figure 5.2: Pixel Value Range for FAI [12]

The FAI mean value for the first approach, as can be seen in Table 4.2, is -0.0037624047. This is somewhere in between the water class and the seaweed class. This suggests that the majority of the satellite image is representing water alone but there is a significant part that represents vegetation in the form of seaweed and harmful algal blooms. Through the two approaches and two indexes, we can conclude that there is some presence of harmful algal blooms in the water body.

5.3 Interpretation of NDTI results

The NDTI mean value for the first approach, as can be seen in Table 4.3, is -0.04293288341883818 and for the second approach, as can be seen in Table 4.6, is -0.04024432642780117. Pixels with NDTI values that are above 0 represent non-water pixels. In essence, alien material in the water body. A mean value of above 0 would indicate high turbidity levels in the water body. The mean values in the two approaches are all below 0 and this suggests lower turbidity levels. Figure 5.3 is from a paper that studies and proves a positive correlation between the total suspended solids concentration and the NDTI values. This correlation validates the point above.

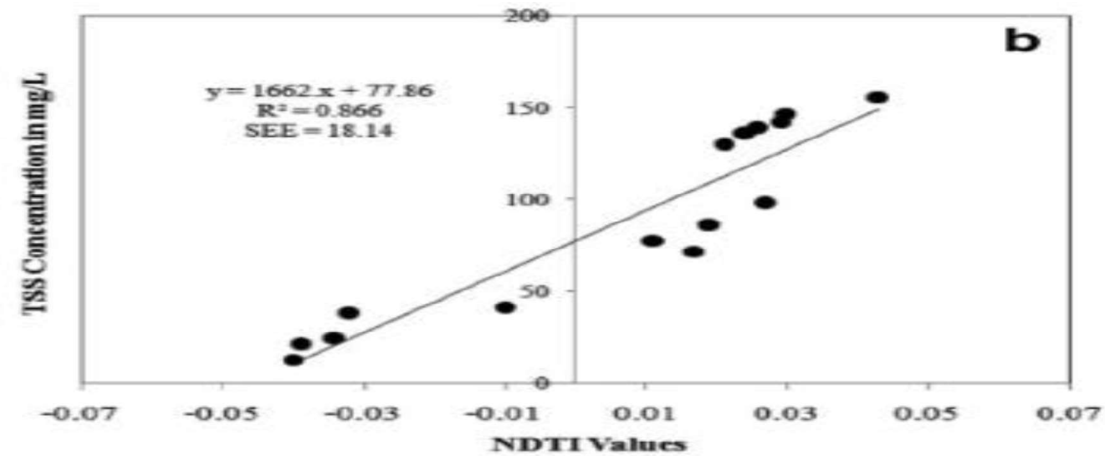


Figure 5.3: Graph of TSS Concentration against NDTI values [13]

5.4 Interpretation of GNDVI results

GNDVI values between -1 and 0 represent water and bare soil but above 0 represent photosynthetic activity and nitrogen content [14]. The mean GNDVI value for the first approach, as can be seen in Table 4.4, is -0.0587754876915526 and for the second approach, as can be seen in Table 4.7, is 0.07469973687434434. This outcome suggests a significant but not so high nitrogen content in the water.

5.5 Interpretation of Water Percentage Cover

The Water Percentage Cover ideally measures the proportion of the area captured by the image that is covered by water. Since the satellite image is of a water body ideally you would be expecting a value of 100%, however, due to impurities and the presence of alien materials the value would be less than that. The larger impurities and the presence of alien materials the lower the water quality. Hence this percentage value can be used as an indicator of the water quality. The idea of using the water percentage cover as an indicator of water quality works on the assumptions that the impurities in the water are not colourless and the model can accurately classify every pixel.

The outcome of the Pearson correlation analysis, as can be seen in Figure 4.20, show a correlation coefficient (r) value of 0.617606 and a p-value of 0.19138. In this implementation, the p-value suggests that there is no significant positive correlation between the Time and Water Percentage Cover. This means the Water Percentage Cover hasn't changed much over time from the 8^{th} of March 2022 to the 7^{th} of April 2022. It can be further deduced that there has been no prominent activity that has reduced the water quality over the time frame nor has there been any activity that has increased the quality of the water body.

Chapter 6: Conclusion and Future Works

6.1 Discussion

The results attained in chapter 4 show that a model was developed using the four indices to analyze the Earth observation data (satellite imagery). The overall intention was primarily to deduce information about water quality.

The interpretation of the results in chapter 5 indicates how the data obtained from the model can be used to gain information about the state of the waterbody regarding the three major parameters. In addition, the translation between attaining the Earth Observation data and knowing what is going on at the waterbody is complete after the interpretation of the results.

The implementation using two approaches allows for high accuracy and also for sustainability and cost-effectiveness. The first approach employs Sentinel-2 data which is open-source data and so can be persistently used without worry of cost. The second approach employs PlanetScope data which is quite expensive but has an extremely high spatial resolution of 3m per pixel and so would lead to very accurate data.

The User Display Interface allows for success in displaying the information obtained in a way that it can be easily understood and accessed.

The Alerting algorithm allows for quick detection of problems with the water body and hence quick resolution of the problem.

6.2 Limitations

In the execution of the project, many limitations prevented the full realization of this project. Firstly, the cost associated with the PlanetScope data limited the amount of Earth observation data that I could access. The Education and Research program had a time limit as well as a limit on the amount of data that could be downloaded. This issue prevented there being sustainability going forward. There were also issues with the consistency of the data available because ideally, PlanetScope data has a temporal resolution of one or two days but sometimes in two days there won't be access to average surface reflectance data of the Area of Interest. Only TOAR (top of the atmosphere reflectance) data would be accessible and this wasn't useful to the project. In addition, sometimes in the span of two days or even more there won't be available data on the entire area of interest. These limitations especially affected the execution of the third approach which used PlanetScope data and sought to see how the water quality changed over time.

Furthermore, the PlanetScope data that I had access to via the Education and Research program did not contain the Shortwave Infrared band and this prevented me from doing the FAI analysis in the second approach.

6.3 Future Work

The future work for this project is to fully automate the processes that make up the system to allow for all of this to be done without human influence. In addition, the alerting algorithm would not only communicate through email but also via SMS/MMS. Lastly, future works would inculcate more water quality parameters and hence more indices would be used to get a better idea of the quality of the waterbody.

References

[1]"Ghana | Safe Water Network", *Safewaternetwork.org*, 2021. [Online]. Available: https://www.safewaternetwork.org/countries-regions/ghana. [Accessed: 10- Oct- 2021].

[2]"Water", *Unicef.org*, 2021. [Online]. Available: https://www.unicef.org/ghana/water. [Accessed: 10- Oct- 2021].

[3]*In Situ Water Quality Monitoring*. Georgia: U.S Environmental Protection Agency, 2018.

[4] J. Tao, Z. Zhang and W. Yu, "Monitoring Taihu water quality by using high-resolution satellite image", *2011 19th International Conference on Geoinformatics*, 2011. Available: 10.1109/geoinformatics.2011.5980743 [Accessed 21 December 2021].

[5]S. Liew, B. Saengtuksin and L. Kwoh, "Monitoring turbidity and suspended sediment concentration of coastal and inland waters using satellite data", *2009 IEEE International Geoscience and Remote Sensing Symposium*, 2009. Available: 10.1109/igarss.2009.5418225 [Accessed 21 December 2021].

[6]A. Muhairi, H. Ghedira, H. Al-Ahmad and A. Dawood, "Using remote sensing satellites for water quality monitoring in the UAE", *2011 IEEE GCC Conference and Exhibition (GCC)*, 2011. Available: 10.1109/ieeegcc.2011.5752634 [Accessed 10 January 2022].

[7]"Satellite imagery access and analysis in Python & Jupyter notebooks", *Medium*, 2022. [Online]. Available: https://towardsdatascience.com/satellite-imagery-access-and-analysis-in-python-jupyter-notebooks-387971ece84b. [Accessed: 12- Jan- 2022].

[8]2022. [Online]. Available: https://eos.com/find-satellite/sentinel-2/. [Accessed: 12- Jan- 2022].

[9]"Sentinel 2 Pre-processing Requirements for coastal and inland waters", *Eurosdr.net*, 2021. [Online]. Available: http://www.eurosdr.net/sites/default/files/images/inline/3-brockmann_ws_sentinel.pdf. [Accessed: 13- Jan- 2022].

[10]2022. [Online]. Available: https://www.satimagingcorp.com/. [Accessed: 11- Mar- 2022].

[11] M. Visitacion, C. Alnin, M. Ferrer and L. Suñiga, "DETECTION OF ALGAL BLOOM IN THE COASTAL WATERS OF BORACAY, PHILIPPINES USING NORMALIZED DIFFERENCE VEGETATION INDEX (NDVI) AND FLOATING ALGAE INDEX (FAI)", *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. -419, pp. 479-486, 2019. Available: 10.5194/isprs-archives-xlii-4-w19-479-2019.

[12] M. Siddiqui, A. Zaidi and M. Abdullah, "Performance Evaluation of Newly Proposed Seaweed Enhancing Index (SEI)", *Remote Sensing*, vol. 11, no. 12, p. 1434, 2019. Available: 10.3390/rs11121434 [Accessed 4 April 2022].

[13]S. Bid and G. Siddique, "Identification of seasonal variation of water turbidity using NDTI method in Panchet Hill Dam, India", *Modeling Earth Systems and Environment*, vol. 5, no. 4, pp. 1179-1200, 2019. Available: 10.1007/s40808-019-00609-8 [Accessed 4 April 2022].

[14]A. Michlig, "What is the GNDVI?", *Auravant*, 2022. [Accessed 4 April 2022].

[15]"Introducing the Spectral Index Library in ArcGIS", *Esri*, 2022. [Online]. Available: https://www.esri.com/about/newsroom/arcuser/spectral-library/. [Accessed: 04- Apr- 2022].

[16]"What is NDVI (Normalized Difference Vegetation Index)? - GIS Geography", *GIS Geography*, 2022. [Online]. Available: https://gisgeography.com/ndvi-normalized-difference-vegetation-index/. [Accessed: 04- Apr- 2022].

[17] C. Hu, "A novel ocean color index to detect floating algae in the global oceans", Remote Sensing of Environment, vol. 113, no. 10, pp. 2118-2129, 2009. Available: 10.1016/j.rse.2009.05.012 [Accessed 4 April 2022].

[18] A. Hossain, C. Mathias and R. Blanton, "Remote Sensing of Turbidity in the Tennessee River Using Landsat 8 Satellite", Remote Sensing, vol. 13, no. 18, p. 3785, 2021. Available: 10.3390/rs13183785 [Accessed 4 April 2022].

[19]"GNDVI—ArcGIS Pro | Documentation", Pro.arcgis.com, 2022. [Online]. Available: https://pro.arcgis.com/en/pro-app/2.8/arcpy/image-analyst/gnvdi.htm. [Accessed: 04- Apr- 2022].

[20]*MathWorks*, 2022. [Online]. Available: https://www.mathworks.com/. [Accessed: 25- Apr-2022].

Appendix

Appendix A: Snippets of Python Code in Jupyter Notebooks

```python
print(numpy.nanmin(GNDVI))
print(numpy.nanmax(GNDVI))
print(numpy.nanmean(GNDVI))
print(numpy.nanstd(GNDVI))
```

```
-0.11153358681875793
0.5135673388100572
0.07469973687434434
0.029047927852804422
```

+ Code + Markdown | Run All Clear Outputs of All Cells Restart Interrupt Variables Outline ··· geo_env (Python 3.9.7)

Opening bands with rasterio

```python
import rasterio as rio
file='20220104_110231_75_105e_3B_AnalyticMS_SR_clip.tif'
satdat=rio.open(file)

GREEN=satdat.read(2)
NIR=satdat.read(4)
```

Computation of GNDVI and Storing in a TIF file

```python
import numpy
numpy.seterr(divide='ignore',invalid='ignore')

GNDVI=(NIR.astype(float)-GREEN.astype(float))/(NIR+GREEN)


meta=satdat.meta
print(meta)
GNDVI_type=GNDVI.dtype
print(GNDVI_type)


kwargs=meta

kwargs=meta

kwargs.update(dtype=GNDVI_type)
kwargs.update(count=1)
kwargs.update(driver='GTiff')
```

+ Code + Markdown | Run All Clear Outputs of All Cells Restart Interrupt Variables Outline ··· geo_env (Python 3.9.7)

GNDVI for Nitrogen Content

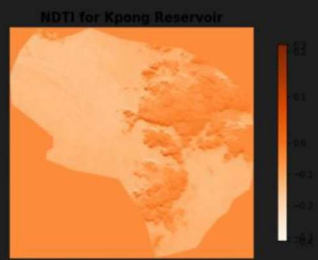
Opening bands with rasterio

```python
import rasterio as rio
BANDS='S2B_MSIL2A_20220103T101319_N0301_R022_T30NZM_20220103T130628.SAFE\GRANULE\L2A_T30NZM_A025216_20220103T101650\IMG_DATA\R10m'
B3_GREEN=rio.open(BANDS+'/T30NZM_20220103T101319_B03_10m.jp2')
B8_NIR=rio.open(BANDS+'/T30NZM_20220103T101319_B08_10m.jp2')
```

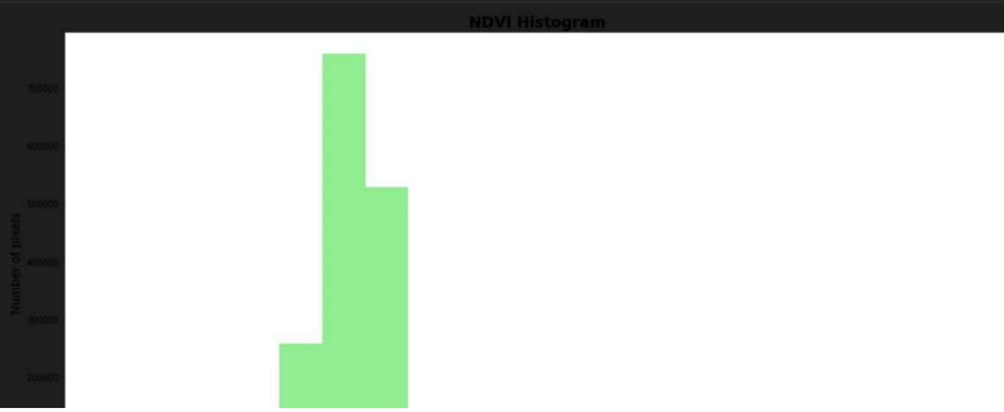
Computation of GNDVI and storing in a TIF file

```python
import numpy
numpy.seterr(divide='ignore',invalid='ignore')
GREEN=B3_GREEN.read()
NIR=B8_NIR.read()
GNDVI=(NIR.astype(float)-GREEN.astype(float))/(NIR+GREEN)

meta=B3_GREEN.meta
print(meta)
GNDVI_type=GNDVI.dtype
print(GNDVI_type)


kwargs=meta

kwargs=meta

kwargs.update(dtype=GNDVI_type)
kwargs.update(count=1)
kwargs.update(driver='GTiff')
```

```python
import matplotlib.pyplot as plt

min=numpy.nanmin(NDTIRI)
max=numpy.nanmax(NDTIRI)

mid=0.0
colormap=plt.cm.Oranges
norm= MidpointNormalize(vmin=min,vmax=max,midpoint=mid)
fig=plt.figure(figsize=(10,5))
ax=fig.add_subplot(111)
NDTIRI=NDTIRI.squeeze()
cbar_plot=ax.imshow(NDTIRI, cmap=colormap, norm=norm)
ax.axis('off')
cbar=fig.colorbar(cbar_plot,orientation='vertical',shrink=0.85)
ax.set_title('NDTI for Kpong Reservoir',fontsize=15,fontweight='bold')
fig.savefig("NDTI_data/ndti-fig",dpi=200,bbox_inches='tight',pad_inches=0.7)
```




```python
fig2 = plt.figure(figsize=(20,10))
ax = fig2.add_subplot(111)
plt.title("NDVI Histogram", fontsize=18, fontweight='bold')
plt.xlabel("NDVI values", fontsize=14)
plt.ylabel("Number of pixels", fontsize=14)
x = NDVI[~numpy.isnan(NDVI)]
bins = 20
color = 'lightgreen'
ax.hist(x,bins,color=color)
fig2.savefig("NDVI_PlanetScope_data/ndvip-histogram.png", dpi=200, bbox_inches='tight', pad_inches=0.7)
plt.show()
```



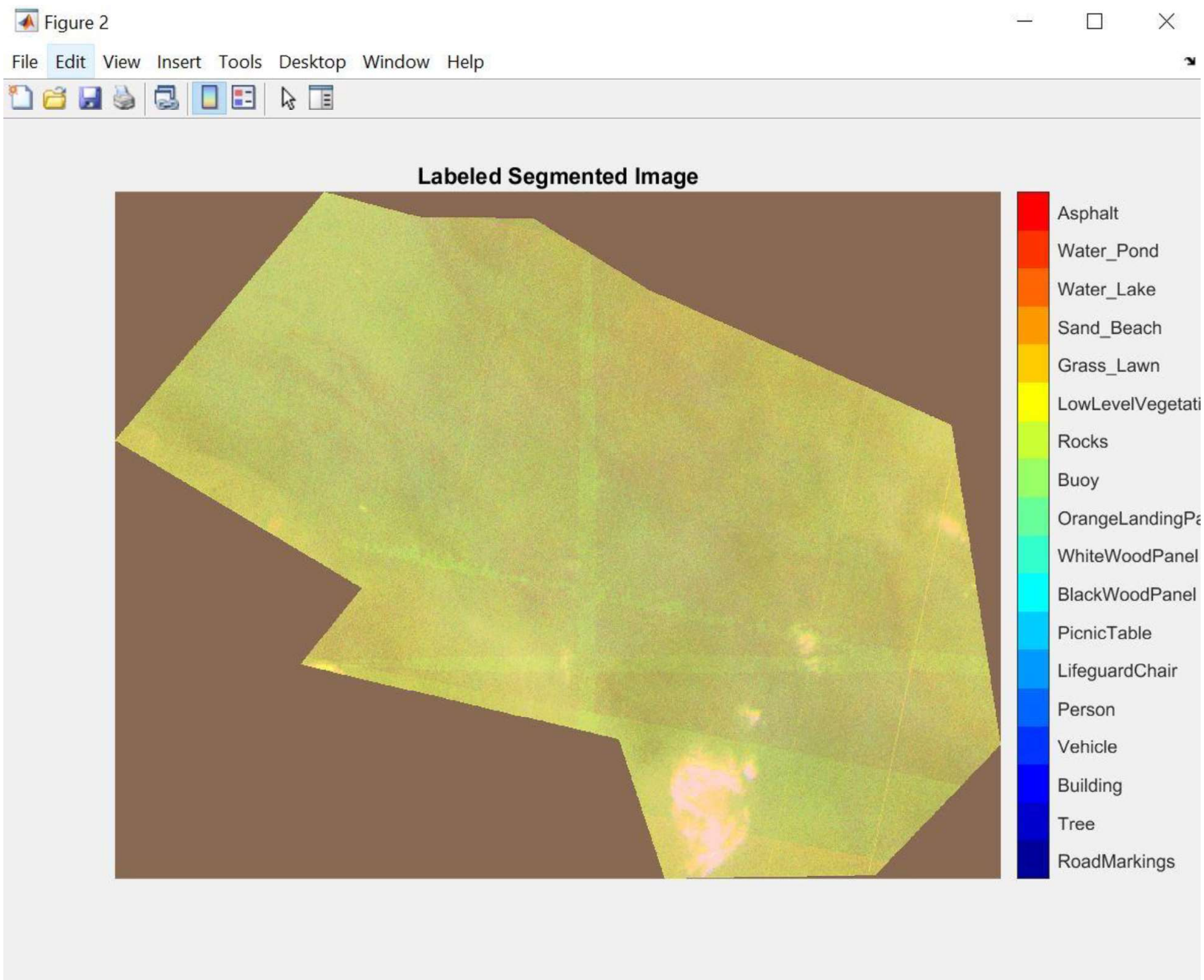
Appendix B: Snippets of MATLAB code

```
Data=imread("20220104_110231_75_105e_3B_AnalyticMS_SR_clip.tif");
redChannel=Data(:,:,1);
greenChannel=Data(:,:,2);
blueChannel=Data(:,:,3);
NIR1=Data(:,:,4);
NIR2=Data(:,:,4);
NIR3=Data(:,:,4);
FinalData=cat(3,redChannel,greenChannel,blueChannel,NIR1,NIR2,NIR3);

dataDir = fullfile(tempdir,"rit18_data");
trainedUnet_url = "https://www.mathworks.com/supportfiles/vision/data/multispectralUnet.mat";
downloadTrainedNetwork(trainedUnet_url,dataDir);
load(fullfile(dataDir,"multispectralUnet.mat"));
```

```
colorbar(TickLabels=cellstr(classNames),Ticks=ticks,TickLength=0,TickLabelInterpreter="none");
colormap(cmap)

WaterClassIds = uint8([16,17]);
WaterPixels = ismember(segmentedImage(:),WaterClassIds);
validPixels = (segmentedImage~=0);

numWaterPixels = sum(WaterPixels(:));
numValidPixels = sum(validPixels(:));

percentWaterCover = (numWaterPixels/numValidPixels)*100;
fprintf("The percentage of water cover is %3.5f%%.",percentWaterCover);
```

Command Window

```
  patch      1024x1024x6            12582912  uint16

  Name          Size                   Bytes  Class     Attributes

  patch      1024x1024x6            12582912  uint16

fx The percentage of water cover is 99.98287%.>>
```



Appendix C: Source Code

https://github.com/YawBrown/Capstone