SPATIOTEMPORAL MODELLING & AUTOMATED *IN-SITU* SENSORS TO MONITOR HARMFUL ALGAL BLOOMS (HABS)

Case Study Lake Victoria

by

Okomo Jacob Okello

Project research report submitted to the department of Geomatic Engineering and Geospatial Information Systems for the award of degree of Bachelor of Science in Geospatial Information Sciences (GIS), 2021.

Department of Geomatic Engineering and Geospatial Information Systems (GEGIS)

DECLARATION

I declare that this project is my own work and has not been submitted by anybody else in any other university for the award of any degree to the best of my knowledge.

Sign…… Date……...

Jacob Okello Okomo

ENC222-0149/2017

Department of Geomatic Engineering and Geospatial Information Systems (GEGIS)

Jomo Kenyatta University of Agriculture and Technology

CERTIFICATION

This project has been submitted for examination with my approval as the candidate's supervisor.

Sign……. Date ……

Dr. Eunice Nduati, PhD.

Senior Lecturer, GEGIS

Department of Geomatic Engineering and Geospatial Information Systems (GEGIS)

Acknowledgements

The achievement(s) in this thesis submission were not fully accomplished solely by the student author. A number of individuals helped me with mentorship and logistical support during the process.

First and foremost, I'd want to express my sincere gratitude to my major advisor and supervisor, Dr. E. Nduati, PhD from the Department of Geomatic Engineering and Geospatial Information Systems-GEGIS. Your enormous energy and enthusiasm serve as a constant reminder of what I enjoy about geoscience. Thank you very much. You deserve special recognition. Both professionally and emotionally, you have been an amazing role model for me. You believed in my ideas and gave me the freedom to pursue them. Working under your guidance has provided me with the resources and confidence I need to accomplish my career achievements.

I sincerely thank Google Earth Engine and United States Geological Survey (USGS) for the provision of Satellite Remote sensing images, particularly Landsat 8 image collection that came in handy for this project realization.

Sincere gratitude is as well sent to Kenya Marine and Fisheries Research Institute (KMFRI) for providing a tone of information that are directly in line with this project.

I would as well like to acknowledge the efforts of both teaching and non-teaching staff from the Department of GEGIS who directly impacted me with knowledge in various aspects and skills that have played key roles in my project. To you all, may you be blessed abundantly.

Abstract

Over the past few decades, there has been increased frequency of harmful algal blooms (HABs) in various water bodies and this has been a major global environmental concern, especially now that the SDGs is the story of the day. Resulting not only in increased waterborne diseases, massive death of aquatic animals especially fish (aquaculture), loss of aesthetic appeal of the affected body, treatment costs for drinking water but also adverse impacts on tourism and therefore the GDP. Classical and traditional approaches like field sampling using ship and boats and later laboratory tests and analysis have been adopted to assess HABs and address on various measures in Lake Victoria basins. Further, there have been complete reliance on the local occupants to notify the authorities in case of such catastrophic phenomena which has proven immaterial. This is so true because these methods (reporting and assessment) have proven labour intensive and costly and do not provide synoptic views of the bloom conditions. In this thesis, a near real-time satellite remote sensing-based approach has been adopted, in combination with near real-time *in situ* IoT system, to come up with a powerful tool that can offer valuable spatiotemporal information about a bloom of the *Cyanobacteria* that has been(lately) a repeating devastation in the Lake Victoria, Nyanza gulf in Kenya. By adopting the Ocean Colour algorithm together with the mono-window Lake Surface Air Temperature algorithm, this has enhanced wide-spread detection and mapping of HABs even at 30m spatial resolution with Landsat mission, revealing surface patches and a heterogeneous distribution of such occurrences. The results obtained were further validated with well-known and accepted products from Copernicus' Sentinel-3 and NASA's MODIS missions. This research also presents the deployment of an *in-situ* based IoT system for continuous near real-time monitoring of LSAT. Therefore, the value provided by the Landsat products in terms of frequency and synoptic observations is of paramount importance for ecological and management purposes at regional and national scales of the said phenomena.

Table of contents

List of figures

List of tables

Acronyms and abbreviations

1.	DO	Dissolved Oxygen
2.	DN	Digital Numbers
3.	GDP	Gross Domestic Product
4.	GEGIS	Geomatic Engineering and Geospatial Information Systems
5.	HAB	Harmful Algal Bloom
6.	FAI	Floating Algae Index
7.	FLH	Fluorescence Line Height
8.	IoT	Internet of Things
9.	KMFRI	Kenya Marine and Fisheries Research Institute
10.	KTN	Kenya Television News
11.	LSAT	Lake Surface Air Temperature
12.	LSWT	Lake Surface Water Temperature
13.	MCI	Maximum Chlorophyll Index
14.	MERIS	MEdium Resolution Imaging Spectroradiometer
15.	ML	Machine Learning
16.	MODIS	MODerate Resolution Imaging Spectrometer
17.	NASA	National Aeronautics and Space Administration
18.	OLI	Operational Land Imager
19.	OC	Ocean Color
20.	Quick-SCAT	Quick Scatterometer
21.	SDGs	Sustainable Development Goals
22.	SeaWiFS	Sea-viewing Wide-Field- of-View Sensor
23.	SST	Sea Surface Temperature
24.	TIRS	Thermal InfraRed Sensor
25.	USGS	United States Geological Survey

1 Introduction

1.1 Background

Richardson K. (1997) defined Algal bloom as "the rapid growth of one or more species which leads to an increase in biomass of the species". This is normally associated with high concentrations of phytoplankton (algae). If the rapid growth is related to a toxic or harmful species, then the phenomenon can be termed as Harmful Algal Bloom (HAB). A species can be harmful due to the release of toxic substances e.g., *Cyanotoxins spp*. which are the most frequent in the Lake Victoria region (Okello et al., 2011). Often termed as "Red Tides", HABs have attracted a significant world-wide attention in research over the last two decades (W. Song et al.,2015; Manuel et al., 2020).

Development, stability, and density of the phenomenon are related to some environmental factors such as wind velocity and direction, Lake Surface Air Temperature (LSAT), Sea Surface Temperature (SST), currents, adequate nutrient concentration, enough sunlight, warm temperatures (Tang et al, 2006) for them to be transported and distributed over a water body e.g., Kisumu Bay especially in large aggregates called *colonies* (Okello & Kurmayer, 2011).

A lot of scholars in the geosciences have put forward a bunch of approaches to detect and monitor HABs in both inland and ocean waters including generating indices from spectral band ratio algorithms. For example, Empirical visible-NIR band ratios (Matthews et al, 2012; Allan et al, 2015), blue-green band ratios (O'Reilly et al., 1998), red-edge (RE) region (685–720 nm) band ratios (Mittenzwey et al., 1992), thermal band-based assessment (Tang et al.,2006) reportedly can be exploited to detect and monitor occurrence of a variety of HABs.

Currently, remote sensing of HAB detection and monitoring methods are primarily designed for sensors like MODIS, MERIS and SeaWiFS, (Kurekin et al., 2014) which despite allowing us to continuously monitor the behavior of the phenomenon (due to high temporal resolution (about 1 day revisit time), limits us from a detailed examination of HABs (due to coarse spatial resolutions (250~1130 meters) and only large scale HABs can be monitored by using them (Blondeau, 2014).

Landsat 8 OLI being able to discriminate finer details, therefore possess the great value and potential to provide for the retrieval of Chl-a at 30m spatial resolution (Allan et al., 2015; Watanabe et al., 2015; Concha and Schott 2016; Manuel et al., 2020).

The advent and uptake of Internet of Things (IoT) further provides for quick development of geo-intelligent automated in-situ systems that collect near real-time water quality data e.g., LSAT which is a thermal proxy, Lake Salinity as another in-situ indicator of HAB presence thereby enabling a step-change in data availability for HAB monitoring in Lake Victoria.

Therefore, an approach that utilizes high spatiotemporal sensor's capability to detect and monitor HABs coupled by in-situ systems technically sounds essential. Landsat 8 high spatiotemporal satellite images (30 & 100 meters for spatial resolution and 16 days for temporal) have advanced the space-based capabilities in detection and monitoring of HABs comparatively more accurately at relatively smaller inland water bodies e.g., Nyanza Gulf of Lake Victoria.

On that regard, this study intends to demonstrate the potential of using spectral bands 2, 3, 4, 5 and TIR band 10 of L8 OLI in detecting and monitoring of HABs using empirical statistical methods. Further extending these space-based capabilities with *in-situ* based IoT systems for near real-time monitoring of the said proxies.

1.2 Motivation and problem statement

Harmful Algal Blooms continue to be a major source of concern, not only because of their huge environmental and socioeconomic consequences, but also because of a recent dramatic increase in the number of cases documented around the world (Hill et al., 2020). HABs have the potential to initiate serious environmental and human health issues, as well as other complications as seen in deterioration of economic status thereby negatively impacting a region's GDP. Environmental impacts include depletion of dissolved oxygen (DO) in the aquatic habitat causing mass fish stock (Tang et al, 2006) (*Fig 1,2,3 below*). Impacts posed to human inhabitants include toxic reactions to affected seafood and even in extreme cases, fatalities. Related economic impacts include adverse effects on coastal based industries e.g., fishing (Smith et al., 2019), tourism (*Fig 4*).

Lately, coastal regions of Lake Victoria especially around Nyanza Gulf (Kisumu Bay) have shown deterioration in its water quality as seen in severe signs of eutrophication with blooms (Simiyu et al., 2018).HABs can be caused by a variety of circumstances, but they are most commonly associated with favorable environmental conditions, such as increased nutrient levels-eutrophication (Santoleri et al., 2003), which are generally driven by urbanization, deforestation and poor agricultural practices (Gohin F. et al., 2006; Hecky et al., 2010).

Another significant problem of interest lies from the fact that despite all the above highlighted health, economic and environmental threats, upon occurrence of the phenomena, there exist a notable challenge in timely-notifying the higher authority of the same. The local maritime authorities need to take immediate responsive action to advice the locals on how to cautiously carry out their activities with minimum contact with the HAB and if any, avert further impacts. The status quo in the study area currently only provides for the authorities e.g., KMFRI and other Lake Victoria sanitations bodies to solely rely on information about the occurrence to be relayed to them by the local citizens and this is normally after the situation is a great mess. (Roegner et al. 2020). The study therefore seeks to address on the gaps.

1.3 Justification

One of the bays of Lake Victoria that is mostly affected by nutrient enrichment is Nyanza Gulf (Gikuma-Njuru, P. 2013) which is discharged from the densely populated catchment with mostly subsistence agriculture (Calamari, D. 1995; Hecky, R.E. 2010). This has led to regular occurrence of bloom-forming cyanobacteria which has been associated with mass fish kills *(Figures 1, 2 & 3 below)* due to depletion of dissolved oxygen (DO) in the aquatic habitats and to some extremes, temporary shutdown of drinking water supply, i.e., from January to March 2004 (Sitoki et al., 2012).

For this purpose, a number of Landsat 8 images acquired in a bloom condition will be quantitatively analyzed for investigation of the capabilities of L8 being capable of detection and extraction of the Chl-a pigments and the rising LSAT values in the affected regions.

Figure 1: Standard Media Kenya reports through KTN about HABs causing Mass Fish stocks in Dunga Beach Kisumu region in February 2021, (Image: Standard Media KE)

Figure 2: Mass fish stock in Lake Victoria Shores impacted by the HABs (Image: Standard Media KE)

Figure 3: High valued Fish lost to HAB in Lake Victoria Shores (Image Source: KMFRI)

Figure 4: Tourism and therefore Income halted to HABs

1.4 Research identification and Objectives

The main objective of this research is to detect, monitor and report the occurrence of Harmful Algal Blooms (HABs) and Cyanobacteria in Lake Victoria. This is achievable through the following specific objectives:

1.4.1 Research objectives

1 To monitor the concentration of chlorophyl-a(Chl-a) concentrations from L8 OLI data as HAB proxies in L. Victoria from 2015 to 2021.

2 To monitor Lake Surface Air Temperature (LSAT) from L8 TIRS images as another HAB indicator in L. Victoria.

3 To develop automated Internet of Things (IoT) *in situ* system, applicable in near real-time to monitor and report geo-tagged Water quality data (e.g., LSAT) from 2021 onwards.

1.4.2 Research questions

The following questions are formulated with respect to aforementioned objectives:

1) Can space based observation systems be used to detect and monitor HABs in Lake Victoria Inland water body?
2) Does HAB occurrence have a direct impact on the LSAT and LSWT at their point of influence?
3) Can IoT be utilized to monitor HAB occurrences inland water Lakes in near real-time?

1.5 Study outline

This research study is divided into 6 chapters whereby the first chapter introduces the study by detailing into the background, motivation and problem statement, justifying the problem, objectives and research questions; Chapter 2 contains the reviewed literature which relate to this thesis. Further, Chapter 3 shows the data and methods used in the study with Chapter 4 highlighting the results for the findings from the methods. Chapter 5 discussed on the findings and finally, Chapter 6 concludes and recommends for future research that might not be addressed at this level of geoscientific exploration and expertise.

2. Literature review

2.1 A Synthesis on the Occurrence and Negative Impacts of HABs

During the last few decades (especially the last two decades), eutrophication of lakes as a result of HABs has been a common environmental problem observed in both freshwater (Jiang et al. 2015; Luo et al. 2016) and salty water bodies (Blondeau-Patissier et al., 2014), raising a global concern due to their ability to rapidly produce various hepatotoxic and neurotoxic substances (W. Song et al.,2015).This has been the case even at local water bodies like Lake Victoria in East Africa which showed a potential significance of algal blooms reported as early as 1980's (Ochumba, 1984).

Proliferation of algae causes eutrophication of freshwater lakes under certain environmental conditions and float on the water surface to cause abnormal watercolor amongst a variety of aquatic-related threats (Huisman et al. 2005; Liu and Yang 2012; Qin et al. 2016). These phenomena have been observed to be severely detrimental to aquatic environment in global inland water bodies which generally follows to the growing discharge of domestic or industrial wastewater as well as agriculture and fertilizer runoff (Glibert et al., 2005; Tang et al, 2006).

Notably, the ecological phenomenon has apparently been related to a variety of socioeconomic havocs like widespread occurrence of mass fish stocks and fish blooms (Ochumba, 1985, 1987; Caballero et al., 2020). Being actively photosynthetic, the blooms, not only deplete the dissolved oxygen (DO) in water bodies that result in mass death to aquatic lives but also release toxins, e.g., microcystin, that cause health risks to wildlife, livestock, humans and their companion pets who get in contact with the HABs (Hallegraeff, 1993), affecting the safety of drinking water supply (Guo 2007; Qin et al. 2010) leading to the disruption of the water food chain as it affects both trophic levels in the chain (Diaz and Rosenberg 2008).

Furthermore, surface foams or scums formed by CyanoHABs and their odorous compounds foul up water quality and surrounding recreational environment (Anderson et al., 2002).

2.2 Space-based Optical Remote Sensing in the Monitoring of Chlorophyll-a as an indicator of HAB event.

Numerous geoscience scholars have observed that space-based technologies are of great significance and importance in quantifying in detailed the spatiotemporal concentrations of HABs in inland water bodies.

Surface-level concentrations of chl-a, a proxy for HAB, has been quantitatively observed in aquatic ecosystems via the exploitation of optical remote sensing for many years (Clarke et al., 1970; Wezernak et al., 1976; Smith and Baker 1982; Gordon et al., 1983; O'Reilly et al., 1998; Tang et al., 2006; Gitelson et al., 2007; Allan et al., 2015; Watanabe et al., 2015; Concha and Schott 2016). This technique has routinely supported the generation of global chl-a concentration maps for the oceans for more than two decades (W. Song et al.,2015; Manuel et al., 2020).

A wide variety of heritage algorithms have been developed in that regard. For example, the blue-green band-ratio models have been used to quantitatively assess chl-a (Gordon et al., 1980; Bukata et al., 1995). However, in optically more complex inland and coastal waters, the color of water is further modified by the dominance of organic and inorganic particles, as well as colored dissolved organic matter (CDOM) (Mittenzwey et al., 1992; Han et al., 1994; Harding et al., 1994) that do not generally co-vary with chl-a.

Consequently, to improve the estimation of chl-a in these eutrophic and turbid environments, more sensitive methodologies have been proposed and implemented. For instance, sensors that image within the Red Edge region (690–715 nm) of the EM spectrum (Vos et al., 1986; Mittenzwey et al., 1992), when combined with red bands have manifested at a greater confidence to correlate well with chl-a in eutrophic and/or turbid waters (Munday and Zubkoff 1981; Gower et al., 1984; Khorram et al., 1987; Gitelson 1992; Rundquist et al., 1996; Gitelson et al., 2007).

The Operational Land Imager sensor aboard Landsat-8 which was commissioned into orbit in 2013, February to capture optical thematic features at 30m spatial resolution which is relatively high (Irons et al., 2012), has offered significant space observatory improvements in both data quality and quantity (i.e., both spectral and spatial

coverage) over other previous heritage optical sensors (Markham et al., 2014; Pahlevan et al., 2014; Markham et al., 2015) and is relatively capable of monitoring bimonthly HAB dynamics (Pahlevan et al., 2014; Allan et al., 2015).

Several methods have been developed to retrieve chl-a from the four OLI visible bands (Markham et al., 2015; Watanabe et al., 2015; Concha and Schott 2016; Manuel et al., 2020).

2.2.1 Retrieval of Chl-a from Landsat 8 OLI Optical Spectral Bands

O'Reilly et al. (1998) observed that for satellite missions that do not inherently support measurements in the Red Edge portion of the EMR Spectrum, chl-a algorithms tend to alternatively rely on blue-green band ratio algorithms coupled by some complex ML approaches. This is true even for Landsat-8 (Le et al., 2013; Markham et al., 2015; Freitas and Dierssen 2019).

Most research on chl-a retrieval has heavily relied on missions like the MERIS, inherently designed to image in the RE portions (Gitelson 1992; Gower et al., 2005; Gitelson et al., 2007); These algorithms have however been noted not to score well to missions that are limited to such measurements (RE) like MODIS (Esaias et al., 1998), VIIRS (Wang et al., 2014), Geostationary Ocean Color Image (GOCI) (Ryu et al., 2012) and Landsat 8 (Snyder et al., 2017) which lack the RE spectral bands. Thus, the adoption of Empirical standard algorithms of the Ocean Color (OC) spectral band family for chl-a estimation e.g., OC2, OC3 (Neil et al., 2019) which has been proven to improve chl-a retrieval from OLI products in highly eutrophic and even clear aquatic regions (Cao et al., 2020) like Kisumu Bay.

The Ocean Color 2 regional algorithms specific to OLI spectral imageries have been successfully implemented in lakes and reservoirs (Allan et al., 2015; Watanabe et al., 2015; Snyder et al., 2017) for chl-a retrieval and would be significantly ideal for this research study.

2.3 Capabilities of Space-based observations in the Monitoring of Lake Surface Air Temperature (LSAT) as an indicator of HAB event.

With the advent of thermal imaging remote sensing, it has been observed that Sea Surface Temperature (SST) observations derived from space borne sensors like the Advanced very high-resolution radiometer (AVHRR) in conjunction with optically derived chl-a data have great potential to detect and monitor HAB at an improved accuracy (Tang et al., 2003a; Thomas et al., 2012).

Thermal infrared bands are potentially capable of quantifying the amount of infrared radiant heat emitted from aquatic surfaces that have environmental and economic import (Anderson, J. et al., 1984; Haakstad, M. et al., 1994; River S et al., 2004).

It has been noted that satellite derived Sea Surface Temperature (SST) data are often used in combination with chlorophyll-a to relate to bloom events (Villareal et al., 2012) at a higher accuracy. This has also been observed in inland water bodies where Lake Surface Air Temperatures (LSAT) are taken into considerations (Thomas et al., 2012; Shi and Wang, 2007). Peñaflor et al. 2007 also examined the correlation of seasonal phytoplankton bloom in the Luzon Strait off the Philippines using MODIS chl-a and SST, in addition to Quick-SCAT wind data. Similar findings were also documented by Binding et al., 2012, 2013 by using MERIS MCI in particular by focusing in the FLH and FAI, as well as other indices, used in conjunction with satellite estimates of Chl-a and SST to improve our image analysis in a very confident way.

Blooms of blue-green algae are reportedly often associated with increase in water temperatures (Hutchinson, 1967). During a bloom observed in Lake Victoria, the average Lake Surface Air temperature of the open waters rose from (23.9-24.8°C) to (26.9–31.5°C) in May of 1986 within two weeks of the bloom period (Kilham, 1991; Gasse, Talling & Kilham, 1994).

2.4.1 A Review of the Potential of Smart IoT Solutions in Location-based Water quality monitoring.

The advent of smart solutions and Internet of Things (IoT) has lately shown an escalating curve in their great capability to monitor water quality particularly with advancement in communication technology (Jan et al., 2021). This has greatly improved the capabilities to remotely gather and further disseminate the in-situ water parameters e.g., Lake Surface Air Temperature (LSAT) (Cloete et al., 2016), which are correlated with an algal bloom event in inland water bodies.

In this regard, legislations have been passed through the relevant government and non-governmental agencies for example the Kenya Marine and Fisheries Research Institute (KMFRI) and African Great Lakes to set thresholding standards in water quality parameters that relate to HAB events. These parameters include Lake Victoria Surface Air Temperature which is set to vary with the proximity from the shore, (25.5 +/- 5°C), Suspended solids (30 mg/L), Pathogens and bacteria (Nil/100ml), Fluoride (1.5mg/L), Total dissolved solids (1200mg/L), Ammonia (0.5 mg/L), Nitrates (10mg/L) (Budyko, M. I. 1974; Yin & Nicholson, n.d.), among many other water quality parameters.

In-situ Water quality monitoring is defined as the collection of information at set locations and at regular intervals in order to provide data which may be used to assess current conditions and even establish trends etc. (Yang et al., 2011; Encinas et al., 2017). This can be used to ascertain the abnormalities in the pre-set standards or provide early warning identification of hazards (Raju et al., 2017)

The proposed power efficient, simple solution *in-situ* monitoring system works to provide near real-time analysis and dissemination of LSAT collected. The system is also designed to connect to a remote user and provide them with an alert based on the pre-configured thresholds. E.g., Lake Water Authorities who are remotely located in the offices when there is a significant deviation of LSAT from the pre-defined set of standard values which is associated with a bloom.

Integration of various methodologies such as the Risk Quotient Approach to show color coded hazards and each hazard level for each parameter assessed at each

location e.g., Surface temperature at a given location mapped using Geographical Information Systems (Wan M. et al, 2019) further cuts-off visualization and interpretation efforts.

Location-based approach in assessing lake and water resource pollution has been implemented by Waspmote in Georgia, USA (Waspmote). Sample points were established using field survey, laboratory analyses and geospatial techniques to monitor the various water quality parameters. The results were tables, graphs and maps showing the concentrations of the parameters.

Boddula et al. (2017) proposed a wireless sensor system, CyanoSense, which provided a low footprint, low power and low-cost solution for the monitoring algal bloom remotely in Lake Oconee, Georgia, USA.

A Smart GIS-based in-situ model for assessing aquifer vulnerability was implemented in Kakamigahara Heights, Gifu Prefecture, central Japan (Babiker et al. 2010)

A GIS-based emergency response system for sudden water pollution was developed in China. (Ma, Xu, and Wang 2014; Zhang 2014). It uses GIS technology and a hydraulic water quality model to represent the levels and extents of pollution. Maps showing the spatial distribution of the sample points are generated.

2.4.2 Importance of IoT concept in this Scope

Beyond providing ground truth data especially for validation of the remotely sensed data (LSAT in this regard), the internet of things:

I. Helps the geoscientific world to work and generally live smarter at large.

II. On the same note, the automated IoT system provides the GIS and Remote sensing researcher with complete control over the entire monitoring with little to no human-to-man or human-to-machine interaction and to some extent, even prediction.

III. In addition to automating the whole process, IoT essentially comes in by providing a near real-time monitoring of how the entire system operates. By sending the near real time GPS position of the entire system and housekeeping data, this saves us the anguish arising from unknown state of affairs.

IV. Consumption of IoT inherently reduces the manual processes involved, thereby cutting down labour costs to be incurred. The amount, in form of money, time and labour that would otherwise be set aside by the authority e.g., the Kenya Marine and Fisheries Institute – KMFRI for physical field collection of that Ground Truthing and validation data is saved on.

V. All said and done, IoT stands out as a few of the chief principal technologies of everyday GIS and Remote sensing research life especially where live and running data is a requirement.

3. Materials and methods

3.1 Study area

Lake Victoria, with an extensive surface area of about 68,800 KM^2 and an average depth of 40m at a maximum depth of 79m ranks the second largest fresh water lake in the world after Lake Superior and the largest in Africa. Lying between 3° S to 0° 30`N latitude and 31° 40`E to 34° 50`E longitude is distributed among these three East African countries viz Tanzania 51%, Uganda 43% and Kenya the remaining 6% (africangreatlakes.org).

That in place, the lake is privileged to serve as economical home of about 40 million residents (Calamari et al.,1995, africangreatlakes.org) in those riparian reserves. These millions of individuals solely bank on the lake for all aspects of their daily economic livelihood ranging from, fishing, agriculture, and industrial applications just to barely highlight but a few. On that regard, it's ecological monitoring should be of great geoscientific interest.

Being located in Equatorial regions of the globe, the lake has an alternating climatic condition varying from tropical rain forest with rainfall over the lake for a better portion of the year to a semi-dry climate with sporadically discontinuous droughts over some locations.

This provides ambient temperatures varying between 18-26°C which therefore provides an optimum host condition for the growth and development of the *Cyanobacteria spp.* in this scope (Okello et al., 2011).

Figure 5 shows the location and extent of the extract of the study area particularly relevant to this study.

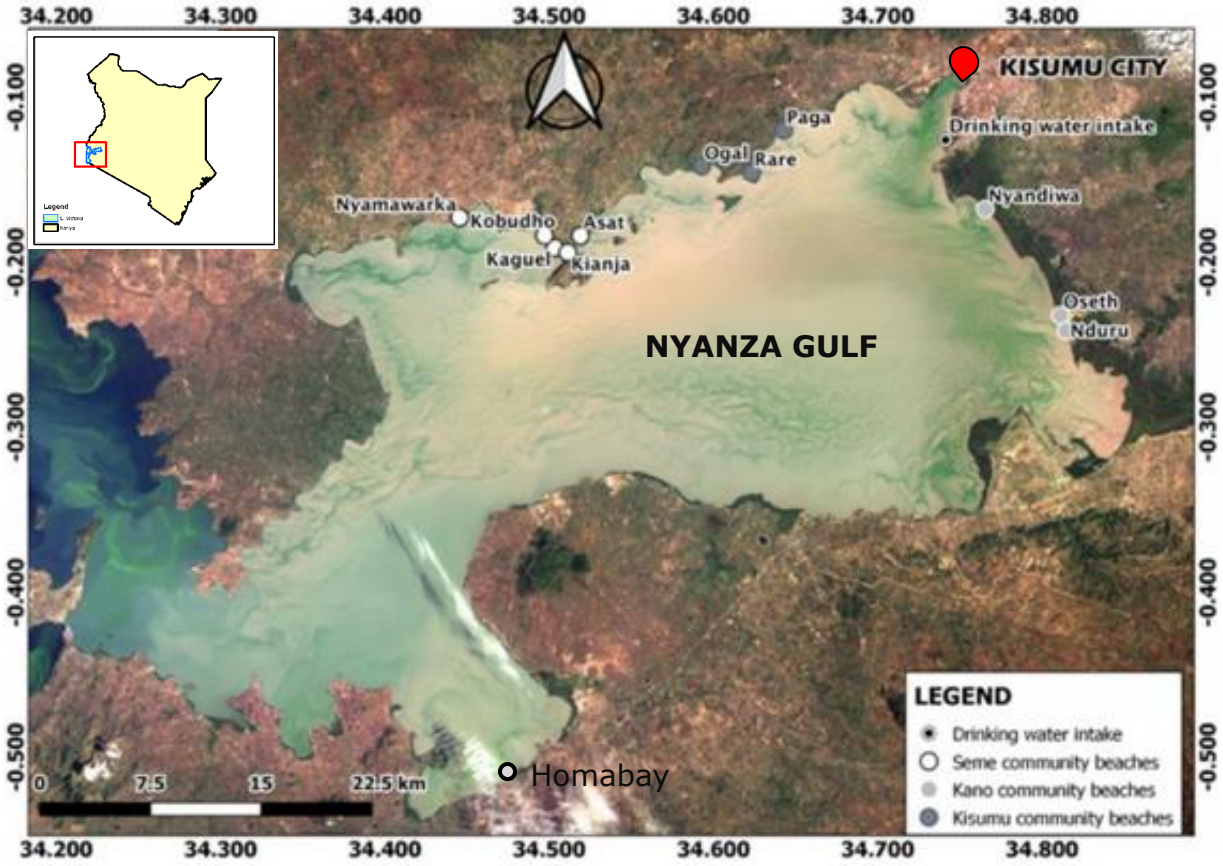

Figure. 5 Map of Winam Gulf with study sites.

3.2 Data

Five (5) datasets are exploited in this study as displayed in table 1 below.

1. Landsat 8 Operational Land Imager Sensor from where spectral bands 2 to 5 centered on Blue $_{(480\ nm)}$, Green$_{(560nm)}$, Red $_{(655\ nm)}$ and NIR $_{(865\ nm)}$ was used in the extraction of Chl-a.
2. Thermal data extracted from the L8 TIR sensor from which TIR1$_{(10,895nm)}$ band10 was used for the extraction of LSAT.
3. In-situ field data form the KMFRI department aided in validate the remote sensing processes by providing the dates which blooms occurred.
4. ESRI GIS Shapefiles were used for delineation of study areas
5. In-situ data from the locally deployable autonomous IoT sensors.

Data Type	Source	Role/Use
Landsat 8 OLI (30m, 16 days)	United States Geological Survey USGS (from 2015-2021)	Spatiotemporal HAB Monitoring
Landsat 8 TIR (100m, 16 days)	United States Geological Survey USGS (from 2015-2021)	Lake Surface Air Temperature Monitoring (LSAT)
Meteorological Data	Kenya Marine & Fisheries Research Institute-KMFRI (2015-2021)	Water Quality assessment
Shapefiles	Geodatabase of Global Administrative areas- GADM	Delineate the Study area
In-Situ Data	In-situ Sensors 2021 Onwards	Continued In-Situ Algal Monitoring

Table 1: Data Sources and their roles

Tool/Material	Role	Availability
Google Earth Engine (GEE)	Download and Processing	Freely Available
QGIS, R, Python	Further Analysis & Maps	Free
Microcontroller & Sensors	In-Situ data Monitoring	Local Purchase

Table 2: Tools and Materials used in the study

3.3 Methodology

3.3.1 Chl-a Estimation from Landsat 8 OLI

3.3.1.1 Relevance of Landsat-8 in Chl-a estimation

The L8 satellite supplies multispectral images comprising of 11 spectral bands, 9 in the optical imaging region and 2 in the thermal region (USGS). Chl-a is known to have distinctly prominent absorption behavioral patterns between certain wavelengths of the visible portion of the EM spectrum within which this space vehicle images. Notably, around the blue (450 - 475 nm) and red (650-675nm) regions, chl-a exhibits high absorption tendency. Further, at the green and NIR regions of the EM spectrum, Chl-a exhibits high reflectance values that could reach 500 and 700 nm, respectively (Beck et al., 2016; Tuuli et al., 2020). This information has been widely exploited by researchers to develop chl-a quantification algorithms, (Richard et al., 2018) and this study is not an exception. The presence of narrower bandwidths and at a finer spatial resolution of 30m improves L8's pigment discrimination ability even in water bodies despite the fact that it was purposely designed for terrestrial applications (Pahlevan et al., 2014). Further, the second Thermal InfraRed (TIR) sensor on board L8 provides for the measuring of the Lake Surface Air Temperature which is another significant proxy for HABs that is investigated in this study.

3.3.1.2 Satellite Data Preprocessing

Landsat-8 OLI level-1 collection-1 data product consists of quantized and calibrated scaled DN values. Retrieval of chl-a from Landsat-8 sensor over the study region involved the following three steps,

(i) Obtaining absolute scaled DN values for all the required bands (B2, B3, B4 and B5).

(ii) The images were then subjected to atmospheric correction to minimize the atmospheric attenuation effects in the quite humid Lake Victoria regions.

(iii) Then finally, chl-a were retrieved by utilizing Ocean Chlorophyll (OC) algorithms (OC-2 shown in equation 1 below).

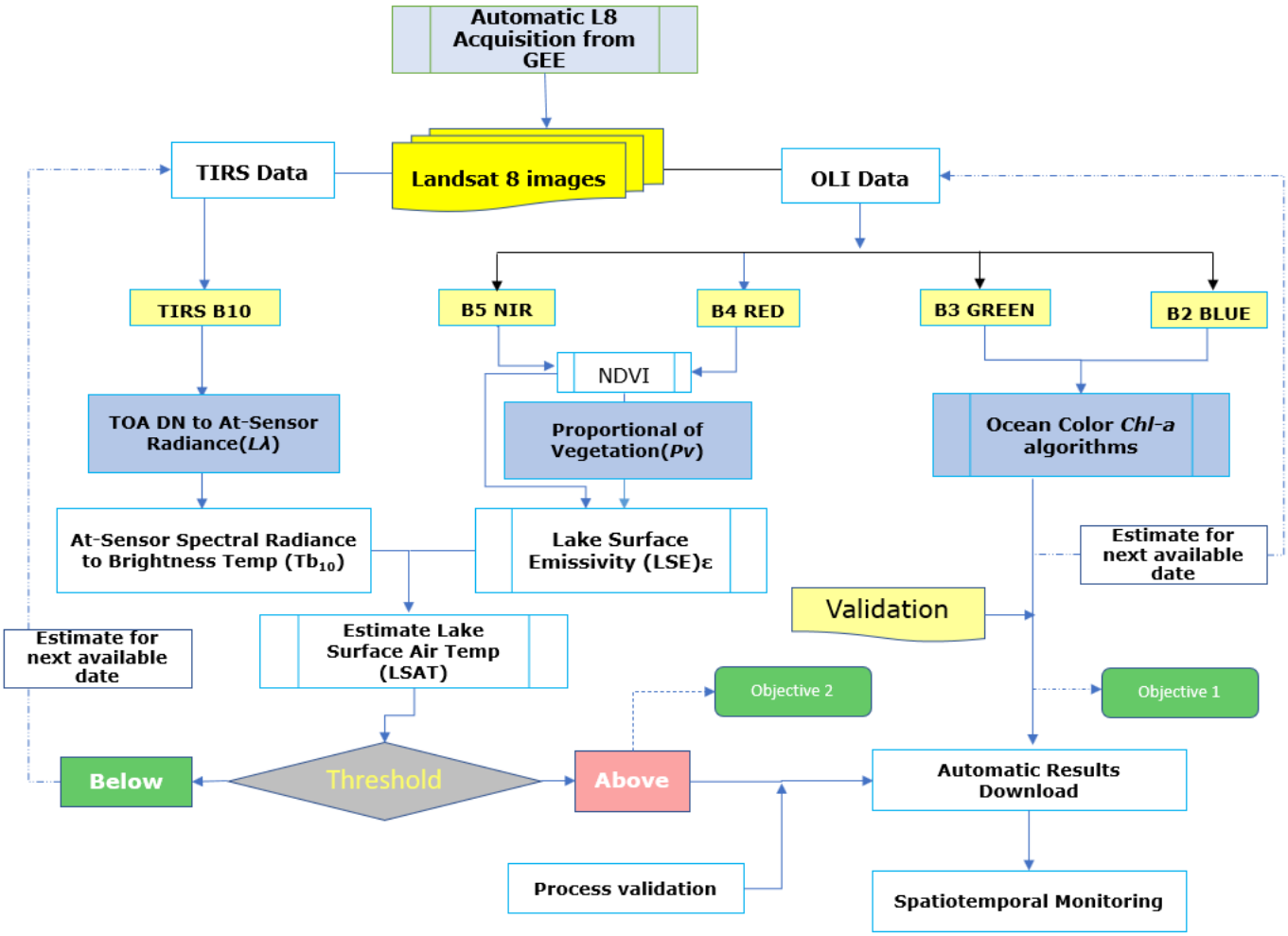

Figure 6: Overall Methodology Workflow for Objectives 1 and 2

3.3.1.3 Estimating Chl-a Concentrations.

Based on band positioning in the OLI sensor and previous good performance in retrieval of chl-a in inland waters (Augusto-Silva et al., 2014) Ocean Color 2 (OC2) algorithm was settled upon for this retrieval purpose.

OC-2 is a modified fourth order polynomial algorithm which was originally developed for the SeaWiFS data (O'Reilly et al., 1998) but can be tuned to perform relatively well with Landsat 8 (Watanabe et al., 2017; Richard et al., 2018)

Mathematically, OC-2 is an empirical two-band ratio algorithm (NASA 2011) for the quantitative estimation of chl-a and is given by:

$$Ch - a = 10^{(0.2511 - 2.0853\mathrm{R} + 1.5035\mathrm{R}^2 - 3.1747\mathrm{R}^3 + 0.3383\mathrm{R}^4)} \quad \text{..... (eq. 1)}$$

Where $\mathrm{R} = \log_{10} {\mathrm{Rrs}(490)}/{\mathrm{Rrs}(555)}$ ..... (eq. 2)

Where:

Ch-a is the estimated quantified value of chlorophyll-a

R is the quantifying coefficient derived from the spectral bands

Rrs is Reflectance in the specified wavelength region

For OC-2, ratio of Rrs at Blue band at wavelength 490nm and Rrs Green band at 555nm are used to estimate the chl-a as reported in the following dates.

Year	Date and Month	Reporting body
2015	12th January, 22nd February	Nasa Earth data, KMFRI
2016	23rd February	KMFRI
2017	04th September	Africa great Lakes
2018	27th January	KMFRI, Nasa Earth Data
2019	18th August	KMFRI
2020	29th August,	KMFRI
2021	No Data	None Reported

Table 3: HABs reported in Lake Victoria, (KMFRI, RCMRD, NASA Earth Data)

The results from the chl-a estimations from this algorithm are as presented in section 4 of this write-up.

3.3.1.4 Accuracy Assessment

To validate the obtained chl-a values, chl-a validation dates and data were retrieved from KMFRI and some from NASA Giovanni Chlorophyll-a portal, managed by Nasa Earth Data. This portal has been for so long a rich archive for global chl-a generated from MODIS, Sentinel-3 OCLI, SeaWiFS just to mention but a few. Chl-a for the study area for years 2015 through 2021 were therefore acquired from the portal through Google Earth Engine. The spatially enabled data was then used to assess the accuracy of the estimated chl-a from Landsat 8 OLI.

The statistical metrics used for validation was coefficient of determination (R^2) as given by the equations 3 below.

$$R^2 = \frac{\sum\left(Modeled_{Chl-a,i} - Mean_{Chl-a,i}\right)^2}{\sum\left(Actual_{ActChl-a,i} - Mean_{ActChl-a,i}\right)^2} \quad \text{..... (eq. 3)}$$

where $Modeled_{Chl-a,i}$ is the estimated Chl-a at point *i*, $Mean_{chl-a,i}$is the mean of the modeled chl-a from Landsat 8 at point *i*, $Actual_{ActChl-a,i}$ represents the reference chl-a from Sentinel-3 OLCI at points *i*, while $Mean_{ActLChl-a}$ is the average of the sampled actual chl-a from Sentinel-3 OLCI sensor.

3.3.2 Lake Surface Air Temperature Estimation from Landsat 8 OLI

This part of the methodology workflow mainly intends to estimate the LSAT as another proxy of HABs (Owen M et al., 2017)

3.3.2.1 Relevance of Landsat-8 in retrieval of LSAT

Thomas et al. (2012), after extensive research in the aquatic environments, observed that certain thermal conditions favor the growth, stability and spread of a wide variety of algal species. Further, after their domination and colonization, these microscopic and photosynthetically active organisms increase the LSAT of the region.

L8 TIR supplies two thermal bands (10 and 11) for the retrieval of LSAT. Wang et al. (2015) proposed the Mono-Window Algorithm for the retrieval of Land surface Temperature using the Landsat 8 band 10. Moreover, this algorithm scores equally acceptable when applied in the retrieval of LSAT in inland water bodies, hence the adoption of the same in this study. It should however be noted that data from the Landsat 8 TIRS Band 11 have large uncertainty and Wang et al., (2015) suggested the use of TIRS band 10 data as a single spectral band for LSAT estimation; for this reason, present work provides the use of L8 band 10 alone.

The information acquired is processed in a multi-software platform, in order to calculate the lake surface air temperature. For this purpose, R programming environment and ArcGIS 10.8 were used.

For the successful retrieval of LSAT, the study includes the major following steps as indicated in the workflow presented in (Figure 7). That is: -

- Conversion of L8 DN pixel values to at-sensor spectral radiance ($L\lambda$)
- Transformation of at-sensor spectral radiance to at-sensor brightness temperature (Tb_{10})
- Estimation of Lake Surface Emissivity
- Estimation of Lake Surface Air Temperature by adopting MW algorithm (Wang et at., 2015).

3.3.2.2 Retrieval of Lake Surface Air Temperature

3.3.2.2.1 TOA spectral radiance ($L\lambda$)

The first step is to convert the raw DN (Digital Number) values of band10 to obtain the TOA (top of atmosphere) spectral radiance ($\mathbf{L\lambda}$) by multiplying the multiplicative radiometric rescaling factor (**ML**) of TIR bands with its corresponding TIR band and adding additive rescaling factor ($\mathbf{A_L}$) using equation (4) below

$$\mathbf{L\lambda} = \mathbf{ML} * \mathbf{QCal} + \mathbf{A_L} \text{ ---------- (4)}$$

Where:

$\mathbf{L\lambda}$ is the spectral radiance in $^{\text{watts}}/_{(m^2\ srad^{-1}\ um^{-1})}$

ML is the $Band_{10}$ multiplicative rescaling factor obtained from metadata (e.g.,0.0003342);

QCal is the DN value for the quantized and calibrated standard product pixel of band 10.

$\mathbf{A_L}$ is the band-specific additive rescaling factor obtained from the metadata (0.1);

3.3.2.2.2 Brightness temperature (Tb_{10}) in ˚C

As Latif 2014 puts it, brightness temperature is the Electromagnetic Radiance travelling upward from the top of the Earth's atmosphere. Therefore, The Brightness Temperature is not a temperature on the ground rather is the temperature at the satellite (M Z Dahiru et al., 2020). The spectral radiance value $\mathbf{L\lambda}$ obtained in equation (4) above, was then converted to brightness temperature BT by adopting equation (5) below

$$Tb_{10} = \frac{K2}{\text{Ln}\left[\left(\frac{K1}{L\lambda}\right)+1\right]} - 273.15 \text{ ---------- (5)}$$

Where:

Tb_{10} is the brightness temperature;

K1 and K2 are thermal constants, obtained from the metadata file of the L* OLI;

Lλ is top of atmospheric radiance.

3.3.2.2.3. Normalized Difference Vegetation Index-NDVI

This mathematical algorithm which ranges between -1.0 to +1.0 is essential to identify different land surface cover types of the study which is further necessary to calculate proportional vegetation (Pv) and Lake Surface Emissivity (ε) (M Z Dahiru et al., 2020). NDVI is calculated on per-pixel basis as the normalized difference between the red band (0.64 - 0.67μm) and near infrared band (0.85-0.88μm) of the images using the formula in equation (6) below.

$$NDVI = \frac{NIR-RED}{NIR+RED} \text{ ---------- (6)}$$

3.3.2.2.4 Proportion of Vegetation (Pv)

The NDVI for the study area derived from equation (6) above was then used to estimate the thermal emitting target under each lake surface cover type denoted as proportional vegetation which is a direct function of NDVI.

The vegetation and bare soil proportions are acquired from the NDVI of pure pixels.

Pv was calculated using the equation (7) below.

$$Pv = \left[\frac{NDVI - NDVImin}{NDVImax - NDVImin}\right]^2 \text{ ---------- (7)}$$

3.3.2.2.5. Lake surface emissivity (ε)

This is the radiative properties of lake features/targets which characterizes the ability of a body to emit thermal radiation energy across the lake surface into the air atmosphere (Rhinane et al. 2012). The knowledge of lake surface emissivity is further exploited to estimate lake surface air temperature as shown in equation (8) below

$$E = 0.004 * Pv + 0.986 \text{ ---------- (8)}$$

3.3.2.2.6. Lake Surface Air Temperature (LSAT)

We finally calculate the retrieval of LSAT using brightness temperature (BT) obtained from band 10 and Lake Surface Emissivity derived above.

LSAT has a linear relationship with at sensor brightness temperature and Lake Surface Emissivity and can therefore be retrieved using the equation (9) below:

$$\text{LSAT} = \frac{\mathbf{Tb}10}{\{1 + [(\boldsymbol{\lambda}\mathbf{Tb10})]\rho * \text{Ln}(\varepsilon)\}} \text{ ---------- (9)}$$

Were,

LSAT is the LSAT in Celsius (˚C),

$\mathbf{Tb}10$ is at- sensor BT (˚C),

λ is the average wavelength of band 10,

ε is the emissivity calculated from equation (v) above

This methodological workflow algorithm finally derives Lake Surface Air Temperature for the region under study and the results were as presented in section 4 of this write-up.

3.3.2.2.7 Accuracy Assessment

To validate the estimated LSAT values, validation data was retrieved from NASA's MODIS which collects global Sea Surface Temperature on a near-daily basis as one of its products. Lake Surface Air Temperature for the study area for years 2015 through 2021 were then used to assess the accuracy of the derived LSAT.

The statistical metrics used for validation was the coefficient of determination (R^2) as given by the equations below

$$R^2 = \frac{\sum\left(Modeled_{LSAT,i} - Mean_{LSAT,i}\right)^2}{\sum\left(Actual_{LSAT,i} - Mean_{ActLSAT,i}\right)^2} \text{ ..... (eq. 10)}$$

where $Modeled_{LSAT,i}$ is the estimated LSAT at point *i*, $Mean_{LSAT,i}$ is the mean of the modeled LSAT from Landsat 8 at point *i*, $Actual_{LSAT,i}$ represents the reference LSAT at points I, while $Mean_{ActLSAT,i}$ is the average of the sampled actual LSAT from MODIS.

3.3.3 Automated In-situ Internet of Things System

The automated *in-situ* system consists of a variety of hardware and software that work in conjunction to collect and disseminate LSAT remotely.

The internet of things, conveniently referred to as IoT, is a network of interconnected computing devices, mechanical and digital machinery or people having unique identifiers (UIDs) and the ability to transfer data without requiring human-to-human or human-to-computer interaction. (2019, Alexander S.)

Narrowing down into the specifics of this research project, the word IoT has been used to refer to a man-made system that has been assigned a unique and private Internet Protocol (IP) address, loaded with sensors and is able to transfer data e.g.,

- Lake Surface Air Temperature
- Lake Surface Salinity
- GPS Location of the system at the time of the collection of such data
- The whole system in-house status including working condition
- Unexpected system failure
- System power level.

All these data are packaged into one socket and transmitted over a wireless network between the system and a remote monitoring computer.

3.3.3.1 Devices and Sensors Used

At least the following components and sensors have been used in the achievability of collection of the In-Situ data:

1. Microcontroller (MCU):

This can be loosely defined as an integrated circuit that contains a brain or rather the microprocessor that has a memory and integrated circuits and that has the potential to control the functionalities of an embedded system or electronic device (such as our Water Quality Monitoring system). There do exist a vast variety of Commercial Off The Shelf-(COTS) MCUs e.g.,

- ATmega (Arduinos)
- ESP32
- BCM
- Single Board Computers (SBC) e.g., Raspberry Pi
- ARM Cortex e.g., STM32

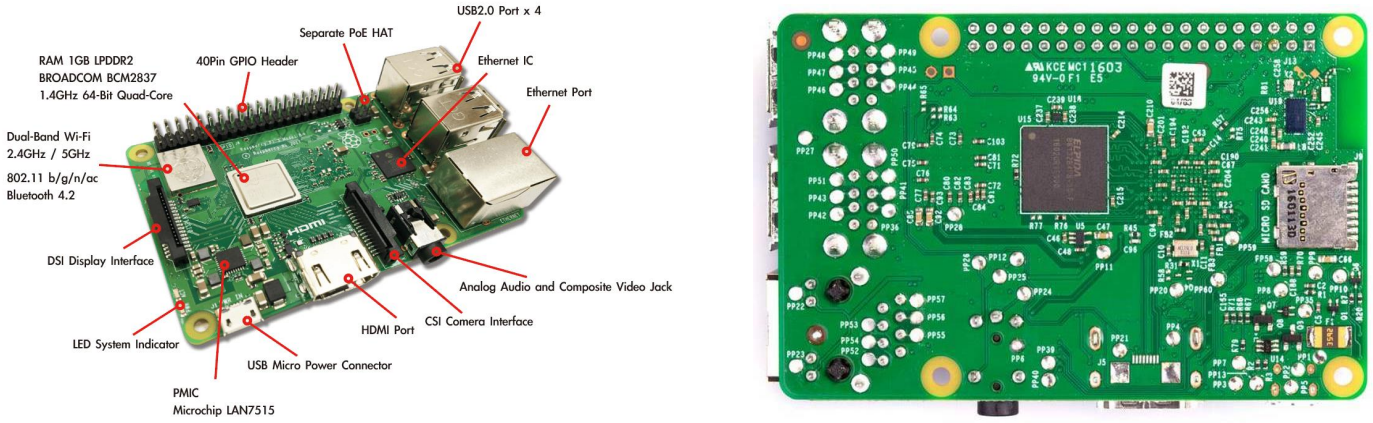

Figure 7: Raspberry Pi model 3B+ used in this project Front and Back views

3.3.3.1.1 Reason for choice of Raspberry Pi in this project

1. **Powerfulness:**

Having the capacity to perform relatively similar to a full-blown computer, it ranks the highest at the time of submission of this thesis by being able to simultaneously run multiple tasks at a go and further operate as a server gives it more potential score and rating.

2. **Little to no electronics prerequisites needed**

Little prior embedded programming languages knowledge and its components is needed here. Unlike for Arduino and other ARM Cortex MCUs, you definitively do not need a very good and deep electronic background, and therefore no need to know about embedded programming languages e.g., Object oriented C++.

3.3.3.2. Sensors Used

At least the following sensors have been integrated and used as part of the in-situ data collection system.

3.3.3.2.1 Lake Surface Air Temperature (*DHT11 Sensor*)

The DHT11 is a widely accepted relatively cheap, a digital and easy-to-use air temperature sensor.

3.3.3.2.1.1 Features of DHT11 Air Temperature Sensor

- Measures air temperature within the ranges 0-50°C with an accepted accuracy of ±2 °C.
- The sensor need be interfaced with a filtering capacitor and a pull-up resistor of 10KΩ connected in series with the VCC output pin of the sensor
- Input voltage is ~3.3 V to 5.0V

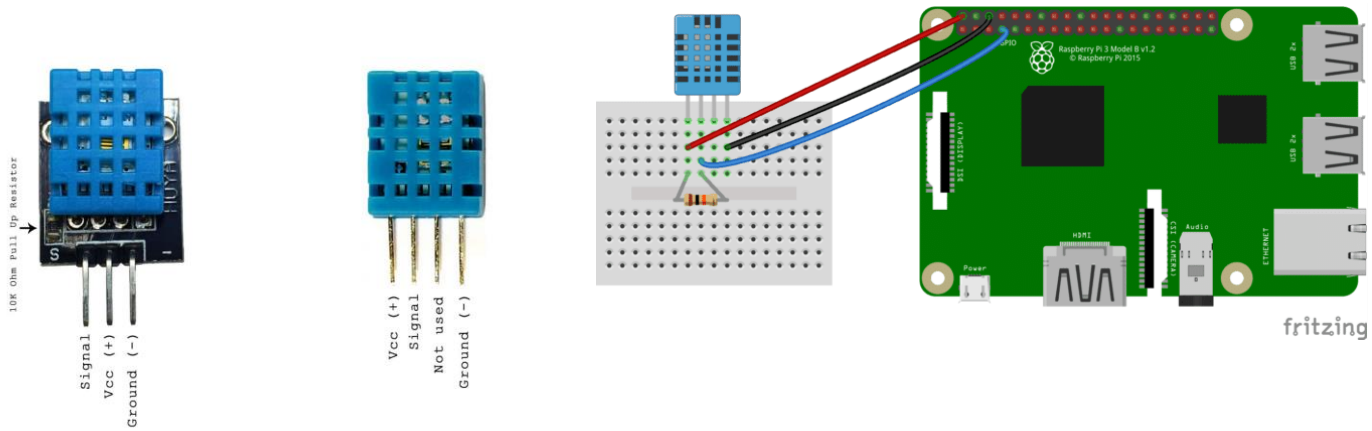

Figure 8. DHT 11 Sensor module and how it hooks up to the Raspberry-pi Microcontroller

3.3.3.2.1.2 Mode of operation of DHT11 Air Temperature Sensor

The DHT11 basically measures the Air Temperature under the concept that, with increase in the air temperatures, the resistance of the Negative Temperature Co-efficient, NTC thermistor installed in it decreases, hence recoding higher temperature values (Electronics Lovers, 2020)

3.3.3.2.2 Global Positioning System (GPS) Sensor.

The *Neo-6M- Ublox GPS* sensor Module was settled on to solely play the role of collection of the (x, y) coordinate location information where the whole system is located at any given time of collection and dissemination of LSAT and other data.

3.3.3.2.2.1 Basic mode of Operation

A basic GPS signal receiver capable of collecting the GPS pseudo frequency signals broadcast at 1575 MHz from GPS space segments and computationally determine the location of any system that it's embedded to. This will tell us precisely where in Lake Victoria there exists abnormality in temperature rises, calling for an emergency mitigation action.

Figure 9 The Neo-6M GPS Sensor Module

3.3.3.2 Working and Operation of the Automated IoT System.

An IoT ecosystem integrates an array of smart devices that use embedded systems technology, such as: -

1. In-built processors for example Intel, BCM (for our case).
2. An array of sensors (as previously discussed in section 3.3.3.2 above) that collect and send the data from their point of attachments.
3. Communication hardware; that disseminates the collected data.

Upon establish a secure connection with the IoT gateway device, the sensors will be able to share the sensor data they collect where data is either sent to the cloud server as commonly seen in IoT systems that run on Raspberry Pi (using POST, GET functionalities) for analysis or the sensor data can just be locally analysed. Our IoT system and functionality are fully automated with little to no human intervention as illustrated in the *Fig. 10* below.

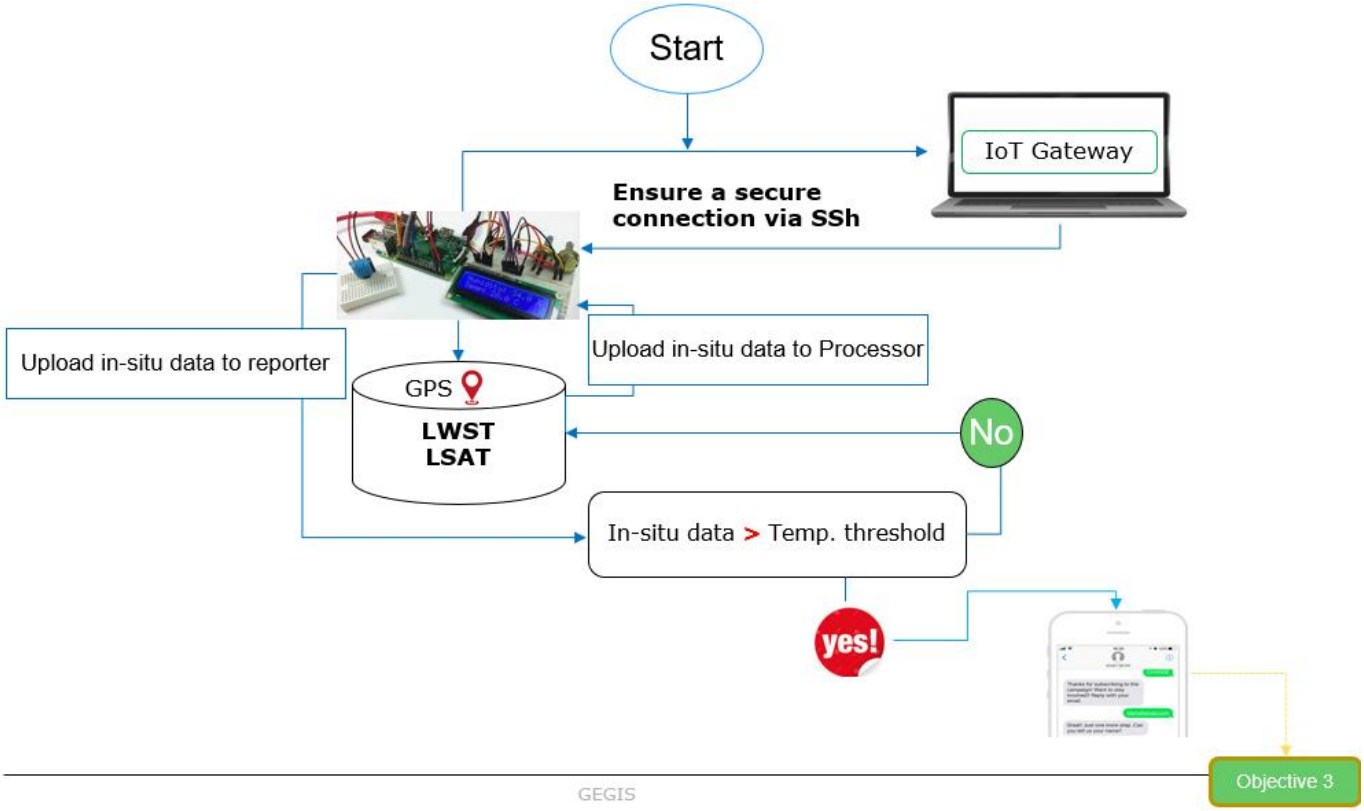

Figure 10: Overall Methodology Workflow for the IoT system.

4. Results

4.1.1 Chl-a concentration Maps

OC-2 algorithm was proposed and used to estimate the HAB concentrations as reported. The following displayed results were obtained.

In figure 11 below, it is noted that on the event of HAB as reported, Chl-a values shoot significantly to up to 31.3 Mg/M^3 in the affected regions as seen in red patched areas. This was unlike the case in the *blue* regions in the western parts on the study area as shown below. They(the western parts of the study area) reported very low chl-a score around 1.2 Mg/M^3 especially for a non-turbid water body like Lake Victoria.

Figure 12 below as well indicates that HAB that occurred especially on the North-Eastern parts of the study area could be well explained by the high chl-a values as reported by the OC-2 algorithm.

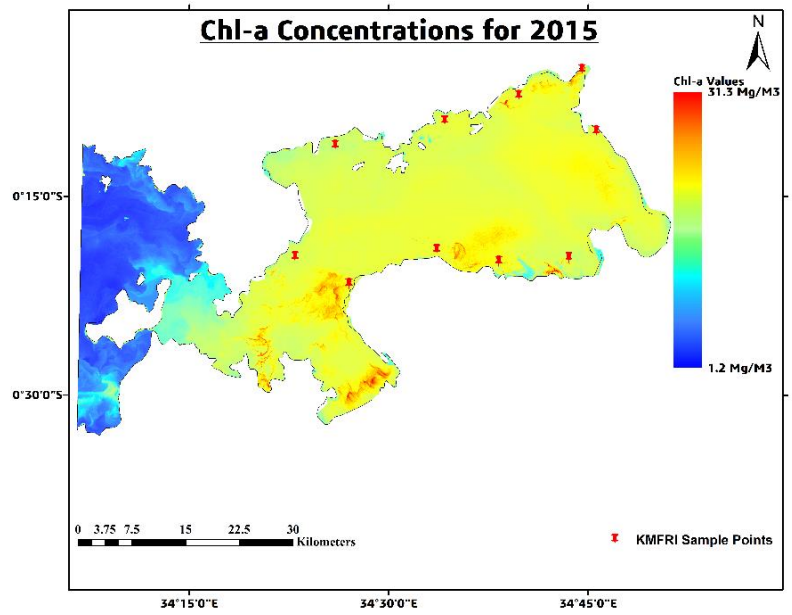

Figure 11. HAB concentration, 2015

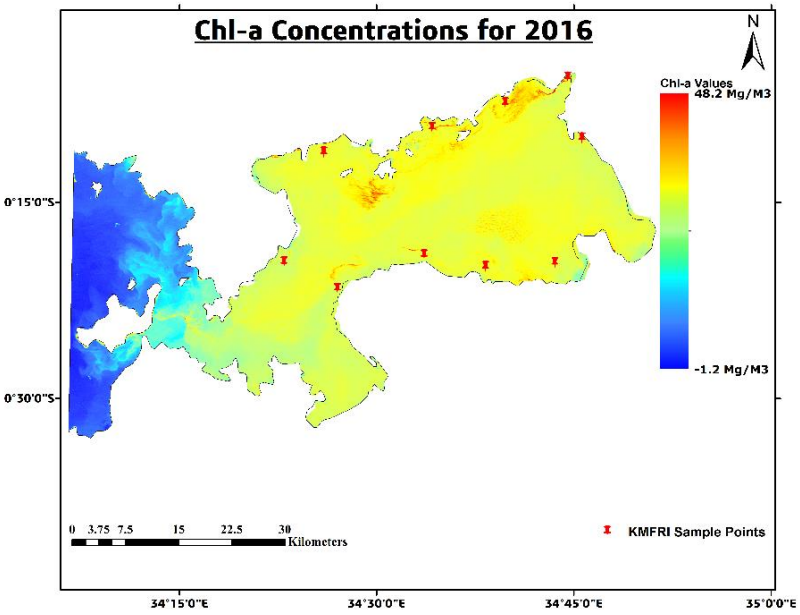

Figure 12. HAB concentration, 2016

In Figures 13 and 14 below, it is noted that on the event of HAB, Chl-a values are seen to rise significantly to up to 39.6 Mg/M^3 and 25.9 Mg/M^3 in the affected regions Notably, in 2018, the severity of the bloom was quite low as the highest chl-a values were 25.9 Mg/M^3.

The western regions of the study area still maintain comparatively low chl-a values indicating they were not significantly affected.

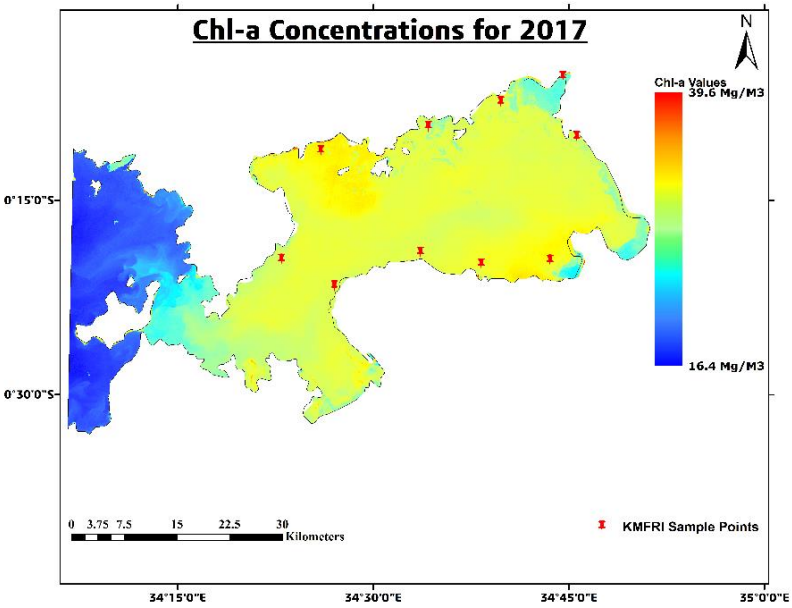

Figure 13. HAB concentration, 2017

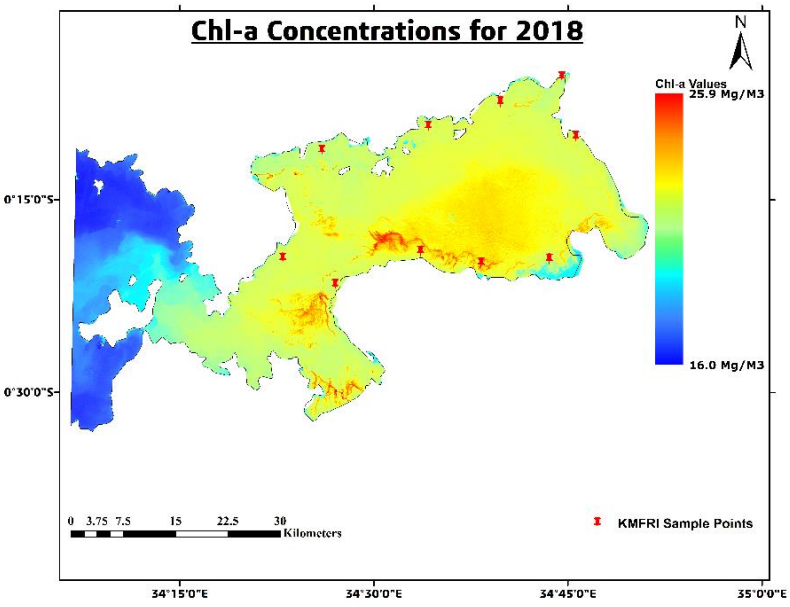

Figure 14. HAB concentration, 2018

Finally figures 15 and 16 for 2019 and 2020 below, show that the severity of the HAB went relatively high. This was realised with the increased eutrophication in the said region as Simiyu et al., 2018 envisioned. This research therefore agrees with the previously well-documented studies. Chl-a values are seen to rise significantly to up to 49.3 Mg/M^3 and 57.1 Mg/M^3 in 2019 and 2020. The cases were more severe.

The western regions of the study area still maintain comparatively low chl-a values indicating they were not significantly affected.

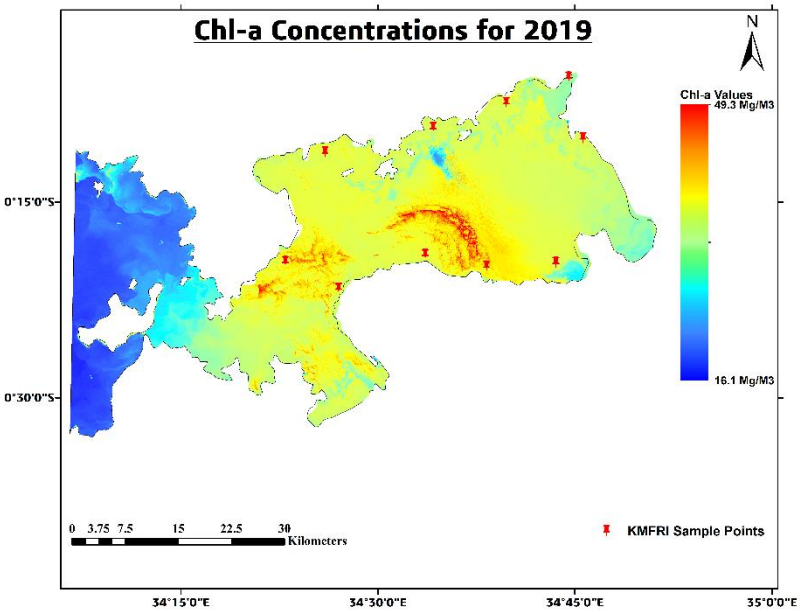

Figure 15. HAB concentration, 2019

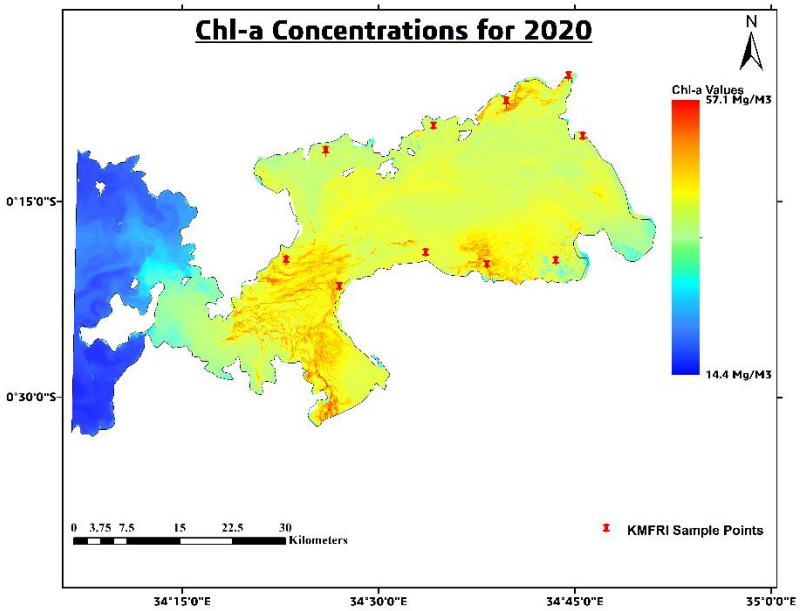

Figure 16. HAB concentration, 2020

4.1.2 Accuracy assessment of the Chl-a Estimates.

As a general rule of remote sensing and quantitative estimations, the chl-a estimates had to be validated. In order to validate the estimated results and analyze the relationship between Sentinel-3 OLCI chl-a data and those retrieved from Landsat8 OLI, scatter plots and linear fitting were done with the help of 500 to 100 samples between chl-a values and the corresponding Sentinel-3 values from the product.

Statistical analysis indicated (Figures 17 and 18) that the coefficients of determination (R^2) were seen to be 0. 836884 (2015) and 0.843381 for 2016.

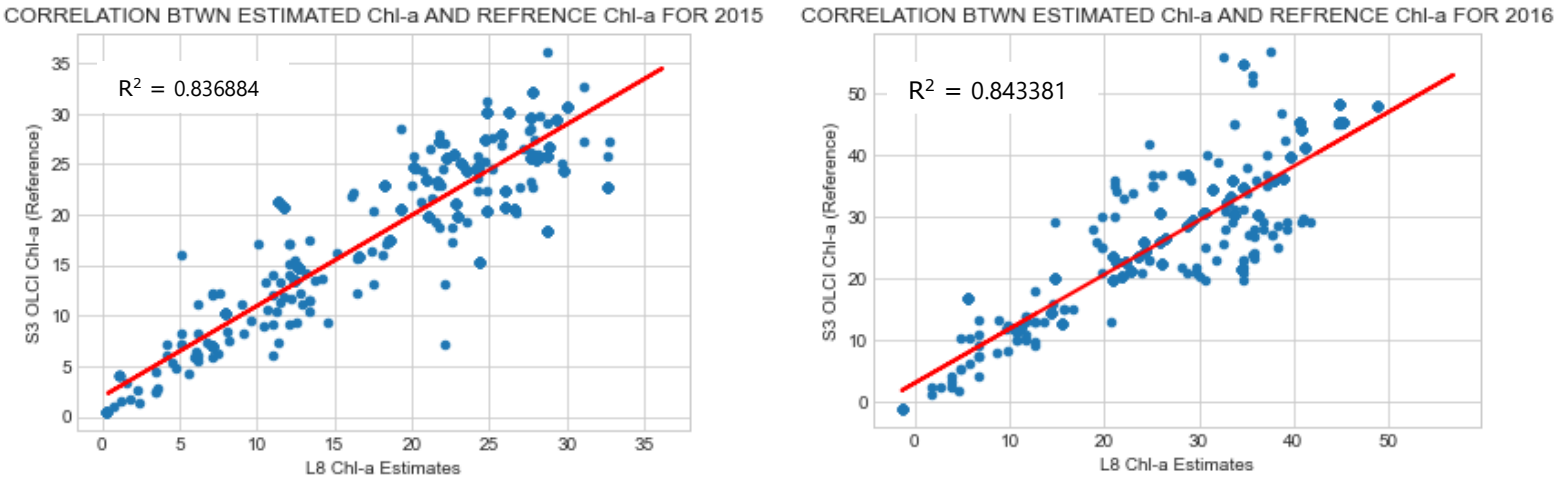

Figure 17 & 18. Scatter plots of correlation of L8 OLI estimates and S3-OLCI products for 2015 and 2016.

For estimated obtained in 2017 and 2018, statistical analysis showed (Figures 19 and 20) that the coefficients of determination (R^2) were seen to be 0. 891653 (2017) and 0. 839463 for 2018, which are of course within acceptable range.

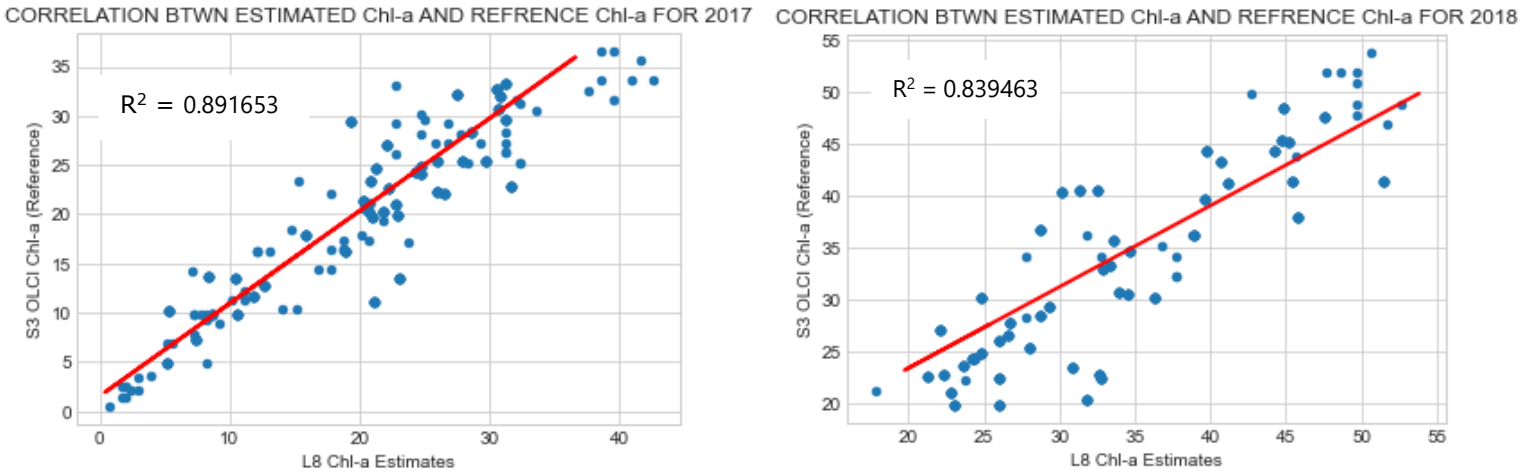

Figure 19 & 20. Scatter plots of correlation of L8 OLI estimates and S3-OLCI product for 2017 and 2018

For estimates obtained in 2019 and 2020, statistical analysis further showed (Figures 20 and 21) that strong positive coefficients of determination (R^2) of 0.899546 (2019) and 0.882167 for 2020.

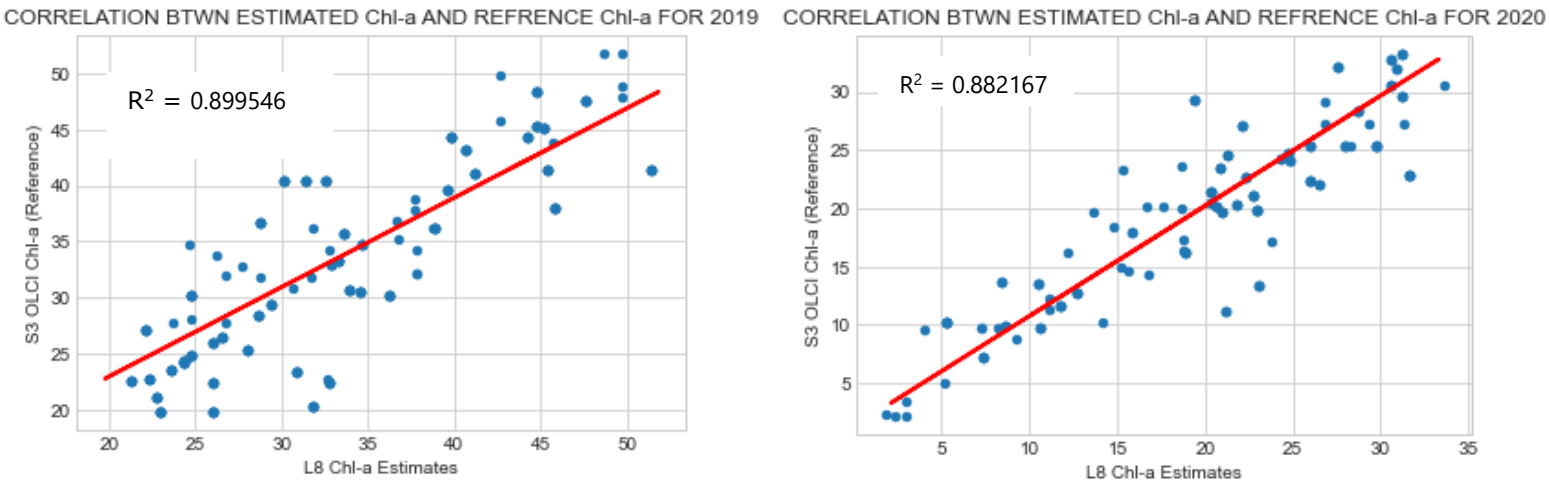

Figure 21 & 22. Scatter plots of correlation of L8 OLI estimates and S3-OLCI product for 2019 and 2020.

This essentially implies that there was a significant strong positive correlation between Chl-a estimated by OC-2 algorithm and ESA's data products provided by Copernicus.

That is to say, in regions with insufficient estimate parameters for chl-a, it is truly reliable and accurate to employ the OC-2 algorithm to estimate using Landsat 8 OLI.

4.2.1 LSAT For corresponding reported HABs Events

As indicated in the methodology, the Moving Window Algorithm was proposed and adopted to estimate the LSAT as another proxy for the HAB as reported. The following displayed results were obtained from the process.

In Figures 23 and 24 below, it can be observed that on the event of HAB as reported, there existed a relatively high thermal emissivity from the Lake surface especially in the affected regions.

This is manifested by the LSAT values rising significantly to up to 35.7 °C in 2015 and 36.6 °C for 2016 especially in the affected regions as shown in the figures below. The unaffected regions however maintained low LSAT values of 23.6 °C and 16.9 °C respectively.

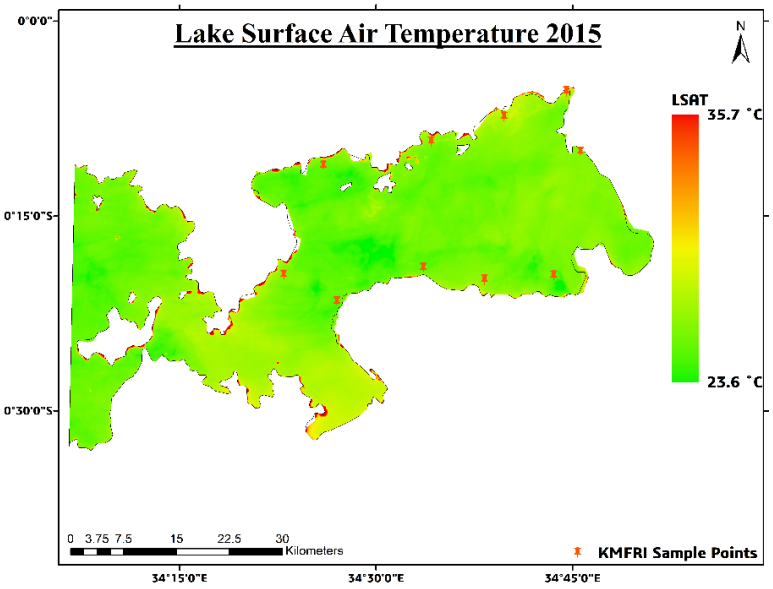

Figure 23. LSAT concentration, 2015

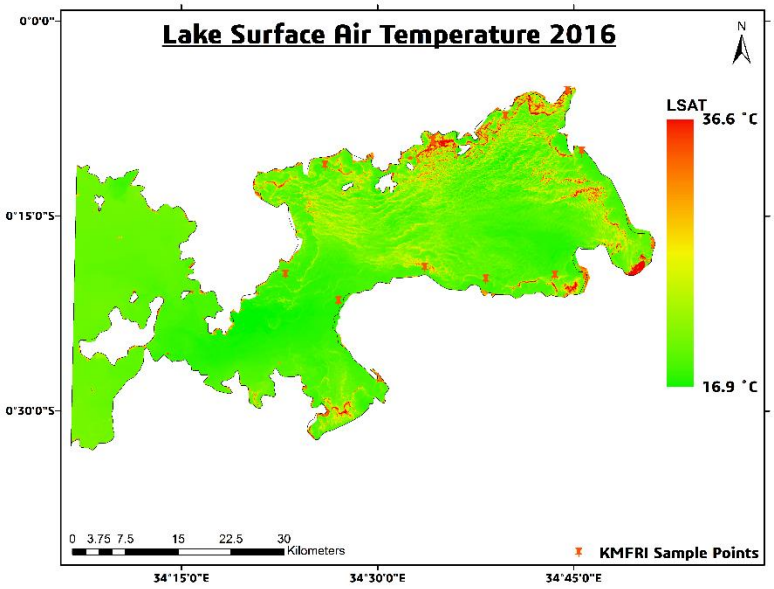

Figure 24. LSAT concentration, 2016

In Figures 25 and 26 below, it is noted that on the event of HAB, LSAT values are seen to rise significantly to up to 35.1 °C and 36.6 °C in the affected regions.

As seen in the chl-a maps in 2017 and 2018, the severity of the bloom was almost evenly distributed in 2017, while in 2018 the distribution mainly dominated the south eastern part as supported by the LSAT maps below.

The western regions of the study area still maintain comparatively low LSAT values indicating they were not adversely affected by the phenomena.

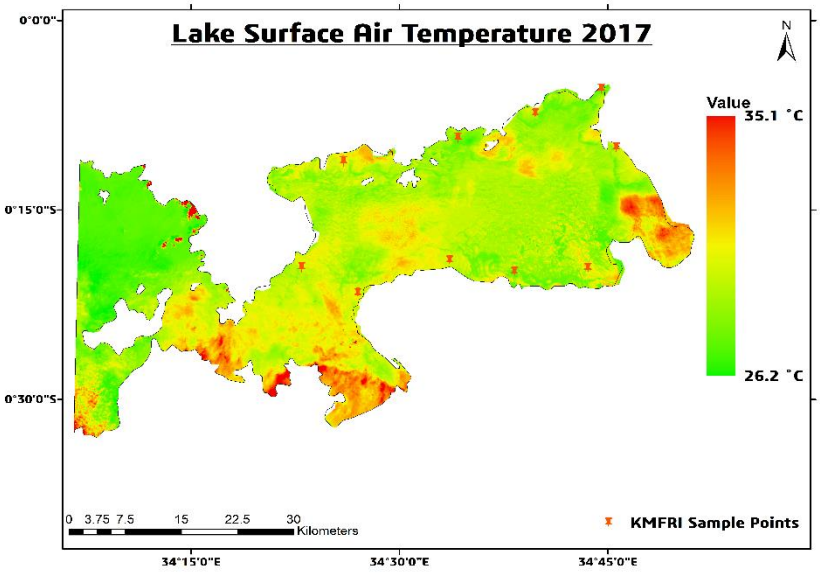

Figure 25. LSAT concentration, 2017

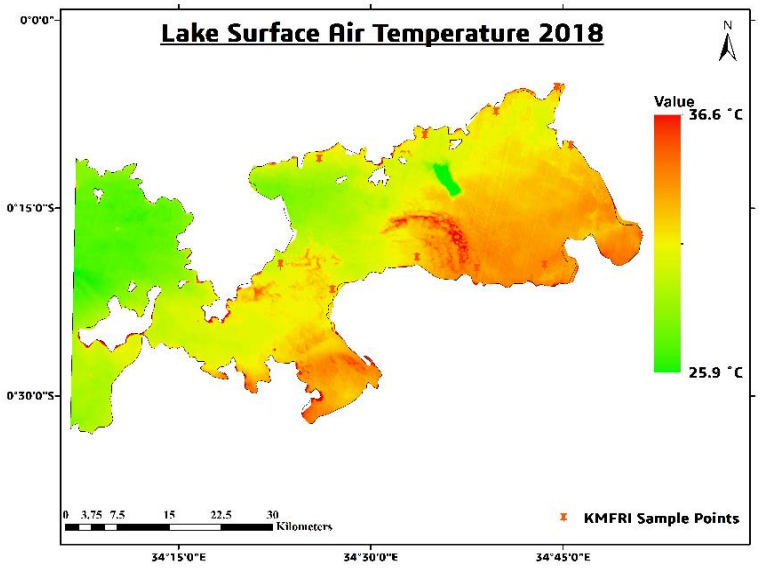

Figure 26. LSAT concentration, 2018

In Figures 27 and 28 below, it is as well noted that on the event of HAB, LSAT values are seen to rise significantly to up to 36.8 °C and 34.8°C in the severely affected regions.

This research therefore agrees with the previously well-documented studies. LSAT values are as well seen to rise significantly in 2019 and 2020. The cases were more severe.

The western regions of the study area still continue to maintain comparatively low LSAT values indicating they were not adversely affected by the phenomena in all the reported dates.

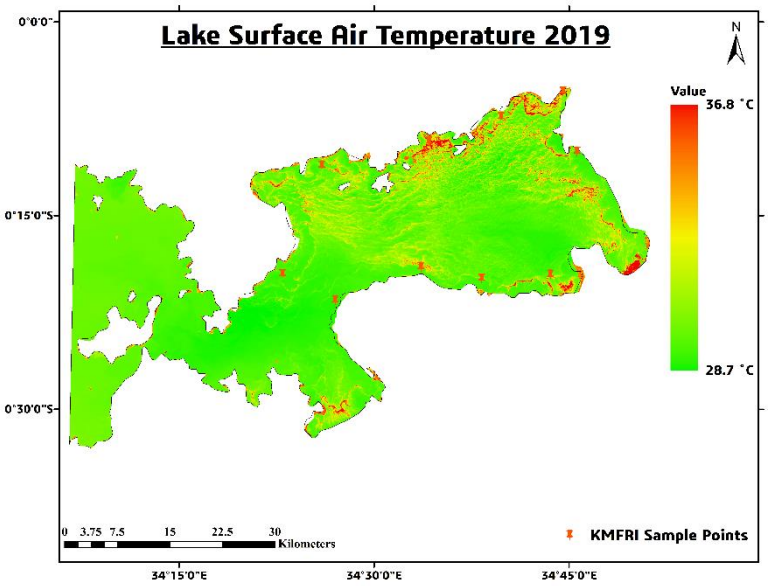

Figure 27. LSAT concentration, 2019

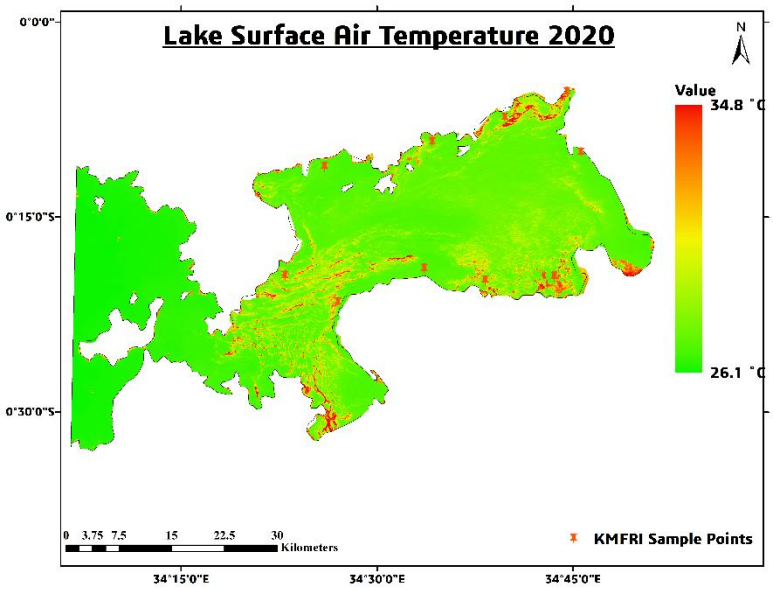

Figure 28. LSAT concentration, 2020

1.1.2 Accuracy assessment of the LSAT Estimates.

In order to validate the estimated results and analyze the relationship between MODIS LSAT data and those retrieved from Landsat8 TIR, scatter plots and linear fittings were done with the help of randomly generated 500 to 1000 samples between LSAT values and the corresponding MODIS values from the product.

Statistical analysis showed (Figures 29 & 30) that the coefficients of determination (R^2) were seen to be 0. 666916 (2015) and 0.843381 for 2016.

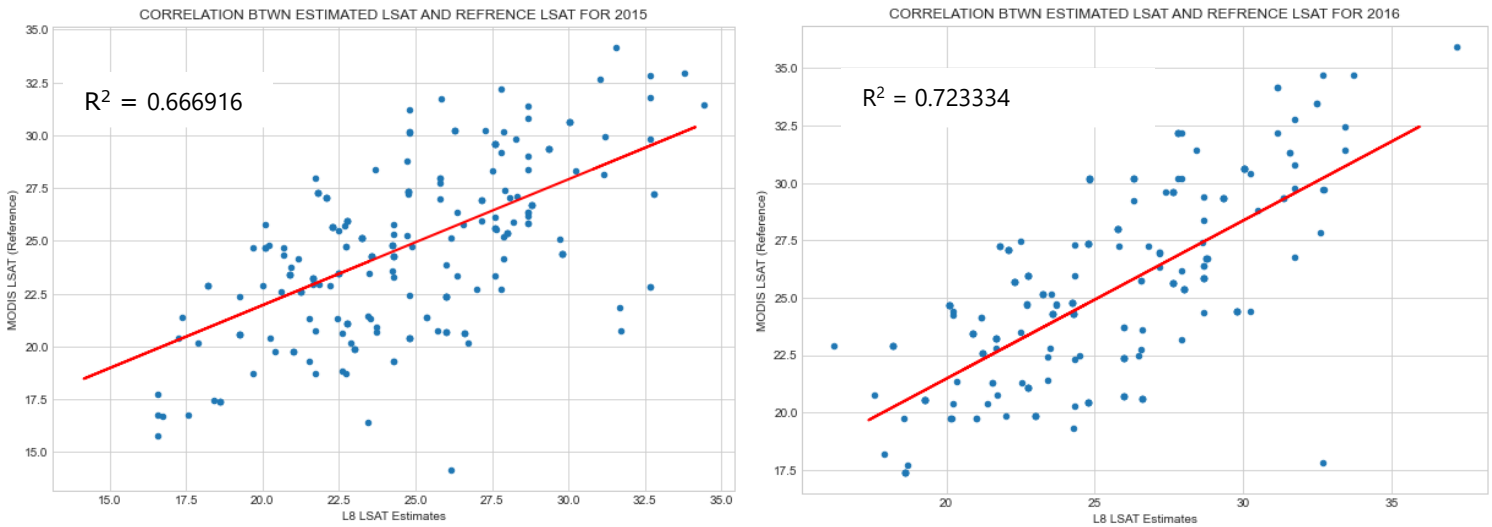

Figure 29 & 30. Scatter plots of correlation of L8 TIRs estimates and MODIS product for 2015 and 2016.

For estimates obtained in 2017 and 2018, statistical analysis showed (Figures 31 and 32) that the coefficients of determination (R^2) were reported to be 0. 696916 (2017) and 0. 781505 for 2018.

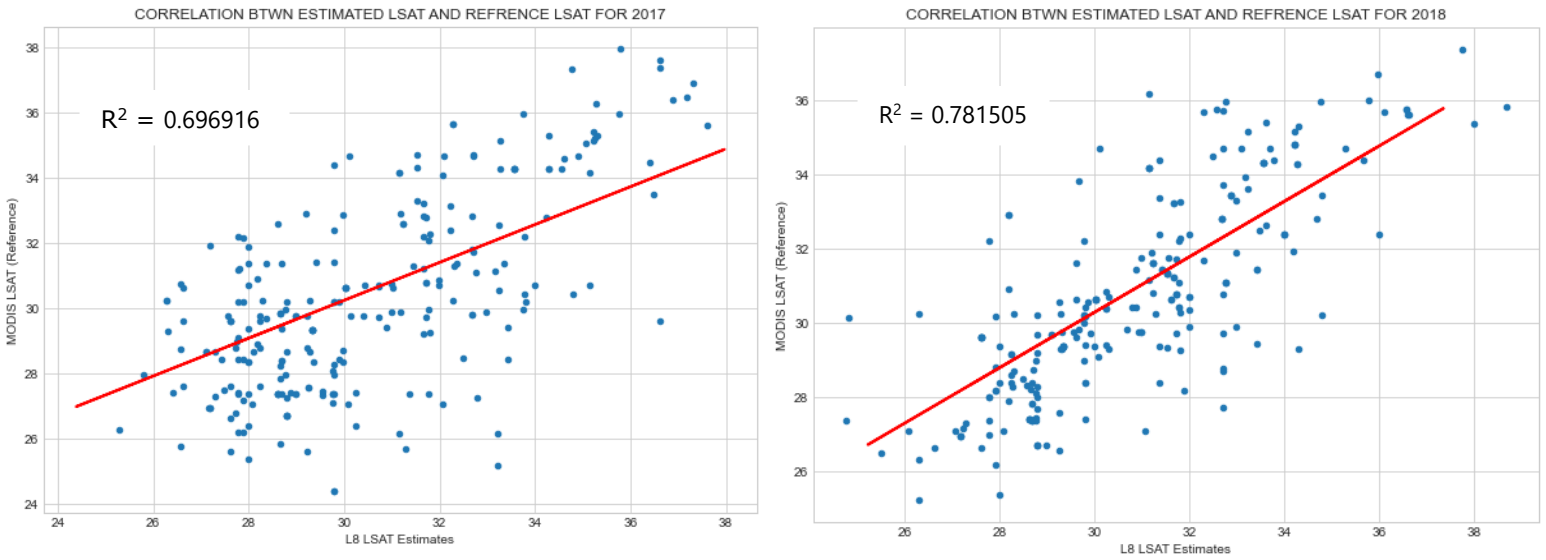

Figure 31 & 32. Scatter plots of correlation of L8 TIRs estimates and MODIS product for 2017 and 2018.

For estimates obtained in 2019 and 2020, statistical analysis further showed (Figures 33 and 34) strong positive coefficients of determination (R^2) of 0.820665 (2019) and 0.754280 for 2020.

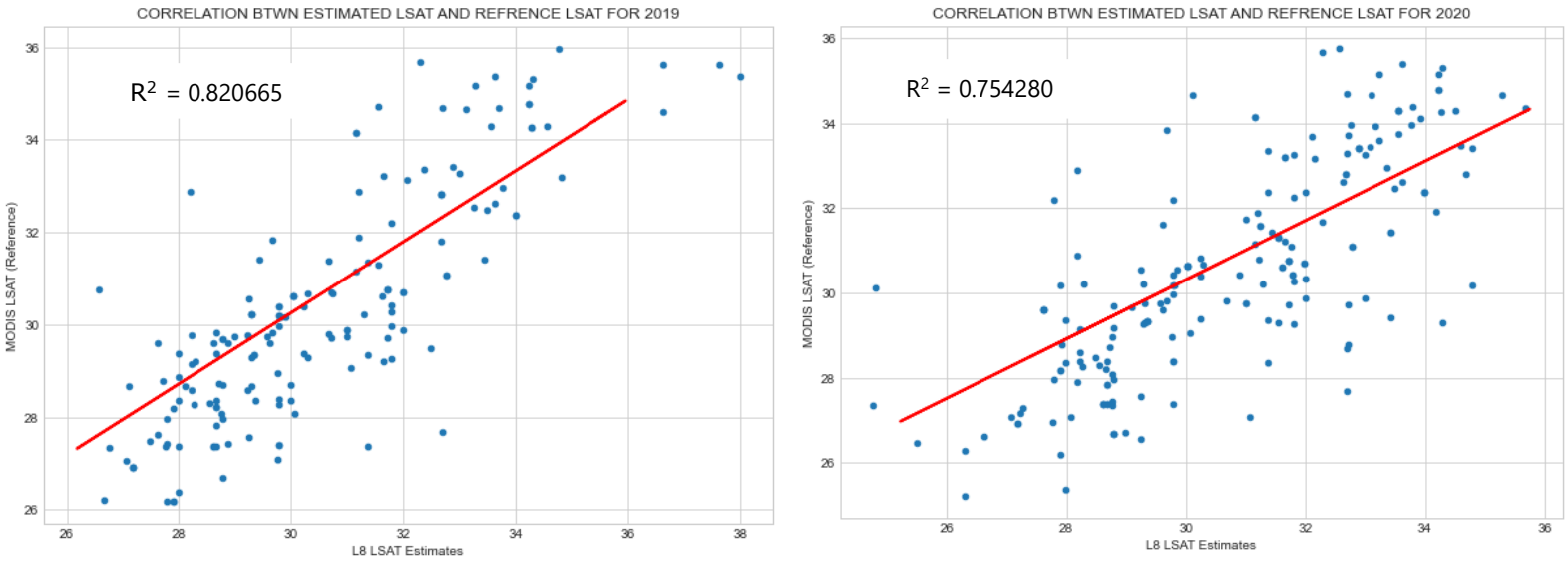

Figure 33 & 34. Scatter plots of correlation of L8 TIRs estimates and MODIS product for 2017 and 2018.

These correlations of determinations essentially implies that there did exist significant strong positive correlation between Lake Surface Air Temperatures estimated by Moving Window algorithm and NASA's MODIS data products provided by the Earth Data.

That is to conclude that, in regions with insufficient estimate parameters for Lake Surface Air Temperature, it is truly reliable and accurate to employ the Moving Window algorithm to estimate using Landsat 8 TIRs.

4.3 Automated *In-Situ* IoT System

In order to come up with a near real-time ground-based monitoring system, an Automated In-situ IoT system was developed, tested and found to be fully operational.

This was primarily to collect geotagged Lake Surface Air Temperature and Lake Surface pH, in this case the salinity being the case.

The system was automatically able to collect the LSAT and send a geotagged notification via text SMS in case of LSAT condition anomalies. The Africastalking text message API was used to relay the near-real time messages from the IoT system to the remote users or authorities.

Figure 35, 36 and 37 indicates how the message upon arrival, appears in the GUI. The text comes along with a Google Maps link indicating the System location from which such data are coming from.

When the user clicks the link, a Google Map UI as indicated in Figures 38, 39 and 40 is realized, showing the Google Map being rendered on the users' screen to show the real-time map locating the system.

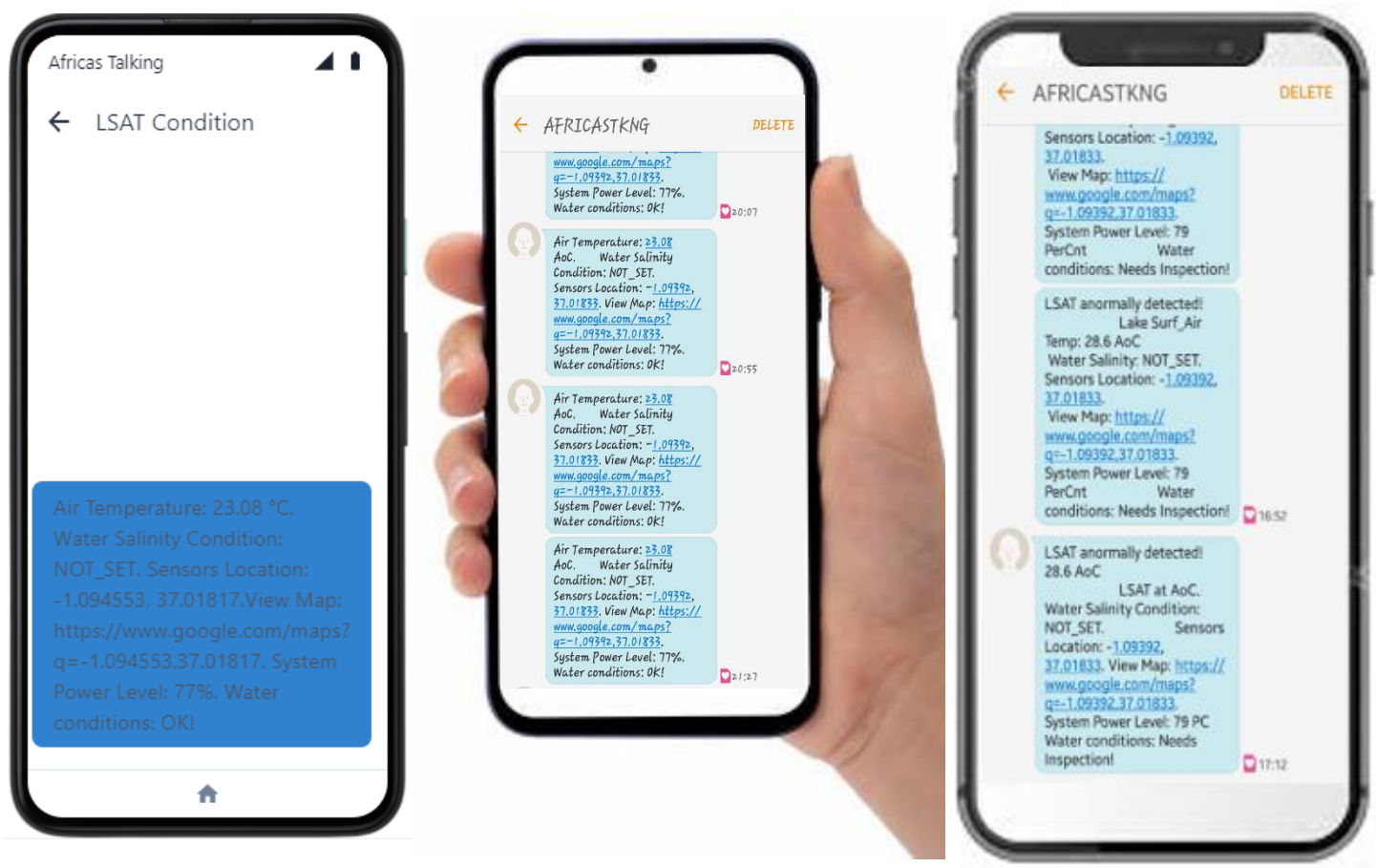

Figure 35, 36 and 37 The text message UI when received by client/user.

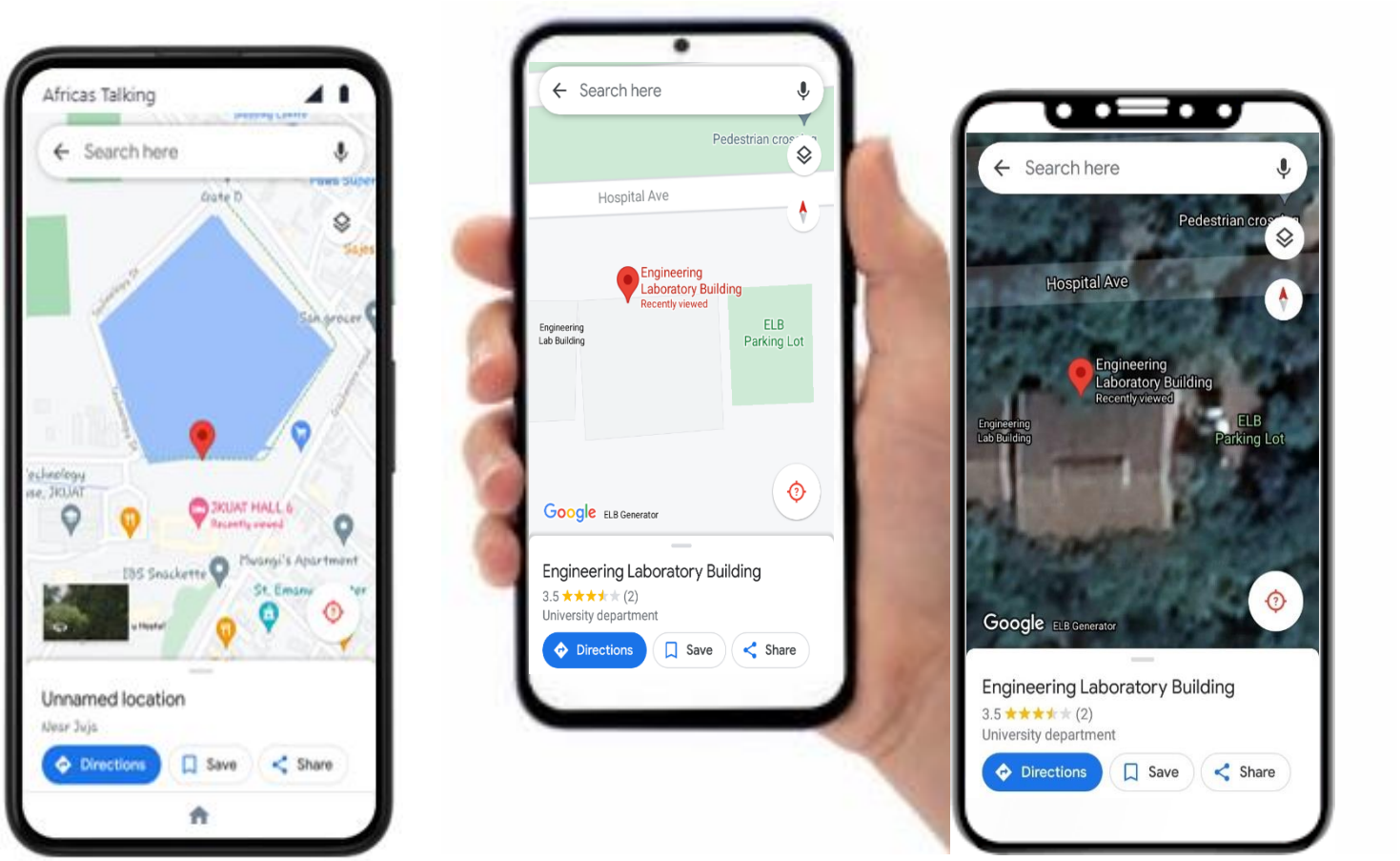

Figure 38, 39 and 40 Client/User viewing sensor located at JKUAT pond and Engineering Laboratory Building (ELB) via Google Maps API.

5.0 Discussion

As seen from the results presented in section four (4) above, chl-a concentrations were reportedly high in the dates assessed and seen to be far above the minimum optimal value of non-turbid waters like Lake Victoria, which has been determined as 20.05m/gm^3. This situation took effect owing to the circumstantial elevated proliferation, predominance and therefore prevalence of the CyanoHABs species that produce the toxins. The used Ocean Color-2 algorithm presented in this literature was able to model and represent this phenomenon.

Out of all the reported incidences, it was noted that 2018 had the lowest magnitude of impact as manifested by the least chl-a concentrations of 25.9 m/gm^3. 2020 bloom occurrence saw the highest concentrations of 57.1 m/gm^3. Generally, the highest concentrations in most cases were located either at the North, South or Central part of the Lake Victoria basin, without necessarily manifesting any predictable trend. In the mentioned sections, the eutrophication caused by a high discharge of pollutions coming from the busy industrial sections of the riparian reserves played the greatest responsibility for the proliferation of harmful algae.

With regard to the Lake Surface Air Temperature (LSAT), the values were seen to corroborate with the chl-a values since high temperatures spots were as well associated with relatively high chl-a concentrations. This agrees with the hypothesis that algae occurrences are closely associated with high temperatures and upon their occurrence, they cause the LSAT to as well rise.

The estimated chl-a and LSAT were correlated with the well accepted Sentinel 3 OLCI Chl-a and MODIS LSAT products. The Chl-a estimates were found to correlate positively strong with R^2 scores ranging from 0.83 to 0.899. Again, the estimated LSAT were correlated with the MODIS LSAT products. The Chl-a estimates were as well found to correlate positively strong with R^2 scores ranging from 0.67 to 0.82. These correlations of determinations essentially implies that there did exist significant strong positive correlation between the estimates and data products highlighted before.

That is to say, in regions with insufficient estimate parameters for Lake Surface Air Temperature and Chl-a, it is truly reliable and accurate to employ the Ocean Color - 2 and Moving Window algorithm to find the required estimates using Landsat OLI and TIRs. These two can then be implemented on board Landsat 8 and enable automatic download of the results in the event of future HAB, which will therefore call for immediate and precautionary corrective measures.

The full-blown IoT system was made up of various components like Lake Surface Air Temperature DHT11 sensor, Neo-Ublox GPS modules with a fixed antenna, Raspberry Pi Microprocessor, breadboard that serves as the circuit prototyping board and power supply that run the software. The system was able to measure and report in real time the Lake Surface Air Temperate, and giving the location in form of WGS84 latitude, longitude, and the time in Greenwich Meridian Time (GMT) and display in a remote hand-held phone as shown in the Figures 35 to 37.

A google map showing the location of the system as at the time of LSAT collection was also implemented as seen in Figures 38, 39 and 40. The LSAT of the test environments in JKUAT ranged from 23.8 to 28.6 which were within optimum were reported. Therefore, this system can help in near-real time to monitor and track the LSAT of Lake Victoria and any other area of study of such related phenomena. The IoT system of hardware and software can therefore improve water quality monitoring and assessment enabling it to be in real time and give the spatial attribute of location which will couple the space-borne monitoring system like Landsat space missions.

6.0 Conclusions and Outlook

In conclusion, the study was able to come up and present a methodological approach to detect and monitor the concentration of chlorophyl-a(chl-a) from Landsat 8 Operational Land Imager data as HAB proxies in Lake Victoria from 2015 to 2020. This was implemented through the Ocean Colour 2 algorithm. Therefore, space-based observations play a significant role in the monitoring of Harmful Algal Blooms.

Further, Landsat 8 Thermal InfraRed (TIR) sensor abord the same space vehicle was able to provide for corresponding thermal images to estimate the rising levels of Lake Surface Air Temperatures in the advent of HAB occurrence.

As displayed in the result section, it is noted that the in occurrence of HAB, Chl-a values rose significantly to 31 to 57.1 Mg/M^3. Similarly, the moving window algorithm for LSAT estimates indicated that on a bloom event, the LSAT rose in the affected areas up to values such as 35.1 °C to 36.6 °C. Conversely, western parts of the study area that were never affected by the phenomena marked the lowest chl-a value (-1.2 mg/m^3) and corresponding temperatures value (less than 27 °C).

With the unfortunate lack of synchronized measured ground truthing data from the authorities like KMFRI, the estimated Chl-a and LSAT products have been validated using in Chl-a products from Sentinel 3 -OLCI and LSAT with the MODIS (MOD11A1) data; Validation results from the scatter plots confirmed that the acceptably high correlation coefficient values of 0.743 to 0.876 for LSAT and 0.845 to 0.902 for Chl-a, this confidently demonstrates the applicability of the proposed and adopted algorithm in Chl-a estimation and LSAT retrieval in the study area.

Finally, to ensure near-real time monitoring at field level, an automated in-situ system was developed to collect and monitor the Lake Surface Air Temperature (LSAT) as another proxy for HAB event. The system was tested and found operational was able to report remotely in case of any abnormal rise in the Lake Surface Air Temperature, calling for further response. This would help the maritime authorities to take effective response to the alert sent and take corrective measures. The system was tested in the local University Pond as indicated in Figures 35 to 40 and was able to report Surface Air Temperature values of ranges 23°C to 28.6°C.

Recommendations

After the implementation of a variety of proposals, the following are areas of improvement that need to be scaled for future betterment of this project:

a) To increase the temporal resolutions of the monitoring and reporting, there need to do a Spatio-temporal image fusion of Landsat 8 and MODIS aqua to provide a near daily 30 M spatial resolution of Chla-a and LSAT estimates instead of waiting for the 16-day 30M Landsat as this current study implements.
b) The Maritime and Environmental management authorities within the Lake Victoria riparian region should take more action in ensuring that the sewage from the Industrial regions is well treated before discharge to minimize the eutrophication of the Lake Victoria Basin.
c) The presented System had relatively low mechanical integrity as there was no protective cover to shield it against environmental hazards like scorching sun, unfavorable rainy weather conditions among others. A protective cover to shelter all the components should therefore be fabricated.
d) Furthermore, including another water quality sensor such as water pH (Salinity) which is another indicator of HAB should be incorporated in the study. The resent study proposed but was not within the might to implement the said.

7. References

Allan, M. G., Hamilton, D. P., Hicks, B., and Brabyn, L. (2015). *Empirical and semi-analytical chlorophyll-a algorithm for multi-temporal monitoring of New Zealand lakes using Landsat*. Environ. Monit. Assess. 187, 364. https://doi.org/10.1007/s10661-015-4585-4

Anderson, D.M., Glibert, P.M., Burkholder, J.M., (2002). H*armful algal blooms and eutrophication: nutrient sources, composition, and consequences*. Estuaries 25, 704–726. Babin, M., Roesler, C.S., Cullen, J.J., 2008. Real-time Coastal Observing Systems for Marine Ecosystem Dynamics and Harmful Algal Blooms: Theory, Instrumentation and Modelling. UNESCO.

Anderson, J.; Wilson, S. (1984) *The physical basis of current infrared remote-sensing techniques and the interpretation of data from aerial surveys*. Int. J. Remote Sens., 5, 1–18.

Augusto-Silva, P.B.; Ogashawara, I.; Barbosa, C.C.F.; de Carvalho, L.A.S.; Jorge, D.S.F.; Fornari, C.I.; Stech, J.L. (2014) *Analysis of MERIS reflectance algorithms for estimating chlorophyll-a concentration in a Brazilian Reservoir. Remote Sens., 6, 11689–117077.*

Babiker, Insaf, Mohamed, Mohamed, Hiyama, Tetsuya, Kato, Kikuo, (2010) *A GIS-based model for assessing aquifer vulnerability in Kakamigahara Heights, Gifu Prefecture, central Japan*.

Beck, R.A.; Zhan, S.; Liu, H.; Tong, S.T.Y.; Yang, B.; Xu, M.; Ye, Z.; Huang, Y.; Shu, S.; Wu, Q.; et al. (2016) *Comparison of satellite reflectance algorithms for estimating chlorophyll-a in a temperate reservoir using coincident hyperspectral aircraft imagery and dense coincident surface observations.* Remote Sens. Environ. 178, 15–30.

Blondeau-Patissier, D., Gower, J. F. R., Dekker, A. G., Phinn, S. R., & Brando, V. E. (2014). *A review of ocean color remote sensing methods and statistical techniques for the detection, mapping and analysis of phytoplankton blooms in coastal and open oceans*. *Progress in Oceanography*, *123*, 123–144. https://doi.org/10.1016/j.pocean.2013.12.008

Boddula, V., Ramaswamy, L., & Mishra, D. (2017). CyanoSense: *A Wireless Remote Sensor System Using Raspberry-Pi and Arduino with Application to Algal Bloom. 2017 IEEE International Conference on AI & Mobile Services (AIMS)*, 85–88. https://doi.org/10.1109/AIMS.2017.19

Budyko, M. I. (1974) Climate and Life. International Geophysics Series, vol.18, Academic Press, New York. Bugenyi, F. W. B. & Magumba, K. M. (1996) *The present physicochemical ecology of Lake Victoria, Uganda*. In: The Limnology, Climatology and Paleocl

Bukata, R. P., Jerome, J. H., Kondratyev, K. Y., and Pozdnyakox, D. V. (1995). *Optical properties and remote sensing of inland and coastal waters*. New York, NY: CRC Press

Caballero, I., Fernández, R., Escalante, O. M., Mamán, L., & Navarro, G. (2020). *New capabilities of Sentinel-2A/B satellites combined with in situ data for monitoring small harmful algal blooms in complex coastal waters*. *Scientific Reports*, *10*(1), 8743. https://doi.org/10.1038/s41598-020-65600-1

Calamari, D.; Akech, M.O.; Ochumba, P.B.O. (1995) *Pollution of Winam Gulf, Lake Victoria, Kenya: A case study for preliminary risk assessment*. Lakes Reserv. Res. Manag. 1, 89–106.

Cao, H., & Han, L. (2021). *Hourly remote sensing monitoring of harmful algal blooms (HABs) in Taihu Lake based on GOCI images*. *Environmental Science and Pollution Research*, *28*(27), 35958–35970. https://doi.org/10.1007/s11356-021-13318-6

Cao, Z., Ma, R., Duan, H., Pahlevan, N., Melack, J., Shen, M., et al. (2020). *A machine learning approach to estimate chlorophyll-a from Landsat-8 measurements in inland lakes. Rem*. Sens. Environ. 248, 111974. https://doi.org/10. 1016/j.rse.2020.111974

Chorus and J. Bartram (1999), Toxic cyanobacteria in water. A guide to their public health consequences, monitoring and management. London: E & FN Spon,388p.

Clarke, G. L., Ewing, G. C., and Lorenzen, C. J. (1970). *Spectra of backscattered light from the sea obtained from aircraft as a measure of chlorophyll concentration*. Science 167, 1119–1121. https://doi.org/10.1126/science.167.3921.1119

Cloete, N.A.; Malekian, R.; Nair, L. *Design of Smart Sensors for Real-Time Water Quality Monitoring.* IEEE Access 2016, 4, 3975–3990.

Concha, J. A., and Schott, J. R. (2016). *Retrieval of color producing agents in case 2 waters using Landsat 8.* Rem. Sens. Environ. 185, 95–107. https://doi.org/10.1016/j.rse.2016. 03.018

Diaz R, Rosenberg R (2008) *Spreading dead zones and consequences for marine ecosystems*. Science 321:926–929

Electronics Lovers, Using the DHT11 Sensor with Raspberry Pi to Measure Temperature and Humidity, Retrieved on 14th July 2021 from https://www.electronicslovers.com/2019/02/using-the-dht11-sensor-with-raspberry-pi.html

Encinas, C.; Ruiz, E.; Cortez, J.; Espinoza, A. (2017) *Design and implementation of a distributed IoT system for the monitoring of water quality in aquaculture*. In Proceedings of the Wireless Telecommunications Symposium, Chicago, IL, USA.

Freitas, F. H., and Dierssen, H. M. (2019). *Evaluating the seasonal and decadal performance of red band difference algorithms for chlorophyll in an optically complex estuary with winter and summer blooms*. Rem. Sens. Environ. 231, 111228. https://doi.org/10.1016/j.rse.2019.111228

Gikuma-Njuru, P.; Hecky, R.E.; Guildford, S.J.; MacIntyre, S. (2013) *Spatial variability of nutrient concentrations, fluxes, and ecosystem metabolism in Nyanza Gulf and Rusinga Channel, Lake Victoria (East Africa)*. Limnol. Oceanogr.58, 774–789.

Gitelson, A. (1992). The peak near 700 nm on radiance spectra of algae and water: relationships of its magnitude and position with chlorophyll concentration. Int. J. Rem. Sens. 13, 3367–3373. https://doi.org/10.1080/ 01431169208904125

Glibert, P.M., Anderson, D.M., Gentien, P., Granéli, E., & Sellner, K.G. (2005). *The Global, Complex Phenomena of Harmful Algal Blooms*

Glibert P, Heil C, Hollander D, Revilla M, Hoare A, Alexander J, Murasko S (2004) *Evidence for dissolved organic nitrogen and phosphorus uptake during a cyanobacterial bloom in Florida Bay*. Mar Ecol Prog Ser 280:73–83

Gordon, H. R., Clark, D. K., Brown, J. W., Brown, O. B., Evans, R. H., and Broenkow, W. W. (1983). *Phytoplankton pigment concentrations in the Middle Atlantic Bight: comparison of ship determinations and CZCS estimates*. Appl. Optic. 22, 20–36. https://doi.org/10.1364/ao.22.000020

Gordon, H. R., Clark, D. K., Mueller, J. L., and Hovis, W. A. (1980). *Phytoplankton pigments from the nimbus-7 coastal zone color scanner: comparisons with surface measurements*. Science 210, 63–66. https://doi.org/10.1126/science.210.4465.63

Gohin F. et al., (2003) *Satellite and in situ observations of a late winter phytoplankton bloom, in the Northern Bay of Biscay*. Continental Shelf Research, vol. 23, pp. 1117–1141,

Gower, J., Lin, S., and Borstad, G. (1984). *The information content of different optical spectral ranges for remote chlorophyll estimation in coastal waters*. Int. J. Rem. Sens. 5, 349–364. https://doi.org/10.1080/01431168408948813

Guo L (2007) *Doing battle with the green monster of Taihu Lake*. Science 317:1166

Haakstad, M.; Kogeler, J.; Dahle, S. (1994) *Studies of sea surface temperatures in selected northern norwegian fjords using Landsat TM data*. Polar Res., 13.

Hallegraeff, G. M (1993). *A review of harmful algal blooms and their apparent global increase.* Phycologia 32, 79-99

Hecky, R. E., Mugidde, R., Ramlal, P. S., Talbot, M. R., & Kling, G. W. (2010). *Multiple stressors cause rapid ecosystem change in Lake Victoria. Freshwater Biology*, 55, 19–42. https://doi.org/10.1111/j.1365-2427.2009.02374.x

Huisman J, Matthijs HC, Visser PM (2005) *Harmful cyanobacteria. Aquatic Ecology Series. Springer, Dordrecht*

Irons, J. R., Dwyer, J. L., and Barsi, J. A. (2012). The next Landsat satellite: the Landsat data continuity mission. Rem. Sens. Environ. 122, 11–21. https://doi.org/10.1016/j. rse.2011.08.026

Hill, P. R., Kumar, A., Temimi, M., & Bull, D. R. (2020). HABNet: *Machine Learning, Remote Sensing Based Detection and Prediction of Harmful Algal Blooms*. ArXiv:1912.02305 [Cs, Eess]. http://arxiv.org/abs/1912.02305

Jan, F., Min-Allah, N., & Düştegör, D. (2021). *IoT Based Smart Water Quality Monitoring: Recent Techniques, Trends and Challenges for Domestic Applications*. *Water*, *13*(13), 1729. https://doi.org/10.3390/w13131729

Jiang GJ, Ma RH, Loiselle S, Su W, Cai WX, Huang CG, Yang J, Yu W (2015) Remote sensing of particulate organic carbon dynamics in a eutrophic lake (Taihu Lake, China). Sci Total Environ 532:245–254

Khalili, M. H., & Hasanlou, M. (2019). HARMFUL ALGAL BLOOMS MONITORING USING SENTINEL-2 SATELLITE IMAGES. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-4/W18, 609–613. https://doi.org/10.5194/isprs-archives-XLII-4-W18-609-2019

Khorram, S., Catts, G. P., Cloern, J. E., and Knight, A. W. (1987). *Modeling of estuarne chlorophyll a from an airborne scanner*. IEEE Trans. Geosci. Rem. Sens. 25, 662–669. https://doi.org/10.1109/tgrs.1987.289735

Kurekin, A.A., Miller, P.I., Van der Woerd, H.J., (2014). *Satellite discrimination of Karenia mikimotoi and Phaeocystis harmful algal blooms in European coastal waters: merged classification of ocean colour data*. Harmful Algae 31, 163–176

Le, C., Hu, C., English, D., Cannizzaro, J., Chen, Z., Feng, L., et al. (2013). *Towards a long-term chlorophyll-a data record in a turbid estuary using MODIS observations*. Prog. Oceanogr. 109, 90–103. doi:10.1016/j.pocean.2012.10.002

Liu JG, Yang W (2012) *Water sustainability for China and beyond*. Science 337:649–650

Luo JH, Li XC, Ma RH, Li F, Duan HT, Hu WP, Qin BQ, Huang WJ (2016) *Applying remote sensing techniques to monitoring seasonal and interannual changes of aquatic vegetation in Taihu Lake*. China. Ecol Indic 60:503–513

Manuel, A., Blanco, A., Tamondong, A., Jalbuena, R., Cabrera, O., and Gege, P. (2020). *Optmization of bio-optical model parameters for turbid lake water quality estimation using Landsat 8 and wasi-2D*. Int. Arch. Photogram. Rem. Sens. Spatial Inf. Sci. 11, 67–72. https://doi.org/10.5194/isprs-archives-xlii-3-w11-67-2020

Markham, B., Barsi, J., Kvaran, G., Ong, L., Kaita, E., Biggar, S., et al. (2014). *Landsat-8 operational land imager radiometric calibration and stability*. Rem. Sens. 6, 12275–12308. https://doi.org/10.3390/rs61212275

Markham, B. L., Barsi, J. A., Morfitt, R., Choate, M., Montanaro, M., Arvidson, T., et al. (2015). "*Landsat 8: status and on-orbit performance*," in SPIE remote sensing. Bellingham, WA: International Society for Optics and Photonics, 963908

Matthews, M.W. et al., (2012). *An algorithm for detecting trophic status (chlorophylla), cyanobacterial-dominance, surface scums and floating vegetation in Inland and coastal waters*. Remote Sensing of Environment 124, 637–652

Mittenzwey, K. H., Ullrich, S., Gitelson, A., and Kondratiev, K. (1992). *Determination of chlorophyll a of inland waters on the basis of spectral reflectance*. Limnol. Oceanogr. 37, 147–149. *https://doi.org/10.4319/lo.1992.37.1.0147*

NASA (2011). *Landsat 7 Science Data Users Handbook Landsat Project Science Office at NASA's Goddard Space Flight Center in Greenbelt.* 186. Available online at: http://landsat.gsfc.nasa.gov/wp-content/uploads/2016/08/Landsat7_Handbook.pdf (accessed July 5, 2021).

Munday, J., and Zubkoff, P. L. (1981). *Remote sensing of dinoflagellate blooms in a turbid estuary*. Photogramm. Eng. Rem. Sens. 47, 523–531.

Neil, C., Spyrakos, E., Hunter, P. D., and Tyler, A. N. (2019). *A global approach for chlorophyll-a retrieval across optically complex inland waters based on optical water types*. Rem. Sens. Environ. 229, 159–178. https://doi.org/10.1016/j.rse. 2019.04.027

Ochumba, P.B.O. (1987) *Periodic massive fish kills in the Kenyan part of Lake Victoria. Water Qual. Bull.12* ,119-122, 130.

Okello, W., & Kurmayer, R. (2011). *Seasonal development of cyanobacteria and microcystin production in Ugandan freshwater lakes: Seasonal development of cyanobacteria and microcystin production. Lakes & Reservoirs: Research &*

Management, 16(2), 123–135. https://doi.org/10.1111/j.1440-1770.2011.00450.x

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., et al. (1998). *Ocean color chlorophyll algorithms for SeaWiFS*. J. Geophys. Res. 103, 24937–24953. https://doi.org/10.1029/98jc02160

Owen M. Dohertyb, Christopher J. Goblera, Theresa K. Hattenrath-Lehmanna, Andrew W. Griffitha, Yoonja Kanga, and R. Wayne Litakerc., (2017). *Ocean warming since 1982 has expanded the niche of toxic algal blooms in the North Atlantic and North Pacific oceans.* https//doi.org/10.1073/pnas.1619575114/

Pahlevan, N., Lee, Z., Wei, J., Schaff, C., Schott, J., and Berk, A. (2014). *On-orbit radiometric characterization of OLI (Landsat-8) for applications in aquatic remote sensing*. Rem. Sens. Environ. 154, 272–284. https://doi.org/10.1016/j.rse.2014.08.001

Qin BQ, Zhu GW, Gao G, Zhang YL, Li W, Paerl HW, Carmichael WW (2010) *A drinking water crisis in Lake Taihu, China*: linkage to climatic variability and lake management. Environ Manag 45:105– 112

Qin BQ, Yang GJ, Ma JR, Deng JM, Li W, Wu TF, Liu LZ, Gao G, Zhu GW, Zhang YL (2016) *Dynamics of variability and mechanism of harmful cyanobacteria bloom in Lake Taihu, China*. Chin Sci Bull 61:759–770

Raju, K.R.S.R.; Varma, G.H.K.(2017). Knowledge based real time monitoring system for aquaculture Using IoT. In Proceedings of the 7th IEEE International Advanced Computing Conference, IACC 2017, Hyderabad, India; pp. 318–321.

Richard, J.; Richard, B.; Jakub, N.; Christopher, N.; Min, X.; Song, S.; Bo, Y.; Hongxing, L.; Erich, E.; Molly, R.; et al (2018) Evaluating the portability of satellite derived chlorophyll-a algorithms for temperate inland lakes using airborne hyperspectral imagery and dense surface observations. Harmful Algae, 76, 35–46.

Richardson, K., 1997. *Harmful or exceptional phytoplankton blooms in the marine ecosystem.* Advances in Marine Biology 31, 301–385

Rhinane H, Hilali A, Bahi H, Berrada A. 2012. Contribution of Landsat data for the detection of urban heat islands areas Case of Casablanca. J Geog Inf Syst. 04:20–26.

River, S.; Sub-Basins, S.R. (2004) *Aerial Surveys Using Thermal Infrared and Color Videography; University of California*: Davis, CA, USA.

Roegner, A., Sitoki, L., Weirich, C. et al. Harmful Algal Blooms Threaten the Health of Peri-Urban Fisher Communities: A Case Study in Kisumu Bay, Lake Victoria, Kenya. Expo Health 12, 835–848 (2020)

Rundquist, D. C., Han, L., Schalles, J. F., and Peake, J. S. (1996). *Remote measurement of algal chlorophyll in surface waters: the case for the first derivative of reflectance near 690 nm*. Photogramm. Eng. Rem. Sens. 62, 195–200.

Ryu, J.-H., Han, H.-J., Cho, S., Park, Y.-J., and Ahn, Y.-H. (2012). Overview of geostationary ocean color imager (GOCI) and GOCI data processing system (GDPS). Ocean Sci. J. 47, 223–233. https://doi.org/doi:10.1007/s12601-012-0024-4

Santoleri R. et al., "Year-to-year variability of the phytoplankton bloom in the southern adriatic sea (1998-2000): sea-viewing wide field-of-view sensor observations and modeling study," Journal of Geophysical Research, vol. 108, p. 8122, 2003

Simiyu, B., Oduor, S., Rohrlack, T., Sitoki, L., & Kurmayer, R. (2018). Microcystin Content in Phytoplankton and in Small Fish from Eutrophic Nyanza Gulf, Lake Victoria, Kenya. Toxins, 10(7), 275. https://doi.org/10.3390/toxins10070275

Sitoki, L.; Kurmayer, R.; Rott, E(2012). Spatial variation of phytoplankton composition, biovolume, and resulting microcystin concentrations in the Nyanza Gulf (Lake Victoria, Kenya). Hydrobiologia, 691, 109–122.

Smith, B., Pahlevan, N., Schalles, J., Ruberg, S., Errera, R., Ma, R., Giardino, C., Bresciani, M., Barbosa, C., Moore, T., Fernandez, V., Alikas, K., & Kangro, K. (2021). A Chlorophyll-a Algorithm for Landsat-8 Based on Mixture Density Networks. *Frontiers in Remote Sensing*, *1*, 623678. https://doi.org/10.3389/frsen.2020.623678

Smith, R. C., and Baker, K. S. (1982). *Oceanic chlorophyll concentrations as determined by satellite* (Nimbus-7 coastal zone color scanner). Mar. Biol. 66, 269–279. https://doi.org/10.1007/bf00397032

Smith, R. B., Bass, B., Sawyer, D., Depew, D., & Watson, S. B. (2019). Estimating the economic costs of algal blooms in the Canadian Lake Erie Basin. Harmful Algae, 87, 101624. https://doi.org/10.1016/j.hal.2019.101624

Snyder, J., Boss, E., Weatherbee, R., Thomas, A. C., Brady, D., and Newell, C. (2017). *Oyster aquaculture site selection using Landsat 8-derived sea surface temperature, turbidity, and chlorophyll a. Front*. Marine Sci. 4, 190. doi:10. 3389/fmars.2017.00190

Song, W., Dolan, J., Cline, D., & Xiong, G. (2015). Learning-Based Algal Bloom Event Recognition for Oceanographic Decision Support System Using Remote Sensing Data. Remote Sensing, 7(10), 13564–13585. https://doi.org/10.3390/rs71013564

Tamatamah, R. A., Hecky, R. E., & Duthie, HamishC. (2005). *The atmospheric deposition of phosphorus in Lake Victoria (East Africa)*. Biogeochemistry, 73(2), 325–344. https://doi.org/10.1007/s10533-004-0196-9

Tang, D., Kawamura, H., Oh, I. S., & Baker, J. (2006). *Satellite evidence of harmful algal blooms and related oceanographic features in the Bohai Sea during autumn 1998*. Advances in Space Research, 9.

Thomas MK, Kremer CT, Klausmeier CA, Litchman E (2012) A global pattern of thermal adaptation in marine phytoplankton. Science 338:1085–108

Tuuli, S.; Kristi, U.; Dainis, J.; Agris, B.; Matiss, Z.; Tiit, K. (2020) *Validation and Comparison of Water Quality Products in Baltic Lakes Using Sentinel-2 MSI and Sentinel-3 OLCI Data*. Sensors, 20, 742.

USGS for a changing world, what is the best band to use in my research https://www.usgs.gov/faqs/what-are-best-landsat-spectral-bands-use-my-research?qt-news_science_products=0#qt-news_science_products

Vos, W., Donze, M., and Buiteveld, H. (1986). *On the reflectance spectrum of algae in water: the nature of the peak at 700 nm and its shift with varying algal concentration*. Delft, Netherlands: Delft University of Technology, Faculty of Civil Engineering

W. Song, J.M Dola, D. Cline and G. Xiong. (2015). Learning-based algal bloom event recognition for oceanographic decision support system using remote sensing data, Remote Sensing, vol. 7, no. 10, pp. 13 564–13 585.

Wang F, Qin Z, Song C, Tu L, Karnieli A, Zhao S. 2015. An improved mono-window algorithm for land surface temperature retrieval from Landsat 8 thermal infrared sensor data. Remote Sens. 7:4268–4289

Wang, M., Liu, X., Jiang, L., Son, S., Sun, J., Shi, W., et al. (2014). "Evaluation of VIIRS ocean color products," in Ocean remote sensing and monitoring from SpaceInternational society for optics and photonics, 92610E

Wan Mohtar, W. H. M., Abdul Maulud, K. N., Muhammad, N. S., Sharil, S., & Yaseen, Z. M. (2019). Spatial and temporal risk quotient-based river assessment for water resources management. Environmental Pollution, 248, 133–144. https://doi.org/10.1016/J.ENVPOL.2019.02.011

Waspmote- Wireless Sensor Networks Open Source Platform. https://www.cooking-hacks.com/documentation/tutorials/waspmote #waspmote ps 2 8 4

Watanabe, F. S., Alcantara, E., Rodrigues, T. W., Imai, N. N., Barbosa, C. C., and Rotta, L. H. (2015). *Estimation of chlorophyll-a concentration and the trophic state of the Barra Bonita hydroelectric reservoir using OLI/ Landsat-8 images.* Int. J. Environ. Res. Publ. Health 12, 10391–10417. https://doi.org/10.3390/ijerph120910391

Watanabe, F., Alcantara, E., Rodrigues, T., Rotta, L., Bernardo, N., and Imai, N. (2017). Remote sensing of the chlorophyll-a based on OLI/Landsat-8 and MSI/Sentinel-2A (Barra Bonita reservoir, Brazil). An. Acad. Bras. Ciênc. 90, 1987–2000. https://doi.org/10.1590/0001-3765201720170125

Wezernak, C., Tanis, F., and Bajza, C. (1976). *Trophic state analysis of inland lakes.* Rem. Sens. Environ. 5, 147–164. https://doi.org/10.1016/0034-4257(76)90045-6

Yang, C.; Wang, X. (2011) *The water quality and pollution character in Qingshuihai lake valley-typical urban drinking water sources*. In Proceedings of the 2011 International Conference on Remote Sensing, Environment and Transportation Engineering, Nanjing, China; pp. 7287–7291

Yin, X., & Nicholson, S. E. (n.d.). *The water balance of Lake Victoria*. 23.

Zhao, Y., Liu, D., & Wei, X. (2020). *Monitoring cyanobacterial harmful algal blooms at high spatiotemporal resolution by fusing Landsat and MODIS imagery*. *Environmental Advances*, *2*, 100008. https://doi.org/10.1016/j.envadv.2020.100008

Appendix

Below are the extra materials that were relevant to my study but could not fit in the main text of my write up.

1. Ocean Color-2 Chl-a Algorithm

```
# Raster, chlorophyll-index and global environment
# Install these packages before usage.
rm(list = ls(all=TRUE)) #clear memory
library(raster)
library(rgdal)

# import default proj dir and Check path contents
path <- setwd("your/working/directory")

# Denote as Multiraster data
aoiRaster <- stack("rasterFile.tif")
aoiRaster

# OC2 Algorithm starts here
bg_Ratio = (aoiRaster[[2]]/aoiRaster[[3]])
R = log10(bg_Ratio)

Chl_a = 10^((0.2511-2.0853*R + 1.5035*R^2-3.1747*R^3+0.3383*R^4))*12.4
# plotRGB(aoiRaster, 4,3,2, scale = 65535, stretch = 'lin')
plot(Chl_a,  main = "Bloom Date")
# Save file on disk
saveFile <- writeRaster(Chl_a, "OC2_Bloomdate.tif", format = "GTiff", datatype =
"FLT4S", overwrite = TRUE)
```

2. Mono-window LSAT Algorithm

```
#Start
================================================================================
// Computing for Lake Surface Air Temperature
// Filter your study area with the LakeVictoria shapefile
var roi = ee.FeatureCollection("users/okomoJacob/finalYrProj/LakeVictoria");
// Cloud masking of the collection
function maskL8sr(col) {
  // Bits 3 and 5 are cloud shadow and cloud, respectively.
  var cloudShadowBitMask = (1 << 3);
  var cloudsBitMask = (1 << 5);
  // Get the pixel QA band.
  var qa = col.select('pixel_qa');
  // Both flags should be set to zero, indicating clear conditions.
  var mask = qa.bitwiseAnd(cloudShadowBitMask).eq(0)
                 .and(qa.bitwiseAnd(cloudsBitMask).eq(0));
  return col.updateMask(mask);
}

//Obtain the L8 image collection in to GEE script.
 {
var col = ee.ImageCollection('LANDSAT/LC08/C01/T1_SR')
    .map(maskL8sr)
    .filterDate('YYYY-MM-DD','YYYY-MM-DD')
    .filterBounds(roi);
}
print(col, 'Image collection');

//Obtain the image median to reduce collection
var image = col.median();
print(image, 'medianImage');
Map.addLayer(image, vizParams2);

//Compute for the NDVI
var ndvi = image.normalizedDifference(['B5',
    'B4']).rename('NDVI');
    var ndviParams = {min: -1, max: 1, palette: ['blue', 'white',
    'green']};
    print(ndvi,'ndvi');
    Map.addLayer(ndvi, ndviParams, 'ndvi');

//select thermal band 10(with brightness tempereature), no calculation
var thermal= image.select('B10').multiply(0.1);
var b10Params = {min: 291.918, max: 302.382, palette: ['blue',
```

```
'white', 'green']};
Map.addLayer(thermal, b10Params, 'thermal');
// Compute for the Maximum and Minimum NDVI
var max = ee.Number(ndvi.reduceRegion({
    reducer: ee.Reducer.max(),
    geometry: roi,
    scale: 30,
    maxPixels: 1e9
}).values().get(0));
print(max, 'minNdvi')

var min = ee.Number(ndvi.reduceRegion({
    reducer: ee.Reducer.min(),
    geometry: roi,
    scale: 30,
    maxPixels: 1e9
}).values().get(0));
print(min, 'minNdvi');

//Computing the Fractional Vegetation
{
var fv
=(ndvi.subtract(min).divide(max.subtract(min))).pow(ee.Number(2)).rename('FV'
);
print(fv, 'fv');
Map.addLayer(fv);
}

//Lake Surface Emissivity
var a= ee.Number(0.004);
var b= ee.Number(0.986);
var EM=fv.multiply(a).add(b).rename('E');
var imageVisParam3 = {min: 0.9865619146722164, max:0.989699971371314};
Map.addLayer(EM, imageVisParam3,'E');

//Compute LSAT in Celsius Degree
var LST = thermal.expression(
    '(Tb/(1 + (0.00115* (Tb / 1.438))*log(Ep)))-273.15', {
    'Tb': thermal.select('B10'),
    'Ep': EM.select('E')
}).rename('LSAT');

Map.addLayer(LST, {min: 20.569706944223423, max:29.328077233404645, palette:
[
'Your palette'
```

```
    ]},'LSAT');

   var LSTFINAL = LST.clip(roi);
   Map.addLayer(LSTFINAL, {min: 20.569706944223423, max:29.328077233404645,
   palette: [
   'your palletes'
    ]},'LSAT FINAL');

    // Export the Final LSAT Result to Google Drive
   Export.image.toDrive({
   //
   });
#End ================================================================================
```

3. Model Accuracy Assessment Programme

```
'''
#Start ==========================================================================
This model accepts a .csv data with at least two columns whose correlation co-
efficient is to be determined.
Install the listed python packages below and import them in to your model as
below.
'''
#1. Import the necessary packages
from matplotlib import pyplot as plt
from sklearn import linear_model
from matplotlib import style
style.use('seaborn-whitegrid')
import pandas as pd
import numpy as np

#2. Import the CSV data into the model and inspect it
data = pd.read_csv('path/to/yourData.csv', index_col = 0)
data.shape
data.dropna(inplace = True)
data.head(3)


#3. Obtain the Correlation Co-efficient and scatter plot the data
data.corr()
data.plot(kind='scatter', x = "L8 Chl-a Estimates", y="S3 OLCI Chl-a
(Reference)")
plt.show()

#4. Change the csv data into a pandas readable dataframe
estimates = pd.DataFrame(data["S3 OLCI Chl-a (Reference)"])
reference = pd.DataFrame(data["L8 Chl-a Estimates"])

#5. Build linear model
lm = linear_model.LinearRegression()
model = lm.fit(estimates, reference)
model.coef_
model.intercept_


#6. Model Evaluation
model.score(estimates, reference)

#7. Predict new value of estimaes
```

```
estimatesNew = 56
estimatesNew = np.array(estimatesNew).reshape(1, -1)
estimatesPred = model.predict(estimatesNew)
estimatesPred

#8. Predict more values
X = ([67, 58])
X = pd.DataFrame(X)
Y = model.predict(X)
Y = pd.DataFrame(Y)
df = pd.concat([X, Y], axis = 1, keys = ['estimatesNew', 'estimatesPred'])
df

#9. Visualize the results
data.plot(kind = "scatter", x = 'L8 Chl-a Estimates', y = "S3 OLCI Chl-a
(Reference)")

#10.Plot the best-fit regression line and append title
plt.plot(estimates, model.predict(estimates), color = "red", linewidth = 2)
plt.title("CORRELATION BTWN ESTIMATED Chl-a AND REFRENCE Chl-a FOR 2015")
plt.show()

#END ============================================================================
```

4. The complete IoT System

The system works but combining two dependent progrems that work simulatoesnly to (a) get the system's location and (b) collect data and send

(a) Get System Location

```
''' #Start=============================================================
This program runs only on a RaspberryPi Operating Sytem. You'll acquire the
Pi Hardware to run the same
Then install the required packages below. Then hook up the GPS sensors.
Power the system at 5V then run the algorithms to get your live GPS location
'''
# Name this file getLocation.py and save both files in the same location.
from signal import signal, SIGTERM, SIGHUP, pause
import time
import pynmea2
import string
import serial

def safe_exit(signum, frame):
    exit(1)
signal(SIGTERM, safe_exit)
signal(SIGHUP, safe_exit)

try:
    getLocation = True
    while getLocation: #run forever
        port = "/../../dev/ttyAMA0" #Nav to where NMEA was initiated(/)
        ser = serial.Serial(port, baudrate=9600, timeout=0.5)
        dataout = pynmea2.NMEAStreamReader()
        newdata = ser.readline() #readline from the serial connection

        if newdata[0:6] == "$GPGLL": #get GPsGlobal data as Lat, Long
            newmsg = pynmea2.parse(newdata)
            lat = newmsg.latitude
            lng = newmsg.longitude
            gps = "Lat = " +str(lat) + ", Long =" + str(lng)
            print(gps)

except KeyboardInterrupt:
    pass
finally:
    pass
#END#===================================================================
```

(b) Package the *in-situ* data, append location information to it and send

```
'''
Complete IoT System:
================================================================================
1.Collects Air Temp from DHT11 LSAT sensor
2.Blinks LED to indicate functionality
3.Populates excel sheet with this data
4.Create an account with Africastalking to get the messaging API
4.Send SMS in case of anomaly
'''

# !/usr/bin/env python
# Import required libraries

from openpyxl import load_workbook #To manipulate spreadsheet
import africastalking as at
from datetime import date
import RPi.GPIO as GPIO
import dht11
import time #to delay your LED, LSAT collection timing
from . import getLoaction as getLoc

#Initialize Africastalking Messaging API
sms_username = "userName"
sms_api_key = "API KEY HERE"
at.initialize(sms_username,sms_api_key)
recepients = recipients = ['RecepinetsPhoneNumbers, emails@mail.mail, etc']

currentLoaction = getLoc.getLocation()
# initialize GPIO
GPIO.setwarnings(False)
GPIO.setmode(GPIO.BCM)

ledPin = 4 #Use pin 7 (GPIO 4)
GPIO.setup(ledPin, GPIO.OUT)

# read dht11 data using GPIO pin 17
instance = dht11.DHT11(pin=17)
# System starts

print("Connection Secured!")
time.sleep(2.5)
print("Connecting to GPS...")
```

```
time.sleep(6)
print("LSAT collection start...")
#Flashing
#while True: #Run forever

# Load data into workbook and select the sheet
wb = load_workbook('/home/pi/testEnv/lsatDb.xlsx')
sheet  = wb['Sheet1']

try:
    while True:
        GPIO.output(ledPin, GPIO.HIGH)
        time.sleep(0.25)
        #  print("Going Off")
        GPIO.output(ledPin, GPIO.LOW)
        time.sleep(0.25)

        result = instance.read()
        if result.is_valid():
            lsat = result.temperature
            today = date.today()
            now = datetime.datetime.now().time()

            print("Last LSAT Updated: " + str(datetime.datetime.now()))
            print("Lake Surface Air Temp: %-3.1f °C\n" % lsat)

            # Append data to sheet
            row = (today, now, lsat)
            sheet.append(row)

            # Save workbook
            wb.save('/home/pi/testEnv/lsatDb.xlsx')
            time.sleep(0.5)

        def send_sms():
            sms = at.SMS
            response = sms.send([
                    "LSAT anormally detected!\n \
                    Lake Surf_Air Temp: %-3.1f °C and rising\n Water
Salinity: NOT_SET.\
                    Sensors Location: %currentLoaction.\nView Map:
https://www.google.com/maps?q%currentLoaction.\ Water conditions: Needs
Inspection!" % lsat], recepients)
            print(response)
```

```
        # main meat execution
        if result.temperature == 28.6:
            send_sms()
except KeyboardInterrupt:
    #Confirm save of Excel Workbook to local Storage.
    wb.save('/home/pi/testEnv/lsatDb.xlsx')
    print("Safe Exit")
    GPIO.cleanup()
#END
#================================================================================
```