

**DERIVING ALGAL CONCENTRATION  
FROM SENTINEL-2 THROUGH A  
DOWNSCALING TECHNIQUE: A CASE  
NEAR THE INTAKE OF A  
DESALINATION PLANT**

**COLLINS ACHEAMPONG**


February, 2018

**SUPERVISORS:**

**DR. IR. MHD. SUHYB SALAMA**

**DR. C.M.M MANNAERTS**





**DERIVING ALGAL CONCENTRATION FROM  
SENTINEL-2 THROUGH A DOWNSCALING  
TECHNIQUE: A CASE NEAR THE INTAKE OF  
A DESALINATION PLANT**

**COLLINS ACHEAMPONG**

Enschede, The Netherlands,

February, 2018

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

Specialization: Water Resources and Environmental Management

**SUPERVISORS:**

Dr. ir. Mhd. Suhyb Salama

Dr. ir. C.M.M Mannaerts

**THESIS ASSESSMENT BOARD:**

Prof. Dr. D. van der Wal (Chair)

Dr. Jaime Pitarch Portero (External Examiner, NIOZ-TEXEL, The Netherlands)

## **DISCLAIMER**

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

## ABSTRACT

Desalinated water contributes to about 90 % of the potable water demand of Gulf Countries due to their characteristic arid climate. However, high level of algal biomass associated with Harmful algal blooms (HABs) have been a major hindrance to the full operation and functioning of the seawater desalination plants. Estimation of chlorophyll-a (Chl-a) concentration from existing satellites images are used as indicators for the monitoring of algal blooms known to be present in coastal waters but at coarser scale. However, high spatial resolution Chl-a maps can be derived directly from Sentinel-2 MSI with Neural Networks algorithms. However, the spectral set-up of Sentinel-2 limits its applicability to accurately estimate Chl-a using such models.

Hence, in this study, a new method of regression analysis model is utilized to retrieve Chl-a by relating Sentinel-2 MSI derived indices with MODIS L2 Chl-a product to obtain Chl-a maps at high spatial and spectral resolution. Three existing semi-empirical algorithms for retrieving Chl-a consisting, the Red Tide Index (RI) and Red Tide Index Chlorophyll Algorithm (RI/RCA), Maximum Chlorophyll Index (MCI) and Normalized Difference Chlorophyll Index (NDCI) based on band ratio are calibrated and validated using two independent datasets of Chl-a data from MODIS Aqua L2 1 km observed over the Arabian Gulf (AG) from 27 December 2015 to 27 October 2017.

Preliminary results revealed 9 x 9 as the optimal window size for extracting data for model development. RI/RCA model run with exponential fit as the best model for estimating Chl-a than the other two models. We obtain an  $R^2$  of 0.507, a root-mean-square error (RMSE) of  $2.75 \text{ mg/m}^3$  and mean absolute percentage error (MAPE) and 43.4 % respectively were estimated for the entire range of Chl-a concentration ( $\sim 0.48 \text{ mg/m}^3$  to  $18.42 \text{ mg/m}^3$ ) modelled compared with MODIS observed Chl-a data. Also, retrieved map with RI/RCA algorithms resolved into more details the spatial variation in Chl-a concentration near the intake of Ras Abu Fontas desalination plant compared with MODIS OC3M Chl-a maps. Results of MCI model also showed potential for Chl-a retrieval.

Seasonal analysis of MODIS observed Chl-a and Sentinel-2 derived Chl-a concentration showed summer maximum and spring minimum. The residual error in upscaled Sentinel-2 MSI derived Chl-a versus the reference MODIS L2 downgraded each at 90 m and 1 km, ranges between  $-7.2$  to  $7.5 \text{ mg/m}^3$ . The results of RI/RCA showed more sensitivity to Chl-a more than MODIS OC3M. In all, the preliminary results revealed that, RI/RCA model derived through the downscaling technique is more sensitive to Chl-a in the AG and has the potential of being used as an indicator for algal bloom monitoring near the desalination plant intakes.

Keywords: Sentinel-2 MSI; Chlorophyll-a, MODIS Chl-a product, HABs; ACOLITE; Arabian Gulf

## ACKNOWLEDGEMENTS

I enjoyed the support, help and encouragement of many people and institutions within the 18 months programme and in conducting this research.

First and foremost, I would thank my supervisor, Dr ir. Mhd. Suhyb Salama, for his advice and thoughtful suggestions on the work. Your relentless support in conceptualizing the idea for this research is much appreciated. I also thank my second supervisor, Dr ir. C.M.M Mannaerts, for his encouragement and guidance were instrumental and helpful for finalising this research.

I am highly indebted to OPEC Fund for International Development (OFID) for the scholarship award offered me to pursue this programme. Not forgetting, University of Twente ITC Excellence Scholarship Award granted me for my entire study period. I say thank you to OFID and ITC for giving me these awards in support of my 18 months study for the Water Resources and Environmental Management programme at the Faculty of Geo-Information Science and Earth Observation, University of Twente.

My heart goes out to my dear family members for their love, care, support, prayers and encouragement all through this point and especially to my dear sister Comfort Acheampong, without you my study would not have been possible. You form a great and unified group of which I am delighted to be part.

Finally, I say thanks to someone special, Lilian Aboagyewaa, and to my best pal Prince Kwabena Boateng who offered me support and advice throughout the study and everyone who in diverse ways contributed to the success of this study.

# TABLE OF CONTENTS

---

ABSTRACT.....	i
ACKNOWLEDGEMENTS.....	ii
1. INTRODUCTION.....	1
1.1. Research Problem.....	3
1.2. Research Objectives.....	4
1.2.1. Specific objectives.....	4
1.3. Research questions.....	4
2. LITERATURE REVIEW.....	5
2.1. Harmful Algal Blooms (HABs) detection in the Arabian Gulf (AG) regions.....	5
2.2. Remote sensing of algal biomass.....	5
2.3. Applications of MODIS for chlorophyll-a mapping.....	5
2.4. Sentinel-2 MSI for Chlorophyll-a retrieval.....	6
2.5. Chlorophyll-a retrieval algorithms.....	6
2.6. Chlorophyll-a concentration estimation through a downscaling approach.....	7
3. Description of Study Area and Data Sets.....	9
3.1. Study area.....	9
3.1.1. General description.....	9
3.1.2. Geographical description of study area.....	9
3.2. Satellite Datasets.....	11
3.3. MODIS Level-2 Chlorophyll product.....	11
3.3.1. Sentinel-2 MSI Datasets.....	11
4. METHODOLOGY.....	13
4.1. Overview.....	13
4.2. Satellite data preprocessing.....	15
4.2.1. Matchup of MODIS Level-2 Chl-a products.....	15
4.2.2. Atmospheric correction (AC) of Sentinel-2 MSI images with ACOLITE algorithm.....	15
4.3. Chlorophyll-a algorithms development.....	16
4.3.1. Maximum Chlorophyll index (MCI).....	17
4.3.2. Normalised Difference Chlorophyll-a Index (NDCI).....	18
4.3.3. Red-tide index Chlorophyll-a Model (RI/RCA).....	18
4.3.4. Aggregation of Sentinel-2 Chl models.....	19
4.4. Calibration of chlorophyll-a retrieval models.....	19
4.5. Accuracy assessment of models validation results.....	21
4.6. Estimation of chlorophyll-a.....	21
4.7. Evaluating the effect of spatially aggregation on spatial variation of retrieved Chl-a products.....	22

4.8.	Consistency check in cross-validation of spatial variation of Sentinel-2/MSI Chl-a product ...	22
4.9.	Chlorophyll-a retrieval with Neural Network models .....	23
5.	RESULTS AND DISCUSSION .....	24
5.1.	Model calibration and Validation .....	24
5.2.	Model validation results.....	25
5.2.1.	Selecting the critical aggregation level .....	27
5.2.2.	Comparison of accuracy results with standards.....	28
5.3.	Chlorophyll-a retrieval with Neural Network models .....	28
5.4.	Spatial variation of retrieved chlorophyll-a (Chl-a) product vs MODIS L2 Chl-a product .....	29
5.5.	MODIS and Sentinel-2 Chlorophyll-a maps as indicators for algal bloom mapping.....	32
5.6.	Chlorophyll-a retrieved with C2RCC Neural Net algorithm .....	33
5.6.1.	Time series of derived Chl-a concentration .....	33
5.7.	Evaluating the Atmospheric Correction method .....	35
5.8.	Comparing results of aggregated and disaggregated Chl-a products.....	36
5.9.	Residual map for MODIS and Sentinel-2 Chlorophyll-a maps .....	37
6.	CONCLUSION AND RECOMMENDATIONS.....	39
6.1.	Conclusion .....	39
6.2.	Recommendations .....	41
	REFERENCES .....	42
	APPENDIX.....	50



# LIST OF FIGURES

---

Figure 1: Map showing the Arabian Gulf, the study area and matchup sites of Multispectral images and MODIS L2 ocean colour product; red box indicates study region ..... 10

Figure 2: Flowchart for deriving chlorophyll-a maps ..... 14

Figure 3: Chlorophyll-a concentration retrieval algorithms based on Sentinel-2/MSI derived indices RI/RCA (a), (b) MCI and NDCI (c) run with 9 x 9 moving window size. .... 25

Figure 4: scatterplots of validation set of Sentinel-2 derived chlorophyll-a by RI, MCI and NDCI against MODIS L2 chlorophyll-a data: a – p (col 1) by RI model; b – q (col 2) by MCI model and c - r (col 3) by NDCI model; ..... 27

Figure 5: Statistical summary of scatter plots of Sentinel-2 derived chlorophyll-a concentration against MODIS L2 chlorophyll-a measurement for the validation datasets of each average spatial window size for RI model. .... 28

Figure 6: Comparison of Chl-a products derived with C2RCC and (a) MODIS L2; (b) Sentinel-2 RI/RCA Chl-a concentration..... 29

Figure 7: Chlorophyll-a concentration (Chl-a) derived from Sentinel-2 RI/RCA model for 11 scenes from 27 December 2015 to 27 October 2017. All images were processes with ESA Snap software (5.00) ..... 30

Figure 8: Seasonal trend of MODIS Aqua chlorophyll-a image derived with OC3M algorithm for 11 scenes from 27 December 2015 to 27 October 2017. .... 31

Figure 9: Sentinel-2 MSI Chl-a images derived with C2RCC Neural Network algorithm. .... 33

Figure 10: Time series of NASA MODIS satellite OC3/OC4 derived chlorophyll-a versus Sentinel-2 MSI RI/RCA retrieved chlorophyll-a ( $\text{mgm}^{-3}$ ) averaged over 36 data stations ..... 34

Figure 11: Comparison between MODIS Ocean color reflectance products (Rrs) and Sentinel-2 atmospheric corrected reflectance product (Rrs) where 443/443 (a); 488/490 and 547/560 for MODIS and Sentinel-2/MSI respectively for all 12 images obtained. .... 35

Figure 12: Comparison of Sentinel-2/MSI against MODIS L2 Chl-a product at 90 m critical aggregate level (a) and at 1 km MODIS spatial resolution (b) for images acquired on 22 September 2017 ..... 36

Figure 13: Comparison of MODIS reference Chl-a and Sentinel-2 derived Chl-a at 90 m for (a) and (b); and 1 km for (c) and (d): Small triangles represent intakes of desalination plants; top (intake of RAS A3 (SWRO system) and down intake 3 is the intake of Ras A2 (multi-stage flash distillation (MSF))..... 37

Figure 14: Spatial variation in residual error of Sentinel-2 derived Chl-a aggregated to (a) 90 and (b) 1 km for image acquired on 22 September 2017 ..... 38

## LIST OF TABLES

---

Table 1: Bands of Sentinel-2 MultiSpectral Imager (MSI) adapted from paper (Drusch et al., 2010; Vanhellemont & Ruddick, 2016). .....	12
Table 2: Downloaded matchup products of MODIS Chl-a product and Sentinel-2 MSI.....	12
Table 3: Structure and bands adopted by the different band ratio algorithms used in the study. Sentinel-2 bands used in this study; where; $\lambda_1=443$ , $\lambda_2=490$ , $\lambda_3=560$ , $\lambda_4=665$ , $\lambda_5=705$ , and $\lambda_6=740$ with MCI =Maximum Chlorophyll index, NDCI=Normalized Difference Chlorophyll index and RI= Red Tide index .....	19
Table 4 : Statistical summary of scatter plots of Sentinel-2 derived chlorophyll-a concentration against MODIS L2 chlorophyll-a measurement for the validation datasets of each average spatial window size...27	

## ABBREVIATIONS

---

AG	Arabian Gulf
Chl-a	Chlorophyll-a
CDOM	Coloured dissolved organic matter
EEZ	Qatar Marine Exclusive Economic Zone
HAB	Harmful algal bloom
IOC	Intergovernmental Oceanographic Commission
MAPE	Mean absolute percentage error
MCI	Maximum chlorophyll index
MERIS	Medium Resolution Imaging Spectrometer
MODIS	Moderate Resolution Imaging Spectrometer
MSI	Multispectral Imager Spectrometer
NN	Neural Networks
NDCI	Normalised difference chlorophyll index
NFLH	Normalised Fluorescence Line Height
NIR	Near Infra-red
RCA	Red tide index chlorophyll algorithm
RI	Red tide index
RMSE	Root mean square error
RT	Red Tide(s)
SNR	Signal to noise ratio
OLI	Operational Land Imager



# 1. INTRODUCTION

The Arabian Gulf (also called the Persian Gulf), surrounded by a hyper-arid and dry region in the Middle-East is a semi-enclosed shallow basin with a mean depth of approximately 35 m and a width of 370 km. The Arabian Gulf is classified as coastal case II water under a dusty environment characterized by extreme climatic weather condition where evaporation exceeds precipitation resulting in hypersaline water (Al-Ansari et al., 2015; Nezlin et al., 2007). Many arid countries along the Arabian Gulf (AG) are increasingly dependent on seawater desalination due to an increasing potable water demand for drinking purpose and the prevailing arid climate. Desalinated water helps in meeting good portions of freshwater needs (about 99% in Qatar, 96% in Kuwait, and 60% in Saudi Arabia) in the Arabian Gulf countries (Darwish, Abdulrahim, Hassan, & Shomar, 2016). A critical threat to this freshwater demand is from harmful algal blooms (HABs).

“Harmful algal bloom (HAB) can be described as any sudden and rapid growth of any monospecific algae or heterotrophic organism as a consequence of ecological perturbation and concomitant imbalance in the nutrients of a given water body” (Anderson, Glibert, & Burkholder, 2002; Heisler et al., 2008; Olalekan & Malik, 2015). The prolific growth in algae found in coastlines responsible for the water surface decolourisation is caused mainly by two types of HABs including; red tides (RT) (dinoflagellates) and cyanobacteria (blue-green algae) blooms. While not all algal species are harmful, some produce potent neurotoxins (such as cyanobacteria and diatoms) which are harmful to aquatic living organisms and poses threats to the health of humans if they persist in the treated water. Majid et al., (2016) reported the presence of cyanobacteria releasing cyanotoxins within 80 % of sampled water bodies as well as other water storage reservoirs (such as impoundments) in the Arabian Gulf. Moreover, samples of desert soils tested revealed the presence of neurotoxins that has the potential to bioaccumulate and the subsequent leaching into the aquifer with a potential of polluting the groundwater resources (Majid et al., 2016). These therefore render water sources unwholesome for consumption or irrigation purpose.

The Arabian Gulf (AG) has historically been susceptible to HABs and HABs have occurred more frequently in recent years both in the AG and around the world (Darwish et al., 2016). Also, cases of massive harmful algal blooms have been reported by Al Shehhi et al., (2014) in Qatar and the Southwest of India between 1996 and 1998, and along the Coast of Kuwait in October 1999 and September 2001. The impact of HABs can be detrimental to desalination plant operations as it was seen during the 2008-2009 severe red tide episode in the Gulf of Oman that triggered concerns of irreversible membrane fouling problems and eventually led to the shutdown of some plants due to clogging and reduce operations for more than 55 days (Anderson & Price, 2015). The global and regional expansion of HABs is occurring at a time when there is also an increase in the construction of desalination plants to produce drinking water (Villacorte et al., 2015) and therefore more HABs impacts are probable to re-occur. The high level of algal biomass associated with

severe algae bloom decomposition at the intake results in the suspension of desalination plant operations due to the high cost involved in the chemical treatment of toxins and malodours in the water. It is therefore very important to mitigate the impacts of such harmful algal blooms before they cause any devastating impacts to humans and the marine ecosystem.

In the Arabian Gulf, it has been a difficult task with the traditional method of constant monitoring of HABs along the Gulf as well as near intakes of seawater desalination plants. Previous studies employing traditional point monitoring or measurements of bio-optical water parameters do not reflect the spatial variability of the water constituents effectively. Moreover, various studies utilizing in-situ measurements and satellite products have been utilized in mapping chlorophyll-a (as proxies for algal blooms) and other bio-optical water properties. But the processes responsible such as nutrient inflow, light penetration, hydrodynamics and water temperature and turbidity for bloom initiation and transport at a localized area are difficult to model and monitor (Conley et al., 2009; Elkadiri et al., 2016). It is therefore of high priority to monitor and forecast the blooms development and movement before they cause any catastrophic effects.

Several studies have used satellite data in monitoring the occurrence of HABs in the Arabian Gulf utilizing these coarser satellite resolution images including Moderate Resolution Imaging Spectroradiometer (MODIS), Medium Resolution Imaging Spectrometer (MERIS) and Seaviewing Wide Field-of-view Sensor (SeaWiFS) (Al-Shehhi, Gherboudj, Jun, Mezhoud, & Ghedira, 2013; Ghanea, Moradi, & Kabiri, 2015; Mohsen Ghanea, Moradi, & Kabiri, 2016; Zhao, Ghedira, & Temimi, 2014; Moradi & Kabiri, 2012; Zhao, Temimi, Al-Azhar, & Ghedira, 2015). In some case, these space-based sensors have also provided huge and varied amount of data used for the geophysical validation of hydrodynamic models by many researchers. However, their coarser resolution does not allow for estimation of water quality constituents at a fine scale. Therefore, the usefulness of high spatial resolution sensors for monitoring coastal and inland water has become apparent with new applications quickly emerging (Vanhellemont & Ruddick, 2016b).

Monitoring HABs at a fine scale is significant for environmental, ecological, and biological management of water regions. With a high spatial resolution (10-60 m) of Sentinel-2 MSI satellite image with improved spectral resolution from MODIS, turbid tidal wakes together with chlorophyll-a and other water quality parameters can effectively be retrieved and analysed in more detail and at a fine scale for optically complex coastal waters. Also, smaller water bodies not captured by coarser satellite images can effectively be delineated together with the retrieval of its bio-optical properties.

## 1.1. Research Problem

Globally, chlorophyll-a concentration plays an important role as an indicator of the quality of water bodies. They are also used as proxies for characterizing harmful algal blooms (HABs). During periods of HABs, algal blooms cover large areas whereby chlorophyll-a of very high concentration exceeding  $10 \text{ mg/m}^3$  are observed. When chlorophyll-a concentration of coastal waters exceeds this threshold ( $10 \text{ mg/m}^3$ ), they are regarded as harmful algal bloom (Comprehensive Studies Task Team of Group Coordinating Sea Disposal Monitoring [CSTT], 1997; World Health Organization [WHO], 2003). For effective monitoring and prediction of harmful algal blooms long period of algal bloom or phytoplankton biomass data are required. However, the scarcity of available in-situ measured data due to low spatial coverage over short sampling periods, time constraints and cost involved limits this application (Al-Naimi, Raitos, Ben-Hamadou, & Soliman, 2017).

Therefore, the integration of Sentinel-2 and or Landsat 8 and MODIS and/or MERIS at high spatial and temporal scale respectively may be more appropriate for retrieving chl-a. However, there are limitations in these existing satellites products since revisit time for the fine spatial resolution sensors is typically poor, while those with a high revisit frequency are characterized by a coarse spatial resolution (Bisquert, Sanchez, & Caselles, 2002). Moreover, the long repeat cycle of MSI, also makes it difficult in observing water quality changes due to its vulnerability to heavy dust storms, clouds, water vapor and haze effects which is common in the Arabian Gulf region especially during summer (June-August) (Al-Naimi, et. al, 2017; Mohsen Ghanea et al., 2016).

Currently (i.e., in mid 2015), there are two well-calibrated global ocean color sensors in operation, Moderate Resolution Imaging Spectroradiometer on Aqua (MODISA) and Visible Infrared Imaging Radiometer Suite (VIIRS) on Suomi NPP (Kahru, Kudela, Anderson, & Mitchell, 2015) that can be used for monitoring and forecasting blooms in data scarce areas. MODIS Ocean colour data at 1km (daily), 4km (monthly) or 9km (monthly) is capable of being used in place of in-situ measurements since it has been successfully used to mapping chlorophyll-a concentration and other water quality parameters globally. The daily revisit time makes it suitable for observing small changes in chlorophyll concentration and the dynamics in algal bloom episodes.

To the researchers' knowledge, downscaling of MODIS Level-2 1 km Chlorophyll-a product with Sentinel-2 MSI data for the reliable estimate and monitoring of Chl-a and mapping of algae blooms at a high temporal and spatial scale near desalination plant intakes is not yet known. The spectral set-up of Sentinel-2 is not good enough in estimating chlorophyll-a in coastal and in-land water compared with MODIS since it was apparently designed for Land applications. Therefore, in this study, MODIS Level-2 chlorophyll-a product will be complemented with the very high spatial resolution data of Sentinel-2 MSI to derive chlorophyll-a products of a high temporal and spectral resolution of the MODIS and the high spatial scale of Sentinel-2.

Also, retrieved chl-a maps will be cross validated with with derived chl-a concentration with the recently available Case 2 Regional coastcolor (C2RCC) neural network algorithm.

## **1.2. Research Objectives**

The main objective of this study is to derive chlorophyll-a concentration maps from Sentinel-2 at high spatial resolution from Sentinel-2 using MODIS 1 km Chl-a product as input and evaluate the spatial variability of algal biomass near the intake of Ras Abu Fontas desalination plant in Qatar.

### **1.2.1. Specific objectives**

The specific objectives of this study include;

- To derive chlorophyll-a concentration maps from Sentinel-2 at 10 m resolution near the intake of a desalination plant.
- To investigate the spatial variation of Chl-a concentration near the intake of the desalination plant from 2016-2017.

## **1.3. Research questions**

1. Can we derive Chl-a maps from Sentinel-2 with an acceptable accuracy w.r.t. MODIS?
2. What is the spatial variation of Chl-a concentration near the intake of the desal plant?
3. How the Chl-a maps are related to the operations of the desalination plants?



## 2. LITERATURE REVIEW

### 2.1. Harmful Algal Blooms (HABs) detection in the Arabian Gulf (AG) regions

Red tide, one type of harmful algal bloom, is caused by the prolific increase in toxic or nuisance algae species (Jun Zhao et al., 2014). According to Thangaraja et al. (2007), incidence of red tide blooms in the Persian Gulf is not something new due to recurrent events for some decades now. It is of great concern due to their adverse effects not only on the health of human and marine organisms, but also severe impacts on the economy of the affected areas (Jun Zhao et al., 2014). The Intergovernmental Oceanographic Commission (IOC), has identify 60–80 of algal species responsible for HABs out of which 75% are (Smayda, 1997; L. Villacorte et al., 2014). Till now, 38 taxa of HABs has been noted as present in the Arabian Gulf (AG) (Moradi & Kabiri, 2012; Thangaraja et al. 2001; Rao et al. 1998). In 2008 fall and winter season in the AG, a catastrophic red tide episode occurred which spread over and affected almost the entire the region. This red tide episode also resulted in the suspension of desalination plant operation and other marine activities.

Some well-known impacts of HABs include; scum formation (can be of very high thick layers) on the water surface causing the blockage of sunlight needed for photosynthetic activities by aquatic plants. And, in extreme cases, the high algal biomass produced depletes the dissolved oxygen content in the water for their metabolic activities leading to the loss of life of fishes and other aquatic animals. Additionally, red tide outbreaks force the shutdown of desalination plants, upon which the potable water supply in the Gulf region relies (Berktaş, 2011). The algal biomass and the associated organic load of red tide blooms lead to fouling of membrane surfaces and clogging of intake filters restricting the effective operations of seawater desalination plants. Therefore, huge costs are invested in cleaning or repairing of damage membranes filters and for flushing of clogged membrane.

### 2.2. Remote sensing of algal biomass

Temporal detection and observation of HABs is critical for environmental assessment, for ecological modelling, and for prediction and mitigation against red tide impacts (Moradi & Kabiri, 2012). Hence, it is very important that, satellite image with a combined high spatial and temporal resolutions is highly required for the detection, forecasting and monitoring of algal bloom dynamics on large scales. Most approaches for remote estimation of phytoplankton biomass are based on the absorption of sunlight by algal pigments in the presence of light scattering by algal cells and non-algal particles (Schalles, 2006). The difficulty is the determination if high Chl-a concentration levels are due to toxic or non-toxic species of algae.

### 2.3. Applications of MODIS for chlorophyll-a mapping

Numerous research have been performed utilizing coarser satellite data such as MODIS and MERIS for the estimation of Chl-a and the monitoring of Red Tides in the Arabian Gulf and the Sea of Oman (Moradi & Kabiri, 2012; Al-Shehhi, Gherboudj, Jun, Mezhoud, & Ghedira, 2013; Ghanea, Moradi, & Kabiri, 2015;

Mohsen Ghanea, Moradi, & Kabiri, 2016; Zhao, Ghedira, & Temimi, 2014; Zhao, Temimi, Al-Azhar, & Ghedira, 2015). Moradi & Kabiri, (2012) found a high correlation between the Normalised Fluorescence Line Height (NFLH) derived from MODIS and in-situ measurements of Chlorophyll-a during harmful algal bloom episodes. They also found high coefficient of determination ( $R^2$ ) of 0.53 between MODIS Modified Fluorescence Line Height (MFLH) and in-situ measurements. However, in some cases, especially during periods of low red tides, there have been reported cases of overestimation when MODIS retrieved results were compared with in-situ Chl-a measurements within the Arabian Gulf (Jun Zhao, Temimi, Kitbi, & Mezhoud, 2016). Al-Shehhi, Gherboudj, & Ghedira, (2017) attributed this poor performance of the coarser satellite image to three reasons namely; “(i) water turbidity (sediments re-suspension), and the presence of coloured dissolved organic matter (CDOM), (ii) bottom reflectance and (iii) incapability of the existing atmospheric correction models to reduce the effect of the aerosols from the water leaving radiance”.

#### **2.4. Sentinel-2 MSI for Chlorophyll-a retrieval**

The use of high spatial resolution satellite images such as SPOT 5, Landsat 5-TM, and Landsat 7-ETM+ for deriving water quality parameters at fine scale have been in existence for awhile now (Kabbara et al 2008). Algorithms designed for the recently launched Landsat-8 OLI and Sentinel-2 MSI in February 2013 and June 2015 respectively have offered an additional capability in retrieving water quality parameters such as chlorophyll-a at high spatial scale. However, it has been argued that, the spectral setup of Sentinel-2 with narrow red-edge spectral bands offers an additional advantage to Landsat-8 OLI for water quality retrieval in optically complex coastal areas (D’Odorico, Gonsamo, Damm, & Schaepman, 2013). Sentinel-2 MSI have successfully been utilised in retrieving water quality parameters including Chl-a, TSS and CDOM nearshore coastal waters and inland waters (Chen et al., 2017; Dörnhöfer, Göriz, Gege, Pflug, & Oppelt, 2016; Kutser et al., 2016; Liu et al., 2017; Toming et al., 2016). (Vanhellemont & Ruddick, 2016a) and Beck et al., (2017), have also recently demonstrated the capability of Sentinel-2 red-edge band at 705 nm in detecting, delineating and monitoring intense algal blooms. (Clark et al., 2017), also suggested integrating MERIS and Sentinel-2 MSI for monitoring water quality and surface algal blooms at high spatial resolution since MERIS is only accurately been able to resolve  $< 1\%$  of water bodies in U.S  $> 1$  ha.

#### **2.5. Chlorophyll-a retrieval algorithms**

Satellite-derived and laboratory measured Chl-a concentrations have widely been used as an index for algal bloom biomass and as a potential indicator of increased nutrients in inland and coastal waters. Remote sensing of algal biomass and its distribution has relied on algorithms to derive the concentrations of chlorophyll-a (Chl-a) in coastal case 2 waters. Traditional algorithms derived for the remote sensing of Chl-a concentration has been based on the blue-green ratio models. These models were first designed for monitoring vegetation using Normalized Difference Vegetation Index (NDVI). The NDVI models were based on bands in the blue-green regions to monitor vegetation health. Moreover, these blue-green spectral regions were used because, there is absorption of energy by Chl-a in the blue and red bands, resulting in

high reflectance in the green spectrum (Kamerosky, Cho, & Morris, 2015). Many studies have therefore use equations similar to the NDVI based on these spectral regions to successfully retrieve Chl-a from in-land and coastal water bodies. However, there have also been literature reports on these blue-green ratio algorithms failing to retrieve chlorophyll in case 2 waters.

Algorithms based on the Red-NIR regions has successfully been used in deriving Chl-a utilizing coarser satellite images in some inland and coastal waters by many researchers and several researches to retrieve high accurate Chl-a products are still on-going. During photosynthesis, Chl-a is known to fluoresce in the near infra-red (NIR) wavelength whereas most in-water algal blooms also show a peak in the red region. The IR reflectance is due to the decreasing absorption by the chlorophyll pigments in the spongy mesophyll cells and an absorption increase by water (Patra, Dubey, Trivedi, Sahu, & Rout, 2017). Therefore, Red-NIR algorithms have been useful in retrieving Chl-a products from satellite images.

There is no regionally tuned ocean colour algorithm for Chl-a retrieval in the Arabian Gulf region. The Red Tide index (RI), Maximum Chlorophyll-a index (MCI) and Normalised Difference Chlorophyll-a index (NDCI) models have been selected for use in testing the downscaling technique for estimation of Chl-a in this study. The models were chosen due to their wide and successful use in previous studies (Zolfaghari & Duguay, 2016; Tao, Mao, Wang, Lu, & Huang, 2011; Ahn & Shanmugam, 2006). These models have been successfully utilized in retrieving Chl-a from satellite images (such as CZCS, SeaWiFS, MODIS, MERIS Landsat sensors) in coastal zones of the Arabian Gulf (Jun et al., 2016; Moradi, Hasanlou, & Saadatersesht, 2016; Jun Zhao, Temimi, & Ghedira, 2015). They have successfully been applied in the retrieval of Chl-a concentration and for algal bloom monitoring globally. Also, because these algorithms are based on physical fundamentals and that they provide indices correlated with Chl-a concentrations. Moreover, previous studies conducted in the Arabian Gulf region recommended the use of blue-green algorithms (RI model) together with NIR-Red edge models (MCI and NDCI) for the accurate retrieval of chlorophyll-a concentration for the monitoring and prediction of red tides in the AG. This therefore influence the selection of these models.

## **2.6. Chlorophyll-a concentration estimation through a downscaling approach**

Information on the spatial extent and variation of algal blooms on a large scale (at both regional and local) is required for harmful algal bloom (HAB) monitoring and assessing health of the ecosystem. In recent years, various remote sensing approaches have been established based on coarser satellite resolution images for chlorophyll-a retrieval (usually used as proxies for HAB monitoring) at high temporal scale however at a low to moderate spatial scale. Ocean colour chlorophyll-a products at 1 km and 300 m at nadir derived from Moderate Resolution Imaging Spectroradiometer (MODIS) and Medium Resolution Imaging Spectrometer (MERIS) respectively (Arun Kumar, Babu, & Shukla, 2015), are used to monitor algal blooms over large coastal areas. Their coarser spatial scale however does not allow the rapid monitoring and quantification of algal bloom episodes near desalination plants intakes and for small water bodies into a

more detail. Therefore, Sentinel-2 MSI at 10-60 m can be used for near real-time monitoring of algal bloom dynamics at a fine spatial scale in the Arabian Gulf. However, the spectral set-up of the Sentinel-2 MSI is not more suitable for estimation of chlorophyll-a concentration compared with MODIS data which has the required spectral band accurate retrieval of water quality parameters.

Hence, a downscaling technique integrating the spectral and spatial properties of the two may be more appropriate (Atkinson, 2013), for the estimation of Chl-a concentration. That is, disaggregating methods provide a means where the coarser resolution satellite images with high spectral resolution and temporal coverage are downscaled to finer spatial scale or vice versa. Hence, in this study we evaluate the possibility of transferring the chl-a data contained in MODIS L2 1 km ocean product derived with OC3M algorithm onto Sentinel-2 MSI to derive Chl-a maps at 10 m resolution. This is tested through a downscaling technique of regression analysis of model development based on a linear, exponential and polynomial model.

## 3. DESCRIPTION OF STUDY AREA AND DATA SETS

### 3.1. Study area

#### 3.1.1. General description

The study area for this research is the Qatar marine Exclusive Economic Zone (EEZ) (Figure 1) of the Arabian Gulf that serves as an intake source for the Ras Abu Fontas A2 and A3 Seawater Desalination Plant. The Arabian Gulf with a geographical location of Latitude 26° 4' 35.47" N and Longitude 52° 37' 28.24" E (see Figure 1) is a small part of the Indian Ocean bounded by six (6) Gulf Coast Countries (GCC) (Oman, United Arab Emirates, Qatar, Saudi Arabia, Kuwait, and Bahrain). The AG has a surface area of 235,000 km<sup>2</sup>, elevation of -57 m, maximum depth of 90 m and a maximum length of 989 km. The Strait of Hormuz separates the AG from the Oman Sea, the Arabian Sea and Indian Ocean. This creates a varying temperature and salinity in these water bodies by inhibiting the exchange of marine water.

#### 3.1.2. Geographical description of study area

Qatar Peninsula lies on the southern shore of the Arabian Gulf (AG) that lies directly on the tropic of Cancer. About 99% of Qatar's freshwater demands comes from desalination plants. This is because at the AG evaporation exceeds precipitation due to the prevailing climate condition (arid) of the region. Evaporation is estimated to be 202.6 cm yr<sup>-1</sup>; where the monthly mean evaporation rate reaches a maximum of ~29.3 cm in June, and a minimum of 8.1 cm in February (Meshal and Hassan, 1986). Temperature exceed 35°C during summer (Al-Shehhi, Gherboudj, & Ghedira, 2017).

Ras Abu Fontas (RAF) Seawater Desalination Plant important for this study consist of the Ras A2 and Ras A3 Desal plant. The Ras A2 is a thermal desalination plant utilizing a multi-stage flash distillation (MSF) technology (<http://www.water-technology.net/projects/ras-abu-fontas-raf-a2-seawater-desalination-plant/>) located at about 10km south of Doha in Qatar along the Arabian Gulf which produces 160,000m<sup>3</sup>/day freshwater amounting to about 10% of Qatar's municipal water needs. While the Ras A3 is a Seawater Reverse Osmosis (SWRO) system that supplies potable water to about 1 million of Qatar inhabitants. The Ras A3 is the first large-scale SWRO system operational in Qatar. The desalination plants lie between latitude 25°06'N and longitude 51°37'E.

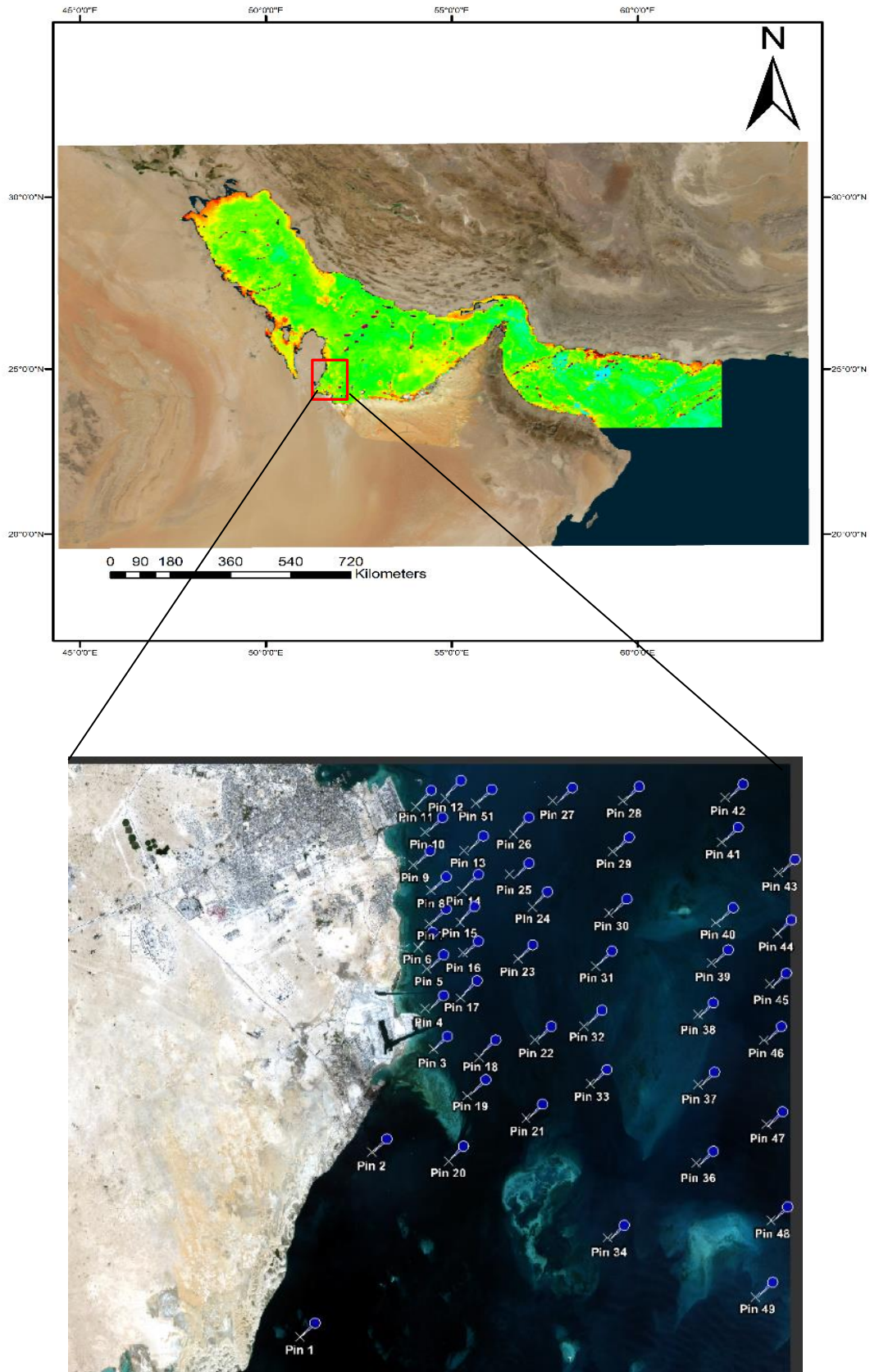


Figure 1: Map showing the Arabian Gulf, the study area and matchup sites of Multispectral images and MODIS L2 ocean colour product; red box indicates study region

### 3.2. Satellite Datasets

In this study, a total of 12 Sentinel-2 L1C images (from 2015 – 2017) coincident with MODIS L2 1 km chlorophyll-a products (already pre-processed) were obtained as shown in Table 2. The available images acquired were grouped into four main seasons representing the Arabian Gulf region (AG); Winter (December - February), Spring (March-May), Summer (June – September) and Autumn (October – November) (Al-Naimi et al., 2017). In all, matchup of Sentinel-2 and MODIS images acquired in the winter period were more than those acquired in summer and in spring.

### 3.3. MODIS Level-2 Chlorophyll product

The MODIS instrument on board the Terra and Aqua satellites has viewing swath width of 2330 km and a revisit period of one day with 36 spectral bands ranging in wavelength from 0.4  $\mu\text{m}$  to 14.4  $\mu\text{m}$  (Al Kaabi, Zhao, & Ghedira, 2016; Che, Feng, Jiang, Song, & Jia, 2015). These two platforms provide two daily passes of ocean color products over an area. A level-2 data product was made using the most recent update of instrument calibration, atmospheric correction and reprocessing algorithms (R2014.0). The ocean color products were freely downloadable at the NASA OceanColor website (<https://oceancolor.gsfc.nasa.gov/>) from late December 2015 to late October 2017. The data is produced by the NASA Goddard Space Flight Center's Ocean Data Processing System (ODPS) and distributed by the Ocean Biology Active Archive Center (OB.DAAC). Gordon and Wang (1994) atmospheric correction scheme based on the NIR (Zhao et al., 2014) are applied on MODIS L1 data from which MODIS L2 ocean colour product is obtained. The default chlorophyll-retrieval algorithm is a merged chlorophyll retrieval algorithm between the standard OC3/OC4 (OC3M) band ratio algorithm and the color index (CI) algorithm developed by Hu et al (2012) that returns the concentration of chlorophyll-a in  $\text{mgm}^{-3}$ , where CI and OC3M occurs at  $0.15 < \text{CI} < 0.2 \text{ mgm}^{-3}$ . The datasets of this product suite include; Rrs, Chl-a, Kd490, PAR, PIC, POC, etc (<https://oceancolor.gsfc.nasa.gov/docs/format/l2nc/>).

#### 3.3.1. Sentinel-2 MSI Datasets

Sentinel-2 Multi-spectral imager is freely downloadable at Copernicus Open Access Hub website (<https://scihub.copernicus.eu/>). Sentinel-2 MSI has 13 spectral bands of central wavelength range from 443–2190 nm (Table 1). The spatial resolution ranges from 10 – 60 m including four visible and near-infrared bands at 10 m, six red edge and shortwave infrared bands with spatial resolution at 20 m and three atmospheric correction bands of 60 m spatial resolution and a revisit period of Sentinel-2 is 2 -5 days (<https://directory.eoportal.org/web/eoportal/satellite-missions/c-missions/copernicus-sentinel-2>). The spectral bands of Sentinel-2 MSI and their wavelengths are shown in Table 1. Also, matchup datasets for the MODIS Level-2 chlorophyll-a product and Sentinel-2 images are given in table 2;

Table 1: Bands of Sentinel-2 MultiSpectral Imager (MSI) adapted from paper (Drusch et al., 2010; Vanhellemont & Ruddick, 2016).

Band	Central wavelength (nm)	Bandwidth (nm)	Spatial resolution (m)	SNR at reference L	Reference L (Wm <sup>-2</sup> sr <sup>-1</sup> um <sup>-1</sup> )
1	443	20	60	129	129
2	490	65	10	154	128
3	560	35	10	168	128
4	665	30	10	142	108
5	705	15	20	117	74.5
6	740	15	20	89	68
7	783	20	20	105	67
8	842	115	10	172	103
8a	865	20	20	72	52.5
9	945	20	60	114	9
10	1380	30	60	50	6
11	1610	90	20	100	4
12	2190	180	20	100	1.5

Table 2: Downloaded matchup products of MODIS Chl-a product and Sentinel-2 MSI

Acquisition Date	MODIS L2 chl product	Time (UTC)	Sentinel-2 MSI image	Time (UTC)
27/12/2015	A2015361094500.L2_LAC_OC	9:47	S2A_OPER_PRD_MSIL1C_PDMC_20151227T131949_R106_V20151227T072331_20151227T072331.SAFE	07:22
16/01/2016	A2016266100000.L2_LAC_OC	9:27	S2A_OPER_PRD_MSIL1C_PDMC_20160116T163256_R106_V20160116T072452_20160116T072452.SAFE	07:24
15/02/2016	A2016046093500.L2_LAC_OC	9:37	S2A_OPER_PRD_MSIL1C_PDMC_20160216T015819_R106_V20160215T071840_20160215T071840	07:28
22/09/2016	A2016296100000.L2_LAC_OC	9:47	S2A_OPER_PRD_MSIL1C_PDMC_20160922T124504_R106_V20160922T070622_20160922T071839.SAFE	07:22
22/10/2016	A2016296100000.L2_LAC_OC	10:12	S2A_OPER_PRD_MSIL1C_PDMC_20161022T122305_R106_V20161022T071302_20161022T071302	07:22
11/11/2016	A2016316094000.L2_LAC_OC	9:47	S2A_OPER_PRD_MSIL1C_PDMC_20161111T171508_R106_V20161111T071302_20161111T071302	07:29
11/12/2016	A2016346100000.L2_LAC_OC	10:02	S2A_MSIL1C_20161211T071302_N0204_R106_T39RWH_20161211T071654	07:24
31/12/2016	A2016366093500.L2_LAC_OC	9:37	S2A_MSIL1C_20161231T071302_N0204_R106_T39RWH	07:29
01/03/2017	A2017060100000.L2_LAC_OC	10:02	S2A_MSIL1C_20170301T071301_N0204_R106_T39RWH	07:29
22/09/2017	A2017265093000.L2_LAC_OC	10:02	S2B_MSIL1C_20170922T070609_N0205_R106_T39RWH_20170922T071248	07:22
27/09/2017	A2017270094500.L2_LAC_OC	9:47	S2A_MSIL1C_20170927T070641_N0205_R106_T39RWH_20170927T071322.SAFE	07:22
27/10/2017	A2017300100000.L2_LAC_OC	9:32	S2A_MSIL1C_20171027T071301_N0206_R106_T39RWH_20171027T105111	07:22



## 4. METHODOLOGY

### 4.1. Overview

This research is aimed at deriving chlorophyll-a concentration (as proxies for algal blooms monitoring) near the intake of the RAS Abu Fontas A2/A3 desalination plant through downscaling Level-2 1 km MODIS ocean colour product with Sentinel-2 data at 10 m resolution. In this study, the downscaling of MODIS L2 Chl-a data consist of mainly two main components shown in Figure 2. This includes first; the derivation of three (3) Sentinel-2 MSI Chl-a semi-empirical algorithms (NDCI, MCI and RI) based on band ratios for the qualitative and quantitative monitoring of algal blooms/biomass. And spatial aggregating (upsampling) the Chl-a indicators to match the spatial resolution of MODIS L2 1km Chl-a data. And, secondly, the establishment of the upscaled Chl-a indices (dependent variables) and MODIS Level-2 chl product (independent variable) relationship based on randomly selected 36-pixel location values (Figure 1) from each satellite image covering the study area. The derived model calibration coefficients based on the pixel domain of sampled locations are applied to the Sentinel-2 images to generate Chl-a maps at 10 m resolution.

The processing steps (Fig 2) followed to achieve the objectives of this study

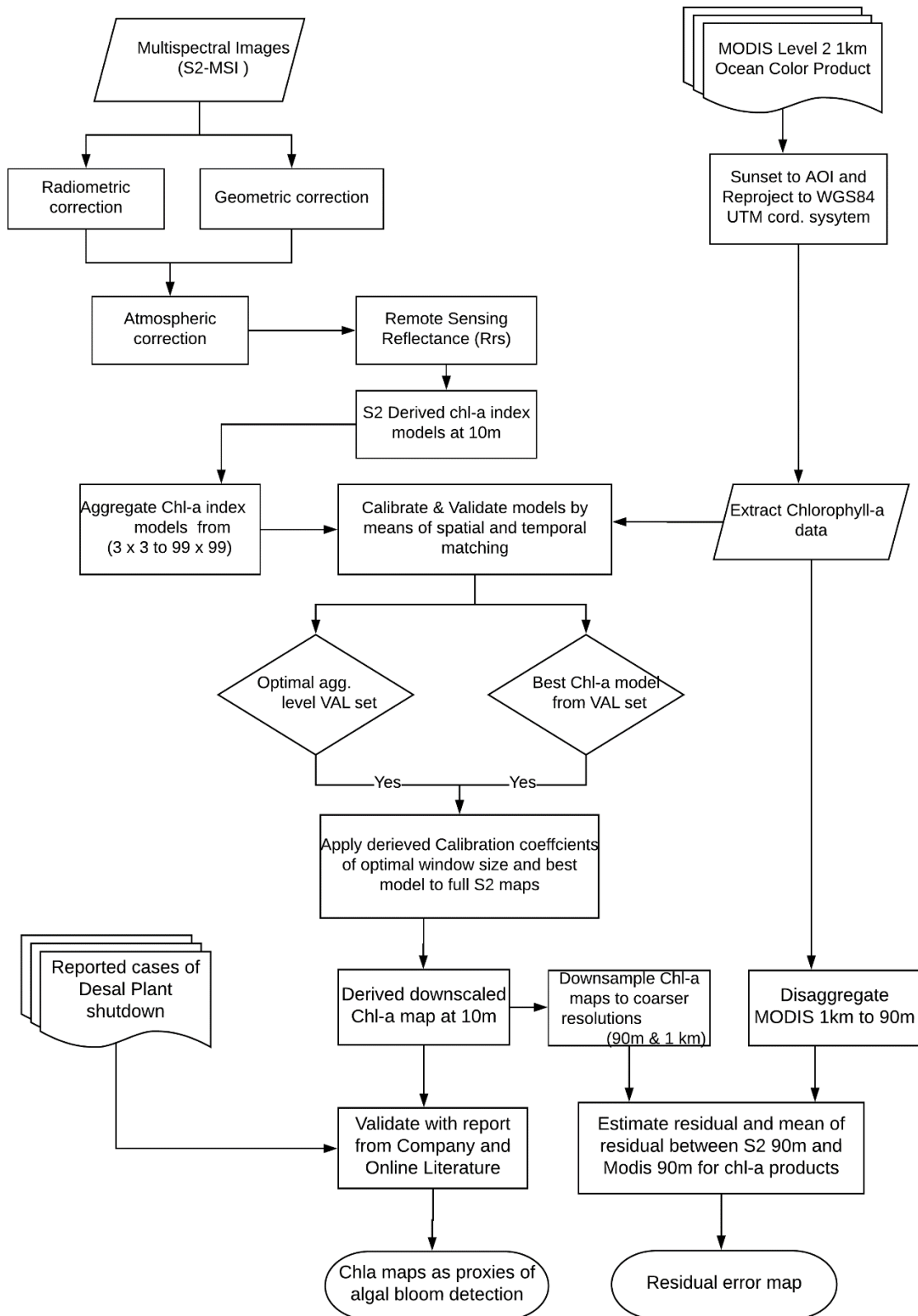


Figure 2: Flowchart for deriving chlorophyll-a maps

## 4.2. Satellite data preprocessing

### 4.2.1. Matchup of MODIS Level-2 Chl-a products

The MODIS chlorophyll-a product which is important for this research downloaded from NASA ocean colour website (<https://oceancolor.gsfc.nasa.gov/>) was derived using a combined algorithm of the current global OC3/OC4 (OCM) algorithm by (O'Reilly et al, 1998 and 2000) given as;

$$\log_{10}(\text{chlor}_a) = a_o + \sum_{i=1}^4 a_i \left( \log_{10} \left( \frac{Rrs(\lambda_{blue})}{Rrs(\lambda_{green})} \right)^i \right), \quad (1)$$

and the CI algorithm proposed by (Hu, Lee, & Franz, 2012) given as;

$$CI = \frac{Rrs(\lambda_{green}) - [Rrs(\lambda_{blue}) + (\lambda_{green} - \lambda_{blue})]}{(\lambda_{red} - \lambda_{blue}) * (Rrs(\lambda_{red}) - Rrs(\lambda_{blue}))} \quad (2)$$

where  $\lambda_{blue}$ ,  $\lambda_{green}$  and  $\lambda_{red}$  are the MODIS-specific wavelengths closest to 443, 555 and 670 nm respectively (<https://oceancolor.gsfc.nasa.gov/>).

The Level-2 MODIS chlorophyll-a product in sinusoidal projection was reprojected to the WGS 84 UTM coordinates of Sentinel-2 MSI data for all images downloaded. A subset of the L2 product covering the area of study is extracted.

### 4.2.2. Atmospheric correction (AC) of Sentinel-2 MSI images with ACOLITE algorithm

Accuracy of the atmospheric correction method is a requirement of almost all bio-optical properties retrieval based on reflectance ( $R_w$ ) of the water surface. The shortwave infra-red (SWIR) bands (20 m) at 1.6 and 2.2  $\mu\text{m}$  on board of Sentinel-2 allow for thorough atmospheric correction, over highly turbid waters (Vanhellemont & Ruddick, 2015; Vanhellemont, Vanderzande, & Ruddick, 2016). Hence, Vanhellemont & Ruddick, (2014) developed ACOLITE (in-built with IDL programming software) for the removal of contribution to reflectance of Sentinel-2 MSI images useful for the application on coastal or marine environments and inland case 2 waters. ALCOLITE uses NIR and SWIR bands for estimating the aerosol scattering on clear and turbid to highly turbid waterbodies respectively while Look up tables (LUTs) from 6SV model (Kotchenova, Vermote, Matarrese, & Klemm, 2006) is use for Rayleigh scattering correction. The Rayleigh-corrected reflectance is used to compute aerosol scattering. The performance of AC by ACOLITE using the SWIR at 1.6  $\mu\text{m}$  over water is due to the high absorption by the water at these wavelengths. Usually, reflectance values at these wavelengths are presumed to be zero (0) or close to zero in clear and even over turbid waters. Hence, the contribution of water to reflectance values may be assumed as negligible.

In ACOLITE the default threshold value for masking non-water pixels (masking out, sun-glints, artefacts, clouds and land) on the 1610 nm (1.6 $\mu\text{m}$ ) SWIR band is at 2.15% (0.0215). However, in this study, the default threshold value for masking non-water pixels was varied depending on the image pixel values for

the 1610 nm band of each Sentinel-2 image downloaded. This because, the water leaving reflectance varied between the summer periods (with high incidence of algal blooms) and that of the winter months. Also, SWIR band at 1.6 $\mu$ m of Sentinel-2 is noted to be of low signal-to-noise ratio (SNR) and therefore, a spatial smoothing window over an extent of 32 m with dilation 16 m was applied to reduce the noise in the output product (Vanhellemont & Ruddick, 2016). After processing the image with ACOLITE, the output product of Sentinel-2 MSI at 10 m grids are obtained.

### 4.3. Chlorophyll-a algorithms development

In this study, two sources of satellite data, Sentinel-2 MSI and MODIS L2 Chl-a product from late December 2015 to October 2017, are obtained and used for the chlorophyll-a model development. The blue-green ratio algorithm (RI model) and Red-NIR algorithms (MCI and NDCI) originally developed based on the spectral bands of SeaWiFS and MERIS respectively for the retrieval of chlorophyll-a were explored in this study. Remote sensing reflectance retrieved after application of ACOLITE AC processor on Sentinel-2 MSI images were used to derive the chlorophyll-a indices for each Sentinel-2 image downloaded. In using the algorithms, small adaptations were made in the specific spectral bands of Sentinel-2 due to changes in band settings. The datasets used in this study encompassed 12 image pairs with 36 randomly selected data points from each image resulting in a total of 432 matchups of MODIS L2 Chlorophyll-a (Chl-a) values and Sentinel-2 MSI derived Chlorophyll-a (Chl-a) indices values. However, the number of matchups of datasets for each model varied depending on the number of flagged pixels in the MODIS L2 product and unqualified datasets in the Sentinel-2 retrieved indices (Salem et al., 2017).

Chlorophyll-a concentration obtained from the pixel locations extracted from MODIS L2 product at the 36 selected data sites for the model development range from  $\sim 0.48$  mg/m<sup>3</sup> to 18.42 mg/m<sup>3</sup> with closet point to shore at about 5-10 m away. And an average chlorophyll-a concentration of 2.11 mg/m<sup>3</sup> was obtained. High level of Chl-a concentration (4.2 - 18.42 mg/m<sup>3</sup>) were observed in summer while low levels ( $\sim 0.48$  – 2.35 mg/m<sup>3</sup>) were observed in winter (January - February) and in spring (March). There were no images obtained between April - August. Model calibration and validation were run for the entire range of Chl-a  $\sim 0.48$  mg/m<sup>3</sup> to 18.42 mg/m<sup>3</sup> and also for a non-bloom water condition when the Chl-a concentration is low ( $< 3$  mg/m<sup>3</sup>) (Mohsen Ghanea et al., 2016).

The mean of the Chl-a indices values was extracted which are centred around each pixel location site for each moving window size in ESA Snap Software (version 5.0.0). Thirty-six (36) randomly selected data points covering the area of study were used of the models. Coincident MODIS data were also extracted at each data point. For statistical accuracy and reasonable results, extracted Sentinel-2 indices values within each moving window size explored, that did not meet a minimum threshold set at 60 % of the total average window size values extracted were excluded from the data for model calibration and validation analysis. Also, areas affected with clouds were also excluded from analysis data.

For NASA MODIS L2 ocean colour processed data, the following data flags made a pixel invalid: ATMFAIL, LAND, STRAYLIGHT, HISATZEN, HIGLINT, HILT, CLDICE, CHLFAIL, SEAICE, NAVFAIL and ATMWARN (Kahru, Kudela, Anderson, Manzano-Sarabia, & Mitchell, 2014; [https://oceancolor.gsfc.nasa.gov/docs/format/l2oc\\_modis/#specs\\_42](https://oceancolor.gsfc.nasa.gov/docs/format/l2oc_modis/#specs_42)). Hence these flags were used to remove MODIS grids of contaminated data sets. After applying the above filters to the MODIS data, almost 99 % of chlorophyll-a products values  $> 3 \text{ mgm}^{-3}$  were filtered. However, during periods of algal bloom, concentration levels in the AG exceeds this  $10 \text{ mgm}^{-3}$  with some areas reaching as high as  $> 30 \text{ mgm}^{-3}$ . Therefore, in applying the MODIS L2 flags, chlorophyll-a products values  $> 3 \text{ mgm}^{-3}$  were excluded from flag filtering after consistency check with nflh, kd and  $\text{AOT}_{865\text{nm}}$  in order to include some data in this range in the model development. For example, most Chl-a values extracted from MODIS L2 in summer periods (September) were between  $\sim > 3 \text{ mgm}^{-3} - 18 \text{ mgm}^{-3}$ . Hence, pixel values that were considered as homogeneous (based on the above criteria) for each moving window explored for extracting data were used for the calibration and validation of each model. In all, a total of 50 - 136 coincident matchups data points between MODIS data and Sentinel-2 derived chl indices were obtained from the combined datasets depending on the average moving window size after removing flagged pixels.

#### 4.3.1. Maximum Chlorophyll index (MCI)

The Maximum Chlorophyll-a index (MCI) was the first MERIS based Chl-a algorithm developed by (Gower et al., 2008) to model chlorophyll-a concentrations. The MCI algorithm is useful for estimating and the detection of phytoplankton biomass, cyanobacterial blooms or biomass, and varied concentrations of chlorophyll-a. The MCI estimates concentration of Chl-a utilizing bands that are centred at the NIR and red-edge regions (681 nm, 708 nm, and 753 nm band) and the most prominent characteristic of the model is the use of 708/709 nm band that is centred over the Chl-a fluorescence. This is because at these wavelengths, the combined effects of TSS and CDOM are minimized (Kamerosky et al., 2015). But, this peak height is adjustable for other sensors and can be redefined by the researcher. However, according to Hu et 2009, high MCI values may be attributed to surface floating algae or may be suspended matter in the column of the water or even might be caused by sediment plumes. The MCI model adopted for use in this study with the specific wavelength bands of Sentinel-2/MSI after atmospheric correction is of the form;

$$\text{MCI} = R_{\text{rs}}(\lambda_2) - R_{\text{rs}}(\lambda_1) \left[ \frac{(708-665)}{(753-665)} * (R_{\text{rs}}(753) - R_{\text{rs}}(665)) \right] \quad (3)$$

The  $(\lambda_2 - \lambda_1)/(\lambda_3 - \lambda_1)$  is a factor which is approximately equal to 0.389 and is the ratio between the wavelengths determining peak (709-681) and baseline (753-681). It has been shown by Binding, Greenberg, & Bukata, (2013), that the MCI model is effective for estimating and monitoring severe algal blooms of concentration of Chl-a in the range of  $10\text{-}300 \text{ mgm}^{-3}$ . It important therefore to critically consider the reflectance spectra of the algal bloom type during an application of MCI model to avoid false alarm of any algal bloom event. This model is available as a plug-in tool in ESA snap (5.0.0) software used in this study.

#### 4.3.2. Normalised Difference Chlorophyll-a Index (NDCI)

The Normalised Difference Chlorophyll-a index (NDCI) model was developed by (Mishra & Mishra, 2012) for Chl-a estimation in optically complex waters to improve the existing two-band ratio algorithm based on MERIS bands. According to Zhang et al., (2015), these two algorithms both assume that: 1) absorption at red wavelengths is dominated by phytoplankton absorption; 2) absorption at NIR wavelengths is dominated by water absorption; and 3) backscattering is independent of  $\lambda$  between  $\lambda_1$  and  $\lambda_2$ . Further, the NDCI was developed following the simplistic structure of normalized difference vegetation index (NDVI) which is used in monitoring the health of vegetation. NDCI was developed by taking the spectral band difference at 708 nm and 665 nm and normalizing by the sum of their reflectance to eliminate any uncertainties in the estimation of  $R_{rs}$ , seasonal solar azimuth differences, and atmospheric contributions at those wavelengths (Mishra & Mishra, 2012). The combined effects of CDOM and TSM are minimized by closely selecting band 708 nm and 665 nm as well as considering the reflectance and the absorption peak close to 700 nm and 675 nm respectively (Pei et al., 2015). Mishra et al., (2014) established that, the NDCI model could be used to analyse the spatiotemporal variability of phytoplankton distribution accurately over a wide range of water bodies starting from low Chl-a ( $<7.5 \mu\text{g/L}$ ) open ocean to moderate Chl-a ( $7.5\text{--}25 \mu\text{g/L}$ ) near-costal to high Chl-a ( $25\text{--}50 \mu\text{g/L}$ ) inland water to extremely high Chl-a ( $>50 \mu\text{g/L}$ ) agricultural ponds such as catfish pond waters. Hence, the NDCI can effectively be used in monitoring Chl-a quantitatively and qualitatively. The structure of NDCI is follows;

$$Chl - a \propto \frac{[R_{rs}(708) - R_{rs}(665)]}{[R_{rs}(708) + R_{rs}(665)]} \quad (4)$$

#### 4.3.3. Red-tide index Chlorophyll-a Model (RI/RCA)

The Red-tide Index (RI) was developed by (Ahn & Shanmugam, 2006) based on SeaWiFS data for the identification of potential HABs in the Korean South Sea. The RI model has also successfully been used to retrieve Chla in East China Sea, the Bohai Sea, and Yellow Sea using satellite measured data. The RI index was developed utilizing three spectral bands in the blue-green region centred at band 443 nm, band 510 nm and band 555 nm. The developed blue-green ratio algorithm yield indices that can be more related to HABs. Hence bloom dominated waters can be separated from non-bloom waters with RI within the range of  $-1$  to  $+1$ . RI values closer to  $-1$  implies the absence of any bloom and closer to  $+1$  means a high possibility of HAB (Shanmugam, Ahn, & Ram, 2008; Chu & Kuo, 2010). Ahn & Shanmugam, (2006) related the RI to band 443 nm to derive the Red Tide Index Chlorophyll Algorithm (RCA). The newly derived RCA algorithm is more sensitive to algal bloom but in-sensitive to non-algal matter and suspended particles and it related the red tide index to in-situ measured or satellite derived chlorophyll-a. Zhao et al., (2016) recently applied the RI algorithm to MODIS imagery to successfully map Chl-a in the South-eastern Arabian Gulf. RI algorithm formulated and used in this study is as follows;

$$RI = \frac{[L_w(510)/L_w(555) - L_w(443)]}{[L_w(510)/L_w(555) + L_w(443)]} \quad (5)$$

Table 3: Structure and bands adopted by the different band ratio algorithms used in the study. Sentinel-2 bands used in this study; where;  $\lambda_1=443$ ,  $\lambda_2=490$ ,  $\lambda_3=560$ ,  $\lambda_4=665$ ,  $\lambda_5=705$ , and  $\lambda_6=740$  with MCI =Maximum Chlorophyll index, NDCI=Normalized Difference Chlorophyll index and RI= Red Tide index

Model	Equations	Sensor	Author
a MCI	$MCI = R_{rs}(\lambda_5) - R_{rs}(\lambda_4) * \frac{(\lambda_5 - \lambda_4)}{(\lambda_6 - \lambda_4)} * (R_{rs}(\lambda_6) - R_{rs}(\lambda_4))$	MERIS	Gower et al., (2008)
b NDCI	$NDCI = \frac{[R_{rs}(\lambda_5) - R_{rs}(\lambda_4)]}{[R_{rs}(\lambda_5) + R_{rs}(\lambda_4)]}$	MERIS	Mishra & Mishra, (2012)
c RI	$RI = \frac{[(R_{rs}(\lambda_2)/R_{rs}(\lambda_3)) - R_{rs}(\lambda_1)]}{[(R_{rs}(\lambda_2)/R_{rs}(\lambda_3)) + R_{rs}(\lambda_2)]}$	SEAWIFS	Ahn and Shanmugam (2006)

#### 4.3.4. Aggregation of Sentinel-2 Chl models

The sensitivity of the average moving window size of Sentinel-2 derived Chl-a models for extracting MODIS L2 1km chlorophyll-a data is not found in literature. That is whether setting the average window size very close to or at the same spatial resolution of the coarser satellite image or using a small window size is required for the aggregation and subsequent downscaling of the coarser Chl-a data. It was therefore important to determine the critical aggregation level needed for downscaling the coarser MODIS L2 Chl-a product with the high spatial Sentinel-2 Chl-a derived indices in order to obtain reliable results of low uncertainties in Chl-a retrieval. A simple spatial averaging method is employed in this study for the downscaling of MODIS Level-2 chl data with Sentinel-2 derived Chl-a indices. Six (6) spatial-averaging-moving-window-size (3 x 3 windows for the 10 m Sentinel-2 resolution pixels at (30 x 30 m<sup>2</sup>), 5 x 5 moving window (50 x 50 m<sup>2</sup>), 7 x 7 moving window (70 x 70 m<sup>2</sup>), 9 x 9 moving window (90 x 90 m<sup>2</sup>), 11 x 11 moving window (110 x 110 m<sup>2</sup>) and 99 x 99 moving window (990 x 990 m<sup>2</sup>) were explored in this research to obtain the optimal moving window size (or critical aggregation level) for obtaining relationship between MODIS L2 Chl-a and Sentinel-2 derived-Chl-a- indices. Hence, the chlorophyll-a retrieval models were calibrated and validated for each extracted MODIS data and their corresponding model indices values under each moving window size explored to obtain the critical aggregation level.

#### 4.4. Calibration of chlorophyll-a retrieval models

In other to achieve the objectives of this research, it was important to calibrate and validate the Chl-a retrieval models used in this study. This is a requirement for generating reliable and accurate results where regionally calibrated and accepted Chl-a retrieval models are unavailable. The atmospheric corrected remote sensing reflectance (Rrs) values of all Sentinel-2 images obtained for the period of 2015 – 2017 (See table 2)

were used to derive the Chl-a indices values for each model (Table 3) by adopting the specific wavelength bands. In all, a total of 50 - 136 coincident matchups data points between MODIS data and Sentinel-2 derived chl indices were obtained from the combined datasets depending on the average moving window size. The calibration and validation method employed in this study was based on the GeoCalVal model developed by (Salama et al., 2012).

Hence to generate calibration coefficients needed for further estimation of the chlorophyll-a, the data sets obtained after data quality checks in this study were numbered from 1 to n based on the MODIS L2 Chl-a product by first arranging the datasets specifically the independent variable (MODIS L2 Chl-a-data) from smallest to highest together with their corresponding Chl-a retrieved indices values at each data point. The data was then divided into two data sets for calibration and validation. Those with odd numbers were used for the calibration and the other with even numbers for validation. From the calibration datasets, extracted under each window size, the RI, MCI and NDCI indices values were related to their corresponding MODIS Chl-a values. Both linear and nonlinear regression (exponential, power and polynomial) models were tested by fitting indices values derived from Chl-a models' equation to their corresponding MODIS Chl-a values in each average moving window datasets.

Prior to calibrating the RI model, a new red tide index (RI) here called  $RI_{II}$  was estimated by establishing a relation between the RI and  $Rrs$  at 443 nm. According to (Ahn & Shanmugam, 2006), the newly developed RI model provides Chl-a indices values that are more responsive to algal bloom variation and is also more resistant to other environmental factors caused by lights and atmospheric conditions which may affect the model performance. The results of initial statistical analysis revealed a polynomial goodness of fit between RI and  $Rrs$  at 443 nm (see equation 6) with a coefficient of determination ( $R^2$ ) of 0.82. The newly developed  $RI_{II}$  values were then use for model calibration and validation.  $RI_{II}$  model showed an exponential fit with the MODIS Chl-a data run with all the average moving window sizes (3 x 3 to 99 x 99). The calibration function of the 9 x 9 average pixel moving window size with the high  $R^2$  is shown as equation (7). The model obtained as equation (7) is called the Red tide index Chlorophyll Algorithm (RCA).

Also, MCI and NDCI showed good relationship with MODIS Chl-a using a second order polynomial fit consistent with results by (Al Shehhi et al., 2017). This is shown as equation (8).

The derived regression equations are of the form;

$$R_{II} = a_0 \times Rrs_{(443)}^2 - a_1 \times Rrs_{443} + a_2 \quad (6)$$

$$Chl - a_M = b_0 \exp^{[b_1 \times RID]} \quad (7)$$

$$Chl - a_M = c_0 \times Model_{Sen2}^2 - c_1 \times Mode_{Sen2} + c_2 \quad (8)$$



Where;  $Chl-a_{(M)}$  = MODIS L2 Chl-a product;  $Model_{(Sen2)}$  = Sentinel-2 derived MCI and NDCI indices and  $a_0, a_1, a_2, b_0, b_1, c_0, c_1, c_2$  are calibration coefficient of model fittings

#### 4.5. Accuracy assessment of models validation results.

The performance of each models (MCI, NDCI and RI) efficiency in the retrieval of chlorophyll-a products run with each window size (3 x 3 to 11 x 11 and 99 x 99) is evaluated with the second (even numbered) independent validation datasets by calculating the following statistical metrics; the root mean squared error (RMSE); and the mean absolute percentage error (MAPE).

The root-mean-squared error (RMSE) is estimated as;

$$RMSE = \sqrt{\frac{(Chla_{MODIS} - Chla_{S2\_model})^2}{n}} \quad (9)$$

where  $MODIS_{chl a}$  = MODIS Level-2 Chl-a data,  $Sen2_{chl a}$  is the Chl-a derived from sentinel-2 by inverting equation of line of  $Sen2_{chl a}$  against  $MODIS_{chl a}$  and n represents the number of observations.

The mean absolute percentage error (MAPE) is given as;

$$MAPE = \frac{100}{n} \times \frac{|MODIS_{chl a} - Sen2_{chl a}|}{MODIS_{chl a}} \quad (10)$$

where  $MODIS_{chl a}$  = MODIS Level-2 Chl-a data,  $Sen2_{chl a}$  is the Chl-a and n represents the number of observations

#### 4.6. Estimation of chlorophyll-a

The downscaling of ocean colour products especially chlorophyll-a at coarser spatial scale to a high spatial resolution requires, a dependent variable (the predictor) and an independent variable (response variables). The dependent variable used here (the Sentinel-2/MSI derived algorithms indices) should specifically be of high correlation with the response variables, in this study, (the MODIS L2 ocean colour Chl-a data). Also, to ensure consistency between downscaled Chl-a product and coarser chlorophyll-a product, it is reported that, the fine resolution downscaled product are upscaled to the coarser resolution and then, the difference (residual) between them is estimated and added to the original fine resolution image (Díaz-Uriarte et al., 2014).

After evaluating the critical aggregation level needed for downscaling the coarser MODIS L2 Chl-a product, none of the models used in this study had a better mean percentage error of variation in the retrieved chlorophyll products. However, the statistical results from RI/RCA model derived with the exponential model under the 9 x 9-pixel moving window size performed better than the other models. It is also more

sensitive to Red Tides which is the main prominent type of HAB present in the AG. Hence, the calibration coefficients derived with the RI/RCA model were used in estimating Chl-a concentration near the intake of the Ras Abu Fontas desalination plant for each month of satellite image downloaded. The equation derived for the downscaling function obtained for application on the Sentinel-2 derived indices for chlorophyll-a retrieval is of the form shown as equation (12).

#### **4.7. Evaluating the effect of spatially aggregation on spatial variation of retrieved Chl-a products**

After applying the calibration coefficients of the RI model run with 9 x 9-pixel window shown as equation (12) to generate fine 10 m resolution chlorophyll-a products, the approaches used to analyse the spatial variability in Chl-a distribution are described. The Chl-a map derived at 10 m for each Sentinel-2 derived Chl-a product are re-aggregated to 90 m and the MODIS L2 is disaggregated to 90 m. Also, Sentinel-2/MSI derived Chl-a are re-aggregated to 1 km (at the same spatial resolution of the MODIS L2 Chl-a product). The difference (residual) between the re-aggregated Sentinel-2 and the MODIS L2 at 90 m for each image were estimated. Average of 3 x 3 pixel was extracted around the selected 36 sampling data points distributed within the area study in each image and used to train the residuals.

Various researchers report adding the mean of the residual to the downscaled product or adding the residuals per pixel basis to ensure consistency (residual correction) between the reference Chl-a product and the downscaled product at fine spatial resolution as in (Bisquert et al., 2002; Kustas, Norman, Anderson, & French, 2003). The assumption is that, the residual introduced is as a result of the spatial variability between the fine and the coarser image, viewing geometry and the differences in the sensitivity of the sensors (Njuki, 2016) and other environmental factors not fully captured in the predictor and response variable relationship. The result obtained from this approach is a chlorophyll product of high spatial resolution obtained through the linear relationship of upscaled 90 m Chl-a indices as prediction variables and MODIS L2 as response variable. However, due to the localized nature of the study and the complexity in modelling chlorophyll-a in the study region because of its shallow and turbid nature, the average of the residuals was not added to the original downscaled product. Because large error may be introduced in the final product and hence affect the capability of the model in predicting Chl-a.

Hence, to assess the consistency in the spatial variability of the derived chlorophyll-a and MODIS observed Chl-a product in the study area at different spatial resolution for cross comparison purpose and the capability of the derived model in Chl-a estimations, scatter plots of Sentinel-2/MSI aggregated to 90 m and 1 km against MODIS L2 Chl-a product disaggregated at 90 m and its 1 km spatial resolution at nadir for images acquired on 22 September 2017 were generated.

#### **4.8. Consistency check in cross-validation of spatial variation of Sentinel-2/MSI Chl-a product**

To further evaluate whether the discrepancy in the retrieved Chl-a products were as a result of the AC method and algorithms used or whether it was caused by uncertainties of MODIS Rrs (Shang et al., 2014).

Two approaches were used to assess the accuracy in the derived Chl-a products. First, we examined the uncertainties based on the atmospheric correction method for the two satellite products and we evaluated the consistency on the atmospheric corrected products of Sentinel-2 and MODIS. The evaluation was performed using band 443, 488, 547 nm of MODIS with apparently close bands 443, 497, and 560 nm of Sentinel-2 MSI. This is because, the derivation of the OC3M and the RI/RCA Chl-a products were based on these bands. Secondly, the 10 m Sentinel-2/MSI retrieved Chl-a products were (i) aggregated to 90 m while the spatial scale of the MODIS L2 was disaggregated to 90 m and (ii) were further aggregated to the 1 km spatial scale of the MODIS L2 ocean color product. The accuracy in retrieved products were evaluated and the residuals and the average of the residuals were estimated for each product using satellite product obtained on 22 September 2017.

#### **4.9. Chlorophyll-a retrieval with Neural Network models**

Estimation of water quality parameters have relied on bio-optical models such as empirical, semi-empirical, analytical, and semi-analytical techniques (Ambarwulan, 2010; Ogashawara, Li, & Moreno-Madrinán, 2016). However, most of these approaches are limited to the representative datasets or parameters with distinctive absorption characteristics (Gholizadeh, Melesse, & Reddi, 2016). Neural network models provides a robust, fast and easy to implement approaches in estimating bio-optical constituents by also taking into consideration the IOPs in highly turbid and eutrophic waters where the conventional bio-optical models fails. These includes Case 2 Regional (C2R), FUB (Free University of Berlin, Eutrophic Lake (EUL) and Coast Color (C2RCC) neural network developed for MERIS. In-situ measurement for robust model calibration and validation is unavailable for use.

Hence, in this study, the performance of Case 2 Regional Coast Color (C2RCC) neural network algorithm compared with results of our downscaling technique for water quality retrieval (Chlorophyll-a retrieval). Case 2 Regional Coast Color (C2RCC) neural network processor recently developed by (Brockmann et al., 2016) by updating Case 2 Regional (C2R) (Doerffer and Schiffler 2007) but with improvement in water quality retrieval as well as Inherent Optical Properties (IOPs) and water leaving reflectance. C2RCC is part of the current ESA SNAP Sentinel toolbox (Version 6.0.0) made available in 15 January 2018. The neural nets model supports Sentinel-2 MSI, Landsat OLI, OLCI, MODIS, MERIS, and SeaWiFS. Output of C2RCC neural nets after processing in SNAP includes (IOPs), water quality parameters including Chl-a concentration and TSM together with atmospheric corrected water leaving reflectance data.

## 5. RESULTS AND DISCUSSION

### 5.1. Model calibration and Validation

Following the atmospheric correction step, the structure of RI/RCA model proposed by (Ahn & Shanmugam, 2006), MCI proposed by (Gower, King, & Goncalves, 2008), and NDCI model proposed by (Mishra & Mishra, 2012) were adopted and applied to the results of the atmospheric corrected reflectance products for chlorophyll-a concentration retrieval. However, some adaptations were made regarding the chl-a indices bands of Sentinel-2. For example; in place of band 709 nm for MERIS MCI, we use band 705 for Sentinel-2 MCI model. Chlorophyll-a concentration obtained from all images the 36 data sites varied from  $\sim 0.48 \text{ mg/m}^3$  to  $18.42 \text{ mg/m}^3$  with average chlorophyll-a concentration of  $2.11 \text{ mg/m}^3$ . It should be noted that, values of Chl-a and indices extracted from all 12 images were used for model calibration and validation divided into two independent datasets. However, after removing flagged pixel and unqualified datasets, the number of matchups of datasets for each model under each average moving window size varied. Regression analysis of linear, exponential, power and polynomial fitting models were all explored in this study for the calibration datasets using least square methods.

The results of initial statistical analysis between RI and Rrs at 443 nm to obtain  $RI_{II}$  revealed a polynomial goodness of fit (See Annex A) with a coefficient of determination ( $R^2$ ) of 0.82 shown as equation 11. Also, the calibration of  $RI_{II}$  using MODIS L2 Chl-a (for all data range) revealed a non-linear fit with an exponential fit as the strongest for datasets with run with all average moving window size. Models run with  $9 \times 9$  was the best for extracting the indices values for model development. For  $RI_{II}$  versus MODIS L2 Chl-a data run with  $9 \times 9$  we obtained an  $R^2$  of 0.507 shown as equation 12. The polynomial fitting model obtained between RI and band 443 nm deriving  $RI_{II}$  and then exponential fitting model between  $RI_{II}$  and MODIS data respectively agrees with findings by Chu & Kuo, (2010) and Ahn & Shanmugam, (2006). However, the  $RI_{II}$  model was found to saturate at Chl-a concentration above  $4 \text{ mg/m}^3$ .

The same approach was also followed for linear, polynomial and exponential model calibration for Sentinel-2 derived MCI and NDCI run for each pixel moving window (from  $3 \times 3$  to  $11 \times 11$  and  $99 \times 99$ ) explored. For MCI model calibration good correlation coefficients were using MODIS L2 Chl-a; For linear model with  $R^2$  of 0.82, polynomial model (2nd order) with  $R^2$  of 0.84 and exponential model  $R^2$  of 0.71 respectively. Also, for relationship between NDCI and MODIS Chl-a,  $R^2$  of 0.67, 0.76 and 0.63 for linear, polynomial and exponential model calibration respectively were estimated. Also, MCI and NDCI models both showed high accuracy with the polynomial relation with the MODIS L2 data run with  $9 \times 9$  average pixel moving window shown in Figure 3b and c.

$$RI_{II} = -10.402x^2 - 1.3872x + 1.0002 \quad (11)$$

Where; x is the remote sensing reflectance at 443 nm.

$$\text{Chl-a}_{(M)} = 8E-56 e [129.69 \times \text{RI}_{(M)}] \quad (12)$$

Where  $e = 2.718281$

$$\text{Chl-a}_{(M)} = 163.06 \times \text{ndci}^2 + 116.67 \times \text{ndci} + 22.166 \quad (13)$$

$$\text{Chl-a}_{(M)} = 1E+06 \times \text{mci}^2 + 823.22 \times \text{mci} - 0.0705 \quad (14)$$

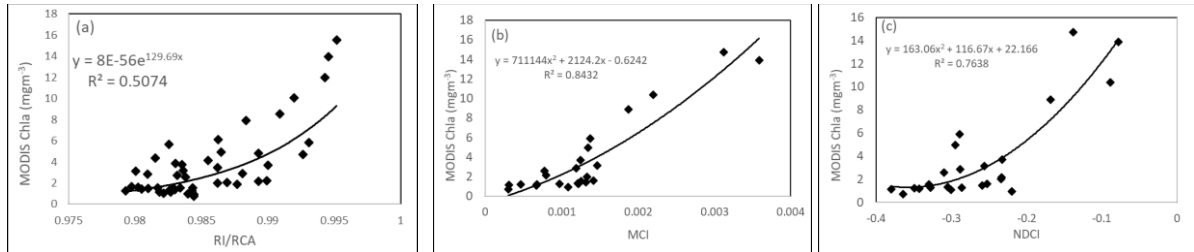


Figure 3: Chlorophyll-a concentration retrieval algorithms based on Sentinel-2/MSI derived indices RI/RCA (a), (b) MCI and NDCI (c) run with 9 x 9 moving window size.

## 5.2. Model validation results

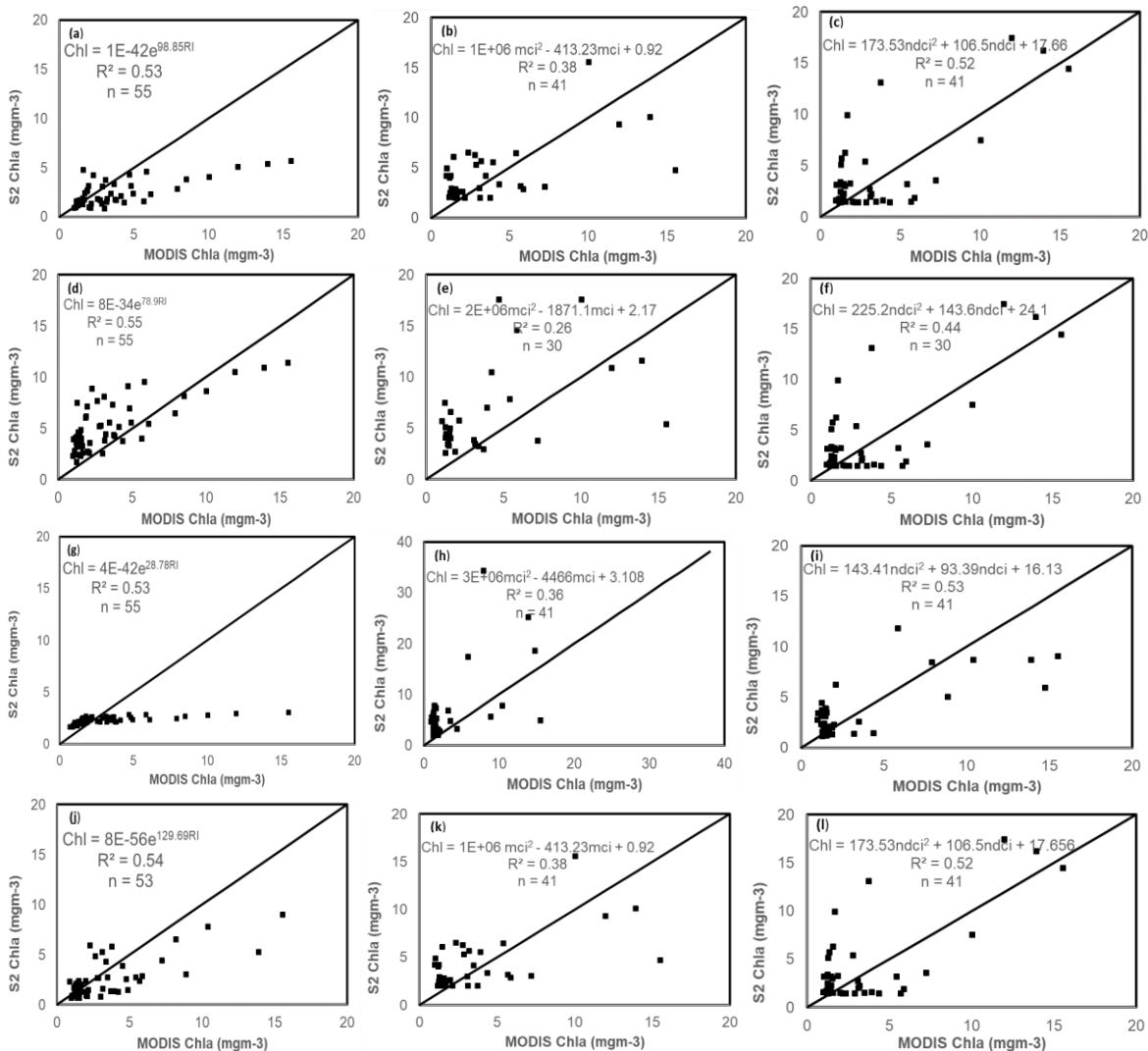
To validate the chlorophyll model performance, the second group of independent validation datasets were used. The validation datasets varied for each model and for each average pixel moving window size explored depending on the number of flagged data. For RI the validation datasets range between 53 to 55. For MCI and NDCI the validation datasets range between 30 to 51 data. Figure 4 shows the validation results of all models run with the different average spatial window size.

For RI/RCA, the model equation shown as Equation 12 derived with the exponential calibration model was applied on the validation dataset for retrieval of chlorophyll-a concentration. Also, results of linear and polynomial calibration coefficients were also applied to the validation dataset for chlorophyll-a retrieval. The validation result of RMSE and MAPE confirmed the exponential fit between Sentinel-2 derived indices and MODIS Chl-a as the best calibration model for chlorophyll retrieval run with all moving window sizes (3 x 3 to 99 x 99). The performance of the RI/RCA model for the validated datasets of all the moving average spatial aggregate level (3 x 3 to 99 x 99) showed similar sensitivity, with RMSE values ranging between 2.75 mg/m<sup>3</sup> to 2.35 mg/m<sup>3</sup>. Also, a mirrored average absolute percentage error (MAPE) values were also obtained ranging (from 104 % to 38.4 %) as shown in Table 4.

The coefficients of the polynomial together with linear and exponential calibration models of MCI and NDCI were also applied to the validation dataset for chlorophyll-a retrieval. From statistical accuracy estimation (shown in Table 4), results of RMSE for the MCI model for the polynomial model ranges between 9.1 to 2.47 mg/m<sup>3</sup> and for MAPE 197 % to 52.5 %. Similar results were also obtained for the NDCI model with RMSE ranging between 4.81 to 2.42 mg/m<sup>3</sup> and MAPE ranging between 150 % to 73.3 %. The accuracy results of all models run for the whole Chl-a range extracted exceeded the acceptable level

of error ( $\pm 35\%$ ) in satellite derived Chl-a by NASA. This is because, RI/RCA model underestimates Chl-a compared with MODIS retrieved Chl-a results. However, MCI and NDCI overestimated Chl-a concentration at low concentration ( $\sim < 3 \text{ mg/m}^3$ ). Summary statistical accuracy estimation results based on RMSE and MAPE of each model under different moving window size are displayed in Table 4.

In contrast to results of RMSE AND MAPE obtained for RI, MCI and NDCI model for the Chl-a for the whole chlorophyll-a range ( $\sim 0.48 \text{ mg/m}^3$  to  $18.42 \text{ mg/m}^3$ ), characterizing the Chl-a model performance for non-bloom conditions (oligotrophic water) of the study area, when  $\text{Chl-a} < 3 \text{ mg/m}^3$ , the errors in Chl-a concentration decreased considerably reaching as low as  $0.385 \text{ mg/m}^3$  for RMSE and  $22\%$  for MAPE for the RI/RCA model. Results of RMSE and MAPE for MCI model ranges between  $0.909$  to  $0.804 \text{ mg/m}^3$  and  $53.79\%$  to  $47.68\%$  run for each average moving window size. The validation results of the Sentinel-2 derived Chl-a concentration by RI/RCA, MCI and NDCI model for each dataset under the spatial moving windows when  $\text{Chl-a} < 3 \text{ mg/m}^3$  are shown in Annex C.



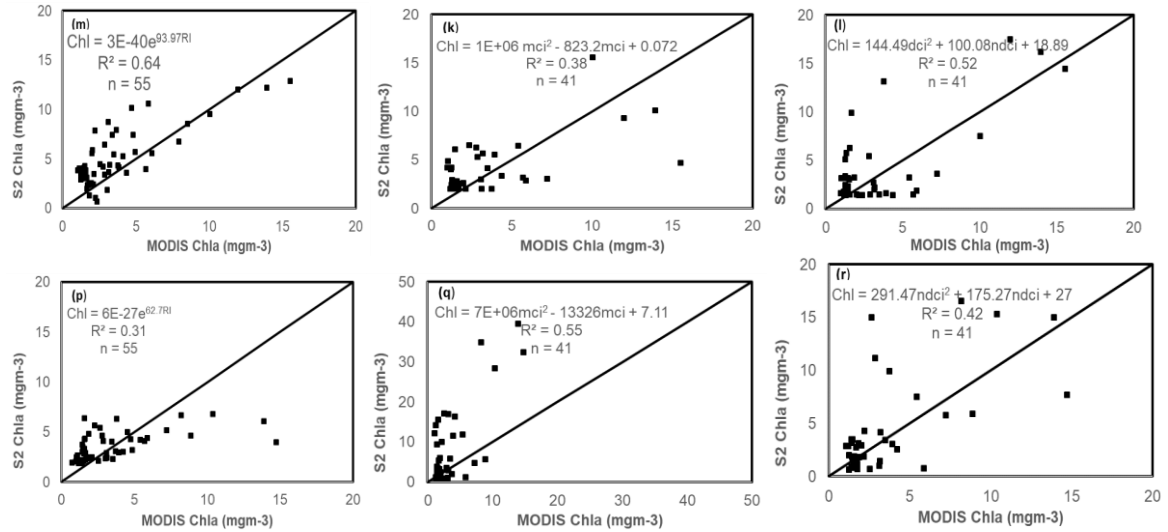


Figure 4: scatterplots of validation set of Sentinel-2 derived chlorophyll-a by RI, MCI and NDCI against MODIS L2 chlorophyll-a data: a – p (col 1) by RI model; b – q (col 2) by MCI model and c - r (col 3) by NDCI model; a, b, c for 3 x 3; d, e, f for 5 x 5; g, h, i for 7 x 7; j, k, l for 9 x 9; m, n, o for 11 x 11 and p, q, r for 99 x 99 average moving window size

Table 4 : Statistical summary of scatter plots of Sentinel-2 derived chlorophyll-a concentration against MODIS L2 chlorophyll-a measurement for the validation datasets of each average spatial window size.

Window size	Model								
	NDCI			MCI			RI		
	R <sup>2</sup>	RMSE (mg/m <sup>3</sup> )	MAPE (%)	R <sup>2</sup>	RMSE (mg/m <sup>3</sup> )	MAPE (%)	R <sup>2</sup>	RMSE (mg/m <sup>3</sup> )	MAPE (%)
3 x 3	0.52	2.85	91.5	0.38	2.96	89.4	0.53	2.70	<b>38.4</b>
5 x 5	0.44	4.81	150.0	0.26	5.26	151	0.55	2.74	104.0
7 x 7	0.53	2.67	78.8	0.36	3.26	54.1	0.53	2.77	47.5
9 x 9	0.52	2.41	74.3	0.38	2.47	52.9	<b>0.54</b>	<b>2.75</b>	<b>44.2</b>
11 x 11	0.58	2.38	95.7	0.38	2.88	71.3	0.64	2.35	84.5
99 x 99	0.42	3.45	73.8	0.55	9.1	197	0.31	2.50	69.0

### 5.2.1. Selecting the critical aggregation level

The results in Table 4 show that, RI model is more sensitive to chlorophyll-a in the Arabian gulf although the MCI model also showed some level of accuracy in retrieved Chl-a. An increase in the average spatial window size to 11 x 11 and 99 x 99 (almost at same spatial scale of MODIS L2 product) did not result in any better improvement of the performance of the model. Therefore, the 90 m x 90 m average pixel moving window size was selected as the critical aggregation level (optimal window size) for the extraction of coincident Chl-a index values important for the estimation of chlorophyll-a concentration. For easier comparison of the RI model performance (RMSE and MAPE) for each moving window size are displayed in Figure 5;

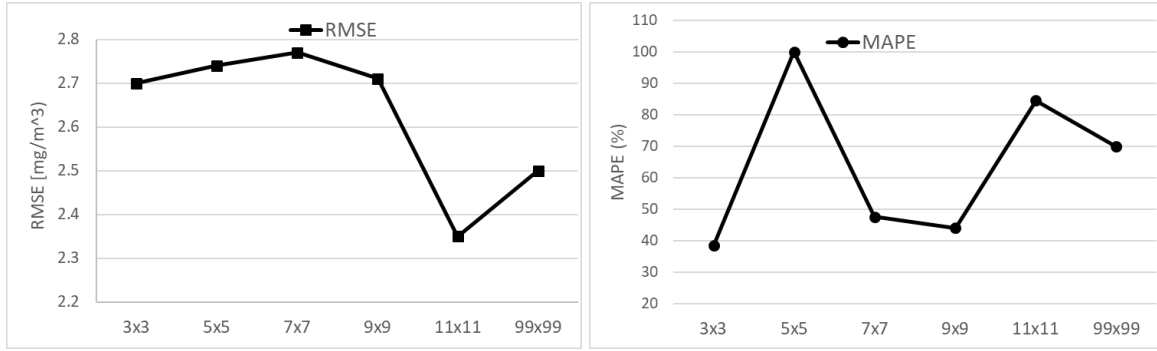


Figure 5: Statistical summary of scatter plots of Sentinel-2 derived chlorophyll-a concentration against MODIS L2 chlorophyll-a measurement for the validation datasets of each average spatial window size for RI model.

### 5.2.2. Comparison of accuracy results with standards

In characterizing the Chl-a model performance with the validation datasets over the whole chlorophyll-a range  $\sim 0.48 \text{ mg/m}^3$  to  $18.42 \text{ mg/m}^3$ , the mean absolute percentage error (MAPE) in Chl-a estimates of all models (RI, MCI and NDCI) of each dataset in each aggregate level explored ( $3 \times 3$  to  $99 \times 99$ ) exceeded the level of accuracy  $\pm 35 \%$  in satellite retrieval of chlorophyll-a values (IOCCG, 2006; Moore, Campbell, & Dowell, 2009) for the shallow region of the Arabian Gulf. The acceptable accuracy  $< \pm 35 \%$  is usually plausible for Chl-a retrieval in the middle of open oceans gyres where contribution of non-algal particles and bottom reflectance is lower (Moore et al., 2009). However, it can be  $> 50 \%$  in ocean gyres and even exceed  $> 100 \%$  for complex coastal waters (Blondeau-Patissier, Gower, Dekker, Phinn, & Brando, 2014), which the Arabian Gulf is not an exception. Although our results exceeded  $\pm 35 \%$  satellite Chl-a measurements, it might still be relevant or acceptable compared with other findings (e.g. D’Ortenzio, et al., (2002), for research conducted over the Mediterranean Sea, and Gregg & Casey, (2004), for research conducted over the Antarctic or the Equatorial Atlantic).

In all, RI/RCA run with  $9 \times 9$  model average moving window size produced the lowest error over the whole chlorophyll-a range although it underestimates Chl-a at high concentration compared with MODIS observed results. The MCI and NDCI however showed overestimation and could not retrieve chlorophyll-a values  $< 1 \text{ mg/m}^3$ . However, results of MCI were better than NDCI. The RI/RCA model was selected for Chl-a retrieval and hence was applied on all images downloaded for use in this study.

### 5.3. Chlorophyll-a retrieval with Neural Network models

The accuracy of RI/RCA retrieved chl-a product from Sentinel-2 was tested by comparing with results of Chl-a concentration derived with C2RCC neural nets as shown in Figure 6. MODIS OC3M derived Chl-a products were also compared with with C2RCC neural nets derived Chl-a. Five chl-a products covering low to high chl-a concentration period were used for the validation test. These include; 22 October 2016, 11 November 2016, 11 December 2016, 31 December 2016, and 27 September 2017. Accuracy of retrieved Chl-a with C2RCC compared with our approach and MODIS OC3M were poor. Our approach of Chl-a



retrieval showed the lowest accuracy with  $R^2$ , RMSE and MAPE of 0.009, 2.54 mg/m<sup>3</sup> and 73.9 % respectively (Figure 6b). For MODIS,  $R^2$ , RMSE and MAPE of 0.0217, 2.34 mg/m<sup>3</sup> and 54.8 % respectively were estimated.

However, there was a an improvement when areas located near-shore were removed thereby generating 0.52, 1.75 mg/m<sup>3</sup> and 53.8 % for  $R^2$ , RMSE and MAPE respectively for RI/RCA model. And for MODIS  $R^2$ , RMSE and MAPE of 0.46, 3.18 mg/m<sup>3</sup> and 55.7 % respectively were estimated. However, C2RCC significantly underestimated chl-a concentration compared with the other two models at  $\sim > 4$  mg/m<sup>3</sup>. It showed more constintency with MODIS chl-a for non-bloom water condition ( $< 3$  mg/m<sup>3</sup>). It was also more correlated with chl-a retrieved product on 22 Septemeber 2017 with RI/RCA model when algal bloom was suspected to have occurred with chl-a concentration exceeding 10 mg/m<sup>3</sup>.

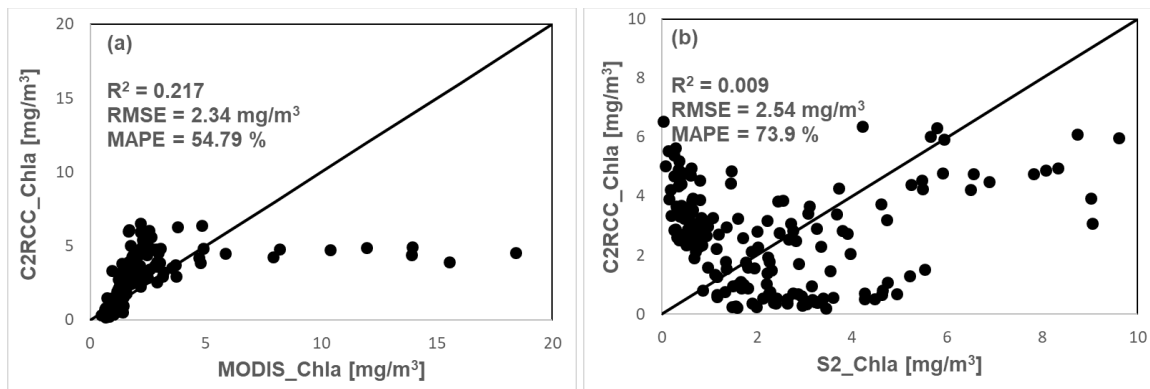


Figure 6: Comparison of Chl-a products derived with C2RCC and (a) MODIS L2; (b) Sentinel-2 RI/RCA Chl-a concentration

#### 5.4. Spatial variation of retrieved chlorophyll-a (Chl-a) product vs MODIS L2 Chl-a product

Figure 7 and 8 show the spatial distribution pattern of chlorophyll-a(Chl-a) from late December 2015 to late October 2017 subset to the area near the intake of RAS A2 and A3 desalination plants. Figure 7 is the results of applying the best performing model, RI/RCA algorithm (shown in Equation (12)), on the Sentinel-2 images. While Figure 8 is the result of MODIS L2 standard ocean color Chl-a products (freely available). From Figure 7 and 8, there was some level of agreement between the MODIS and Sentinel-2 RCA derived Chl-a on-shore. However, disparity in estimated Chl-a concentration was observed at near-shore in all images. There was an overestimation of Chl-a near-shore for the MODIS Aqua OC3M Chl-a products. This because, at nearshore, where the water depth is low, the contribution of CDOM and increase sediment concentration are predominant over the Arabian Gulf (Al Shehhi et al., 2017). At these areas, retrieved Chl-a from MODIS exceeded 3 mg/m<sup>3</sup> with some areas recording as high 10 mg/m<sup>3</sup> in low tide periods. But, Chl-a results from Sentinel-2 showed Chl-a concentrations lower than 3 mg/m<sup>3</sup> with some areas reaching as low as  $< 0.1$  mg/m<sup>3</sup>. However, Chl-a results from MODIS and Sentinel-2 where were in agreement for Chl-a image obtained on 22 September 2017. At this period there was an elevated level of observed Chl-a from both data. There is also a rough trend in the spatial distribution of all MODIS L2 Chl-a products near-

shore of the study area and therefore MODIS OC3M does not fully capture the spatial variability in Chl-a. But from the available derived maps, clearly, the Sentinel-2 derived Chl-a product show into more fine details, the spatial variation in Chl-a concentration allowing for easy identification of areas with high to no patches of algal bloom concentration in these maps.

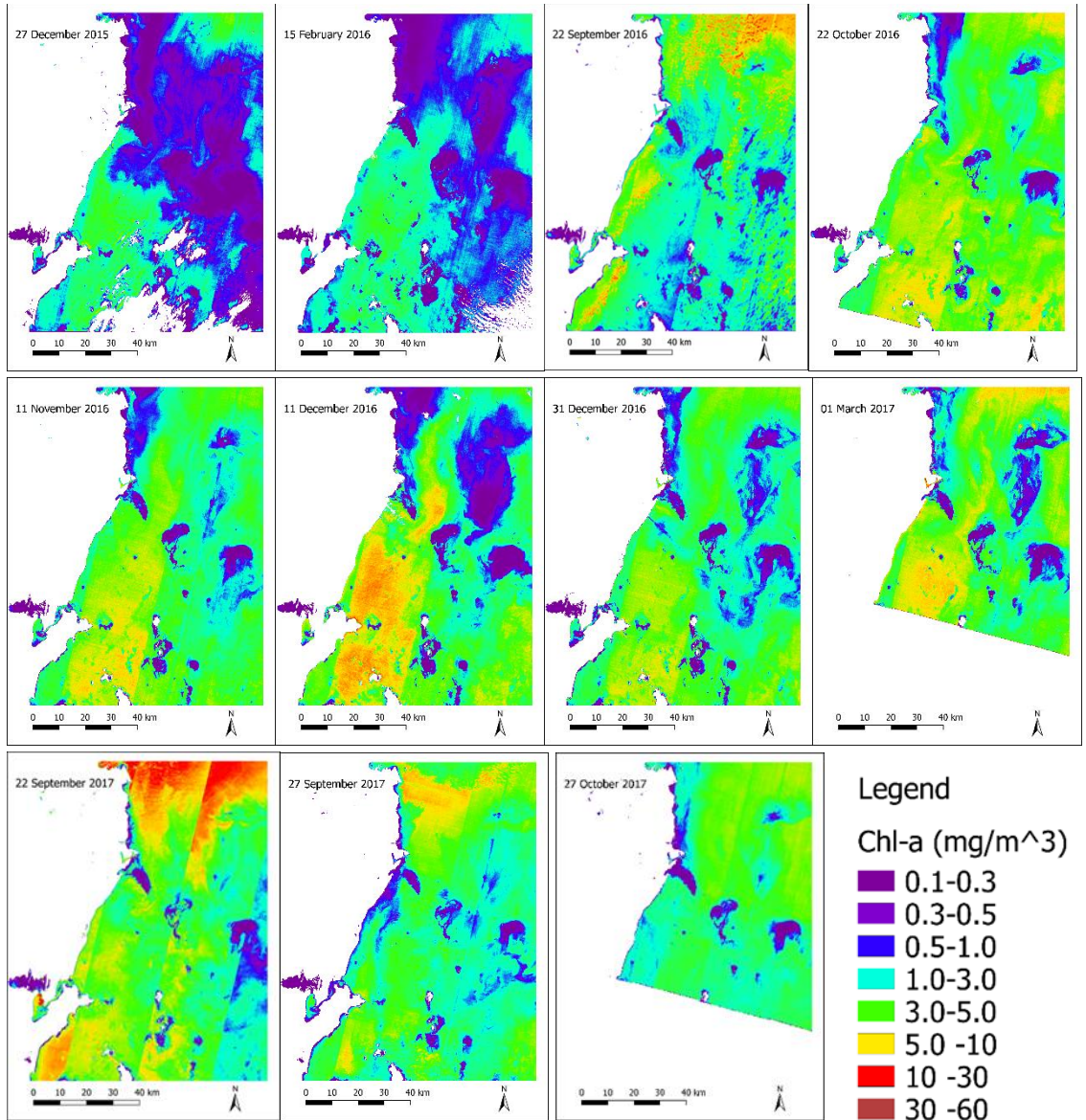


Figure 7: Chlorophyll-a concentration (Chl-a) derived from Sentinel-2 RI/RCA model for 11 scenes from 27 December 2015 to 27 October 2017. All images were processed with ESA Snap software (5.00)

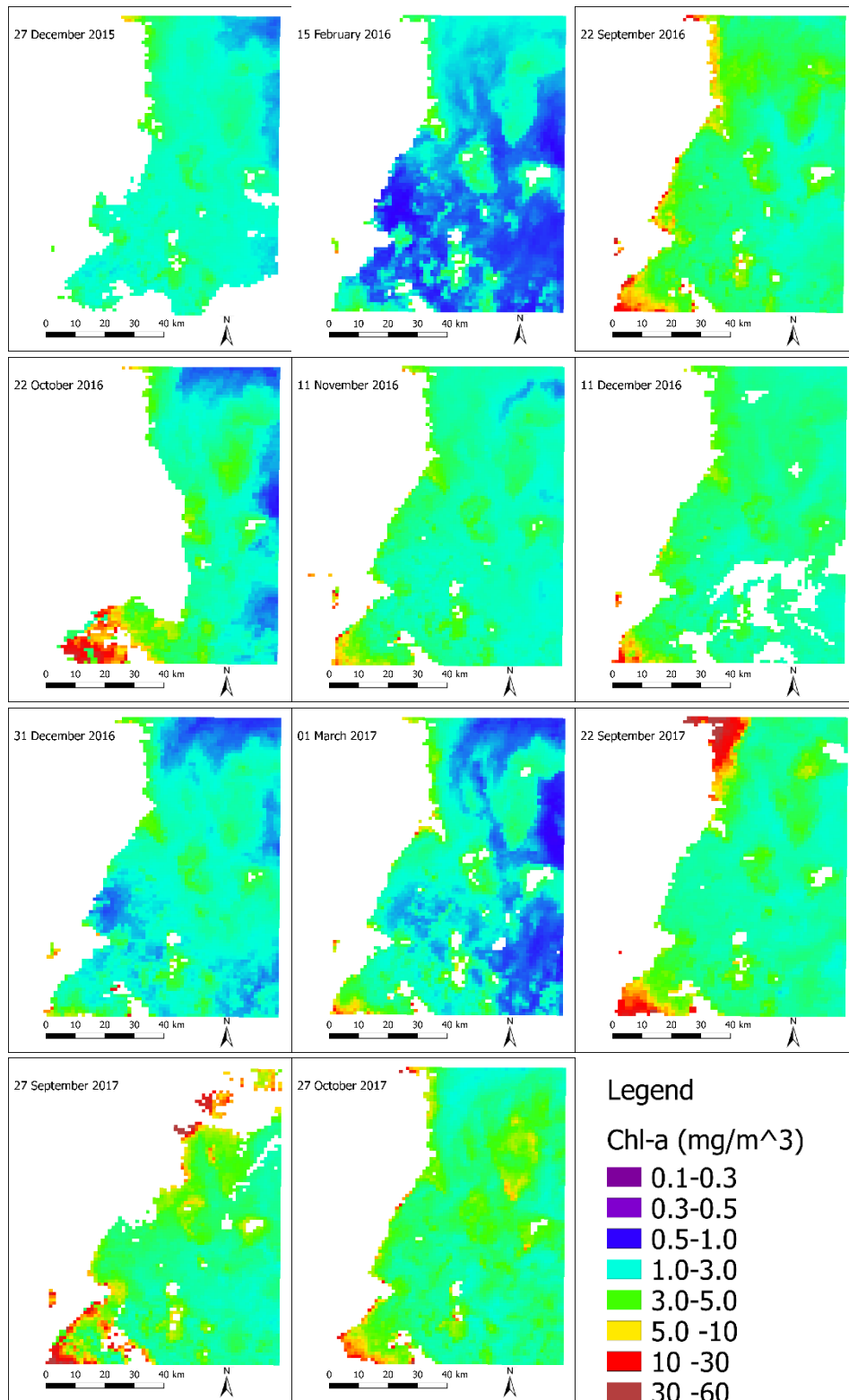


Figure 8: Seasonal trend of MODIS Aqua chlorophyll-a image derived with OC3M algorithm for 11 scenes from 27 December 2015 to 27 October 2017.

### 5.5. MODIS and Sentinel-2 Chlorophyll-a maps as indicators for algal bloom mapping

Chlorophyll-a products derived from Satellite images can be used as indicators of the occurrence and presence of algal blooms in coastal waters. Figure 7 and 8 shows the spatial distribution of chlorophyll-a over the study area. The development and spread of the patches of algal blooms in the study area can clearly be traced from Sentinel-2 maps shown in Figure 7. From Figure 7, the bloom patches are more visible for Chl-a maps processed on 11 December 2016 and 22 September 2017. The observed Chl-a concentration exceeded  $3 \text{ mg/m}^3$ , indicating the presence of algal bloom (Ghanea et al., 2016). However, this is not apparent in the MODIS observed Chl-a products shown in Figure 7. Hence, the Sentinel-2 derived Chl-a product resolve into more fine detail the variation in observed Chl-a concentration. Hence, this gives an indication that, the RCA model developed through the downscaling technique as being more sensitive to Chl-a than OC3M algorithm.

On, 22 September 2017, the chlorophyll concentration derived from MODIS L2 ocean color product showed a much more elevated level of observed Chl-a concentration with the minimum concentration of  $\text{Chl-a} > 1 \text{ mg/m}^3$  and maximum Chl-a concentration exceeding  $10 \text{ mg/m}^3$ . The high level of retrieved Chl-a was observed at the northern part of the study area. The Sentinel-2 derived RCA chlorophyll-a also showed a similar trend in the chlorophyll-a concentration during this period. Hence, confirming MODIS results of probable algal bloom occurrence during this period. The high levels in chlorophyll-a concentration may be due to a prolific growth in algae and increased nutrients concentration. We do not have extensive coincident/simultaneous field measurements of Chl-a data to validate the result if whether indeed there was truly an algal bloom event.

However, report received from management at the project site indicates that, during the early period between September 2017 to middle of October 2017, desalination operations at the RAS Abu Fontas A3 (SWRO) desalination plant was shut-down (See snapshot of report in Appendix D). Similar incidence occurred within this period (August – December) in 2008 - 2009 which forced several desalination plants to partially halt operation or completely shut-down due to an increased red tides together with its associated high algal biomass resulting in increased turbidity near the plant intakes (Al Shehhi, Gherboudj, & Ghedira, 2017; Darwish et al., 2016). This is because primary productivity in the AG reaches its peak in summer and sometimes in spring season when the sun energy is more available (Al-Naimi et al., 2017). Although the OC3M algorithm is designed for detection of algal blooms but non-algal materials and bottom reflectance affect its ability in retrieving Chl-a from the shallow-basin of the Arabian Gulf resulting in overestimation of observed Chl-a (Al Shehhi, Gherboudj, Ghedira, et al., 2017).

Overall, it is not easy to accurately distinguish areas with patches of blooms from those of low or no patches of algal bloom during the winter and spring periods (when chlorophyll concentration is lower than  $3 \text{ mg/m}^3$  and especially at offshore) for MODIS observed Chl-a products. This could be caused by the default parameterization of the OC3M algorithm for the MODIS sensor (Al Shehhi et al., 2017; Jun Zhao et al.,

2016). It might also be due to the large spatial scale of MODIS incapable of resolving into detail small changes in Chl-a concentration. However, the RI/RCA model proved to capture small changes in the diverse community of small algal bloom patches even though it underestimates at high concentration compared to MODIS observed Chl-a products.

### 5.6. Chlorophyll-a retrieved with C2RCC Neural Net algorithm

Figure 9 shows the results of chlorophyll-a concentration retrieved with the C2RCC Neural Net in ESA SNAP software (Version 6.0,0). The Chl-a image processed with the C2RCC processor on the 22 September 2017 also showed similar trend in obtained results to that from MODIS OC3M and Sentinel-2 RI/RCA algorithms. However, it showed an overestimation near-shore as compared to RI/RCA. Hence, the effects of CDOM and TSM were also predominant in the retrieved Chl-a products. Moreover, the Chl-a retrieved through the downscaling technique were more spatially homogeneous and more sensitive to Chl-a in the area of study. Therefore, the spectral features introduced by MODIS onto Sentinel-2 for the retrieval Chl-a is important for accurate detection and mapping of Chl-a due to limitations in the spectral set-up of Sentinel for Chl-a retrieval.

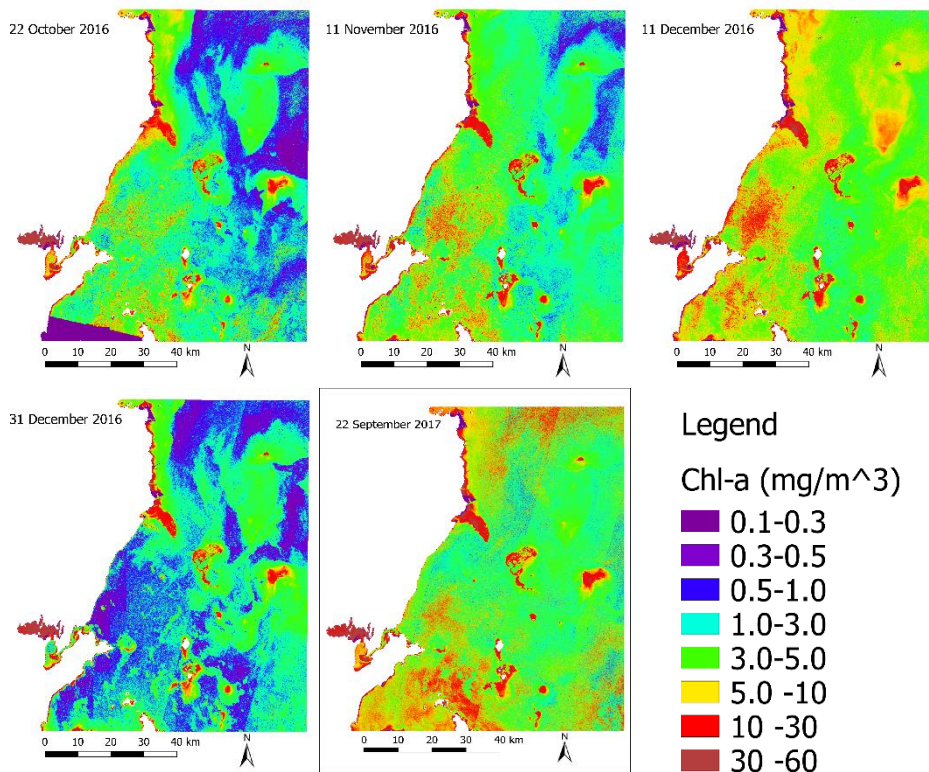


Figure 9: Sentinel-2 MSI Chl-a images derived with C2RCC Neural Network algorithm.

#### 5.6.1. Time series of derived Chl-a concentration

Figure 10 shows time series of monthly-averaged seasonal trends in Chl-a distribution over the study areas between 2015 – 2017 estimated from maps shown in Figure 7 and 8. Although it was difficult to accurately distinguish areas with small to large patches of algal bloom from non-bloom waters from the MODIS



derived Chl-a data, the monthly averaged Chl-a results over the study area obtained in Figure 9 however shows a similar trend with the spatial distribution of Chl-a in the study area.

Minimum concentration levels of Chl-a ( $1.04 - 1.53 \text{ mg/m}^3$ ) and ( $0.45 - 0.73 \text{ mg/m}^3$ ) were observed between late December 2015 and middle of February 2016 for MODIS and Sentinel-2/MSI respectively. High concentration levels of observed Chl-a reached  $9.63 \text{ mg/m}^3$  and  $7.53 \text{ mg/m}^3$  in September 2017 for MODIS and Sentinel-2/MSI respectively. The concentration then increased with maximum ( $3.50 \text{ mg/m}^3$ ) in early winter (December 2016) to a minimum ( $1.42 \text{ mg/m}^3$ ) in March 2017 (spring). High Chl-a in the winter was probably related to the decreases stratification of the water column (Jun Zhao et al., 2016). Hence, high levels of Chl-a were found in summer and minimum concentration levels in late winter.

Our result of the temporal trend in Chl-a is similar to finding by (Nezlin, Polikarpov, Al-Yamani, Subba Rao, & Ignatov, 2010). They found that, in the shallow western and southern regions of the AG, minimum chlorophyll-a concentration was observed in spring and maximum in late summer and Autumn. Nezlin et al., (2010), for their research conducted over the AG observed that, these maxima coincided with the seasonal minimum of wind mixing over the Gulf suggesting that in shallow waters, turbidity decreased and bottom reflectance increased. Here the contribution of bottom reflectance may be pronounced in the Chl-a estimates than turbidity. Hence, algorithms designed for retrieving ocean color products should be able to overcome bottom reflectance contribution to water leaving reflectance especially the blue band (Cannizzaro & Carder, 2006). However, this temporal trend in observed Chl-a for the study area deviates from previous findings for research conducted over the whole Arabian Gulf. The found high levels of Chl-a in winter and minimum concentration levels in summer. Mezhoud et al., (2016) reported that, the spatiotemporal variability of Chl-a together with SST, salinity, DO, pH, SDD are particularly affected and driven by the climatic condition of the AG.

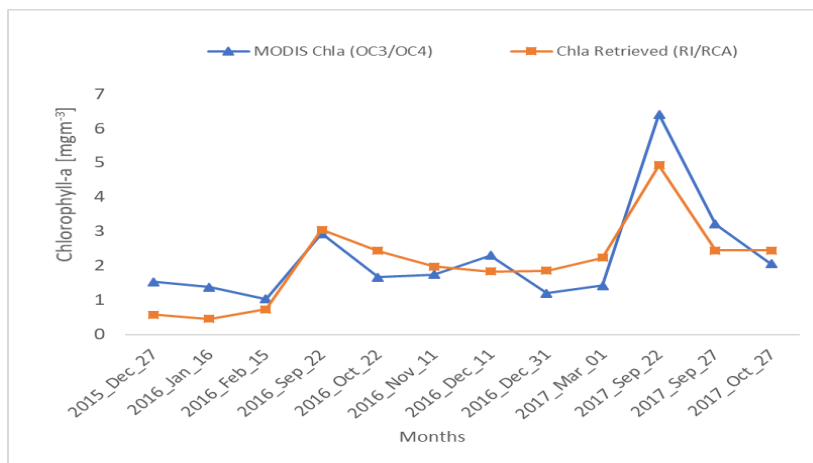


Figure 10: Time series of NASA MODIS satellite OC3/OC4 derived chlorophyll-a versus Sentinel-2 MSI RI/RCA retrieved chlorophyll-a ( $\text{mgm}^{-3}$ ) averaged over 36 data stations

### 5.7. Evaluating the Atmospheric Correction method

Several atmospheric correction models have been used in the removal of atmospheric artefacts for both coastal and in-land water application. ACOLITE atmospheric processor algorithm developed by (Vanhellemont & Ruddick, 2014) was applied on Sentinel-2 images used in the study while MODIS 1 km ocean color reflectance product are atmospherically corrected using the NIR algorithm for coastal case 1 waters (Gordon and Wang, 1994) implemented in the satellite data processing software (SeaDAS: SeaWiFS Data Analysis System software, <https://seadas.gsfc.nasa.gov/>).

From the statistical results of the atmospheric correction methods shown in Figure 11, the mean absolute percentage error (42 %) estimated for band 443/443 was high than the uncertainty and 28 % (at 18 %) estimated for band and 488/490 nm and 547/560 nm respectively. The ACOLITE processor derived atmospheric corrected reflectance products are high than the MODIS ocean atmospheric corrected reflectance products for band 443 nm and 488 nm. However, we do not have available in-situ remote sensing measurement to validate the results if overestimation is by ACOLITE processor or underestimation by NIR algorithm (Al Shehhi, Gherboudj, Zhao, & Ghedira, 2017). But, Al Shehhi, et al., (2017) found that by MODIS NIR implemented in the satellite data processing software (SeaDAS: SeaWiFS Data Analysis System software, <https://seadas.gsfc.nasa.gov/>) underestimates the water leaving reflectance after AC at wavelengths 412 nm, 443 nm, and 531 nm compared with in-situ measurements over the turbid and shallow region of the Arabian Sea including the Arabian Gulf.

Since the uncertainties in the observed remote sensing reflectance high especially the blue bands, the results suggest that the uncertainties in the retrieved Chl-a data by the two models (MODIS OC3M and Sentinel-2 MSI RI/RCA) is due to the parameterization of the OC3M based on band ratio which is mostly not applicable in most coastal complex case 2 waters (Blondeau-Patissier et al., 2014). Also, the RI/RCA model may need an intense model tuning with in-situ reflectance measurements to improve results in Chl-a estimation over the shallow and turbid region of the Arabian Gulf.

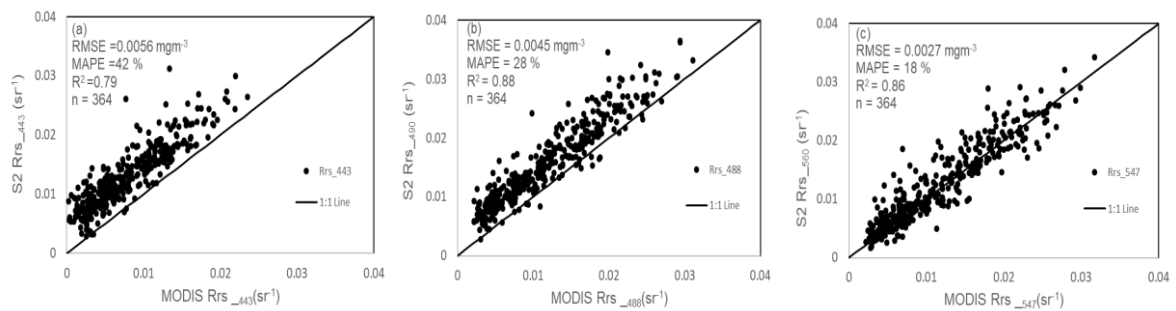


Figure 11: Comparison between MODIS Ocean color reflectance products (Rrs) and Sentinel-2 atmospheric corrected reflectance product (Rrs) where 443/443 (a); 488/490 and 547/560 for MODIS and Sentinel-2/MSI respectively for all 12 images obtained.

### 5.8. Comparing results of aggregated and disaggregated Chl-a products

Figure 12 shows the scatter plot of Sentinel-2 MSI aggregated to 90 m and 1 km against MODIS L2 Chl-a product disaggregated at 90 m and its 1 km spatial resolution at nadir for images acquired on 22 September 2017. For the aggregated and disaggregated Sentinel-2 and MODIS Chl-a data at 90 m respectively, a single MODIS pixel contains a large number of varied Chl-a data. This may have been caused by spatially downgrading the 1 km pixel to 90 m which is about 109 x 109 pixels of Sentinel-2 shown as Figure 10.a. Estimated RMSE value of 2.38 mg/m<sup>3</sup> and MAPE value of 97 % were estimated for the upscaled Sentinel-2 MSI Chl-a map and the downgraded MODIS L2 Chl-a map. However, spatially aggregating Sentinel-2 to approximately the same resolution of MODIS data at 1 km, this large variability in Chl-a range within a single pixel reduced but with an increased RMSE (2.68 mg/m<sup>3</sup>) and MAPE (90 %). Hence more or large number of pixel values in the disaggregated MODIS and aggregated Sentinel-2 product a be due to the spatial variability in sensor comparison and not due to noise of the sensor.

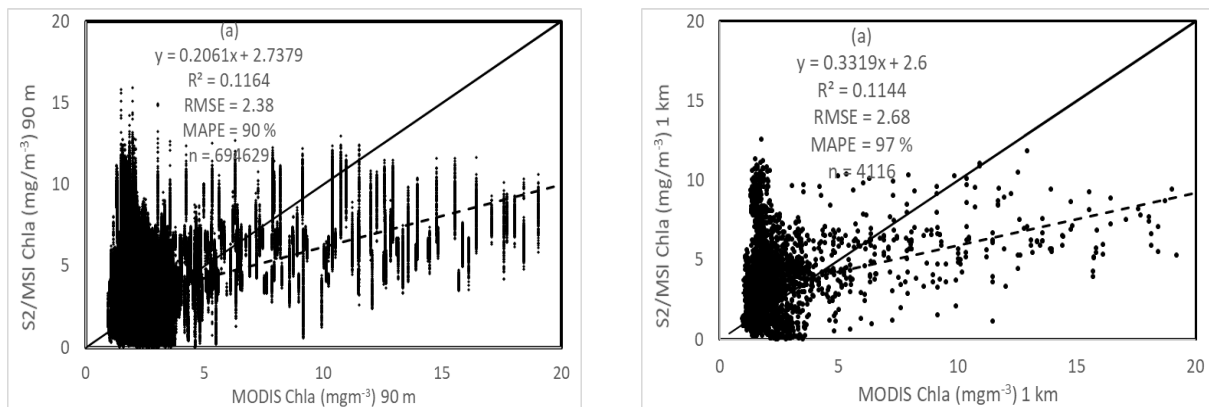


Figure 12: Comparison of Sentinel-2/MSI against MODIS L2 Chl-a product at 90 m critical aggregate level (a) and at 1 km MODIS spatial resolution (b) for images acquired on 22 September 2017

Figure 13 shows the comparison of MODIS reference Chl-a and Sentinel-2 derived Chl-a maps aggregated and disaggregated at 90 m; and (1 km: From Figure 13, it can be clearly seen that, the Sentinel-2 RI derived Chl-a show into a much fine spatial detail than the MODIS Chl-a and hence algal bloom patches and patterns easily be delineated. Moreover, at concentration less than 4 mg/m<sup>3</sup>, the OC3M and RI/RCA derived Chl-a concentration are comparable. However, there were few over-estimations for Sentinel-2 derived Chl-a which may need further investigation. However, large disparities arise at high MODIS OC3M derived Chl-a concentration. It was found that, Sentinel-2 RI/RCA underestimates at chlorophyll concentration greater than ~10 -15 mg/m<sup>3</sup> when the area is near-shore. However, it should be noted that, these high MODIS observed Chl-a concentration were all found near-shore (See Figure 13). Hence, the contribution of bottom reflectance, CDOM and non-algal materials might be pronounced in the results.

Although the OC3M algorithm, has been used previously by many researchers to qualitatively derive Chl-a concentration in oligotrophic waters bodies, however, poor results have been obtained in and some few



successful results in case 2 waters due to its inability to distinguish Chl-a from non-algal materials. Moreover, it has been reported that, it is incapable of differentiating bloom from non-bloom water conditions. Hence, we could not qualitatively compare the results of 10 m downscaled Sentinel-2/MSI from the OC3M derived Chl-a from MODIS. Nonetheless, the use of in-situ measurement to accurately validate this is needed.

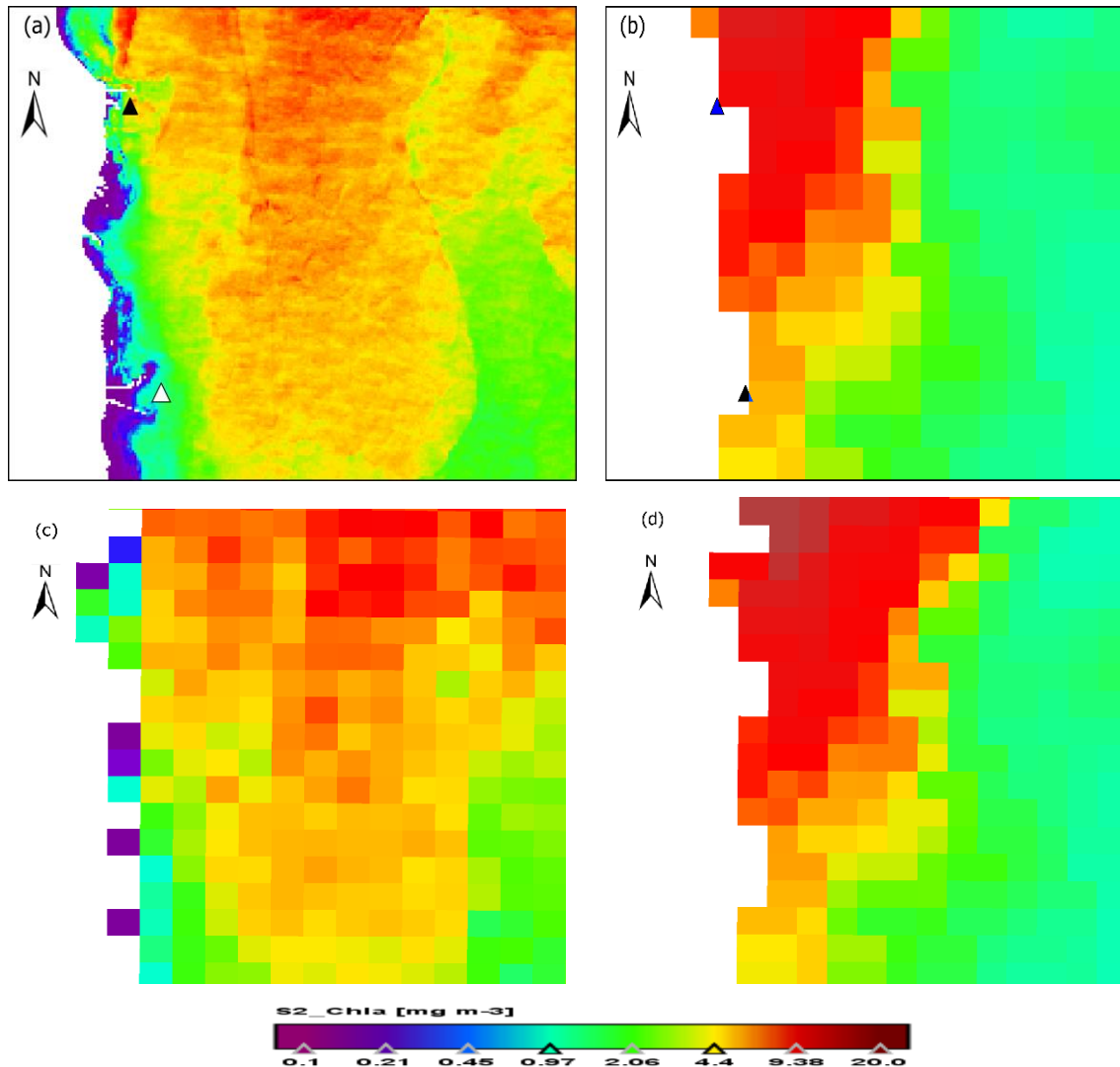


Figure 13: Comparison of MODIS reference Chl-a and Sentinel-2 derived Chl-a at 90 m for (a) and (b); and 1 km for (c) and (d): Small triangles represent intakes of desalination plants; top (intake of RAS A3 (SWRO system) and down intake 3 is the intake of Ras A2 (multi-stage flash distillation (MSF)).

### 5.9. Residual map for MODIS and Sentinel-2 Chlorophyll-a maps

In many cases, the downscaling algorithm/equation derived from regression analysis of model fitting used for the downscaling may not be enough for the downscaling (Wang, Shi, Atkinson, & Zhao, 2015), hence residual correction methods are usually applied to compensate for the residual in order obtain the final downscaled product. The result of subtracting Sentinel-2 derived Chl-a aggregated to 90 and 1 km shown in Figure 14 a and b respectively for image acquired on 22 September 2017. The residual error obtained

through the downscaling approach of deriving the final Chl-a product is evident in Figures 14. Area of map that are red colored shows areas where MODIS observed Chl-a were high than those derived from Sentinel-2 MSI while blue colored areas indicates areas where RCA estimates of Chl-a were greater than the OC3M derived Chl-a. It is evident from the two images that, MODIS OC3M algorithm, overestimates Chl-a near-shore. Hence the contribution of bottom reflectance and CDOM may be abundant in the Chl-a estimates revealed as deep red colors. The errors in retrieved Chl-a product were within the range of -7.23 to 7.43 mg/m<sup>3</sup> for the image acquired on 22 September 2017. In all, the average residual for the whole study period range between -0.5 to 1.6 mg/m<sup>3</sup> (See Annex G). The values of errors obtained are in line with similar findings by (Pereira, Harari, & de Camargo, 2014) in which the residual error in estimated Chl-a determined from MODIS and numerical modelling for the Tropical and South Atlantic Ocean were within the range of <0.005 to 10 mg/m<sup>3</sup>. The errors obtain were too large during bloom periods, which may be due to the difference in sensitivity of the models (RI/RCA and OC3M) to Chl-a in the study area and also due to the complexity and localized nature of the region. We currently do not have in-situ measurement to accurately determine which of the models can best be used to delineate algal blooms. Hence adding residual to the final downscaled Chl-a product may not be appropriate in this instance due to errors which may be introduced to the original derived Chl-a maps at 10 m.

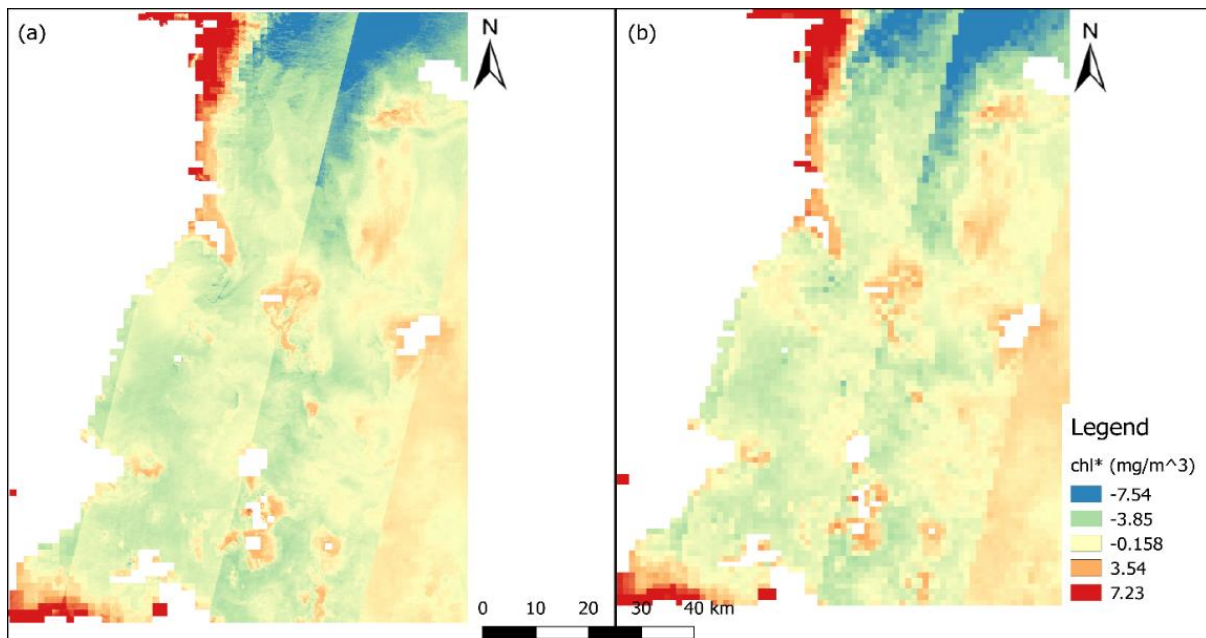


Figure 14: Spatial variation in residual error of Sentinel-2 derived Chl-a aggregated to (a) 90 and (b) 1 km for image acquired on 22 September 2017

## 6. CONCLUSION AND RECOMMENDATIONS

### 6.1. Conclusion

In this study, we investigated a newly developed approach using regression analysis as an attempt to downscale MODIS L2 1 km ocean color (Chl-a) product with ACOLITE atmospheric corrected Sentinel-2 MSI indices data extracted with several average spatial moving window sizes ranging from 3 x 3 to 99 x 99 to obtain Chl-a maps at high spatial resolution. Although MODIS ocean colour products have been freely available for a long time and provides twice daily Chl-a data, its usage in the monitoring of complex coastal case 2 waters have been limited due to reported cases of poor performance of the standard OC3M algorithm affected by CDOM, non- algal particles and bottom reflectance over such waters and also its spatial resolution is too coarse to track and observe small variations in water quality parameters that may lead to HABs. Nonetheless, accurate Chl-a results from been retrieved using the MODIS ocean colour product by previous researchers over coastal oceans. Three semi-empirical algorithms based on band ratios comprising RI, MCI, and NDCI models derived from Sentinel-2 MSI data (predictor variables) tested with linear, polynomial, exponential and power models were used to relate them with MODIS L2 1 km Chl-a product (response variable) to develop final algorithm for Chl-a retrieval.

The RI/RCA algorithm derived with an exponential model and run with a 9 x 9 average moving window size proved to be the best model and the critical aggregation level respectively for relating Sentinel-2 MSI indices data with MODIS L2 1 km Chl-a product. High resolution Chl-a maps at 10 m spatial resolution were obtained from the application of the RI/RCA algorithm on the Sentinel-2 images with a resultant high accuracy in Chl-a retrieval for the study area with an estimated RMSE and MAPE values of 2.75 mg/m<sup>3</sup> and 43% respectively for the whole Chl-a range. The MCI model also showed a good potential for use as Chl-a retrieval algorithm for the area of study. However, poor results were estimated for the NDCI model and hence is inappropriate for use in the study area.

There was high consistency in retrieved Chl-a products for Sentinel-2 MSI data compared with MODIS L2 Chl-a data for chlorophyll-a at low concentration at  $\sim < 3$  mg/m<sup>3</sup> and with a decreasing trend in Chl-a concentration from near-shore to off-shore in both maps. But large disparities in retrieved Chl-a products were seen at high Chl-a concentration at  $\sim > 10$  mg/m<sup>3</sup>. Nonetheless, these areas were very found off-shore and at the very shallow region of the study area. However, the newly developed RI/RCA algorithms generated 10 m Chl-a maps resolved into more details the spatial variation in chlorophyll-a concentration near the intake of Ras Abu Fontas desalination plant in Qatar off the complex coastal waters of the Arabian Gulf region. Hence, the RI/RCA algorithm was also capable of differentiating patches of bloom from non-bloom water conditions other than the MODIS OC3M algorithm Chl-a derived maps and hence could be used as proxies for HAB monitoring and development.

The Sentinel-2 MSI retrieved Chl-a data showed a similar seasonal trend in observed Chl-a as MODIS L2 Chl-a product. The seasonal variation of observed Chl-a concentration follows similar finding by Nezlin et al., (2010) for research conducted over the shallow western and southern regions of the AG, observing minimum chlorophyll-a concentration in spring and maximum in late summer and Autumn. However, failure of our newly developed Chl-a algorithm for the study area to accurately estimate extremely high Chl-a values at the high spatial resolution was still unaccounted for in this study. High chlorophyll-a concentration was observed at the Northwestern part of the study area close to the RAS A3 (SWRO system) intakes  $\sim > 10 \text{ mg/m}^3$  for image acquired on 22 September 2017 for both Sentinel-2 MSI derived Chl-a and MODIS L2 observed Chl-a data confirming an incidence of algal bloom. However, there were no in-situ measurements to validate these results. But report from project site indicated a possible shut-down of the RAS A3 (SWRO system) desalination plant within this period.

In relating upscaled Sentinel-2 MSI derived Chl-a and the disaggregated MODIS L2 observed Chl-a product with each at 90 m and at 1km with scatterplots, revealed that, the spatial variation between the two products may have caused an increased vertical stripping in 90 m scatter plots rather than noise in the Sentinel-2 MSI. Compared with scatter plot analysis for upscaled Sentinel-2 Chl-a data and the reference MODIS L2 Chl-a data at 1 km showed a reduction in vertical stripping. This is because the Sentinel-2 MSI data was spatially smoothed to reduce noise in observed reflectance data during atmospheric correction application processing. Although both sensors showed different sensitivity to HAB with accuracy results exceeding the maximum tolerable error ( $\pm 35 \%$ ) in satellite retrieved Chl-a. But, the residual error estimated using upscaled Sentinel-2 MSI derived Chl-a versus the reference downgraded MODIS L2 each at 90 m and 1 km range from  $-7.2$  to  $7.5 \text{ mg/m}^3$  which is within typical uncertainty in modelled and satellite retrieved chlorophyll-a measurements reported to be  $\sim 10 \text{ mg/m}^3$  (Pereira et al., 2014). Moreover, the combined effect of bottom reflectance and total suspended particulate matter led to an increase in error in retrieved Chl-a found near-shore. Hence there were large disparity in observed Chl-a concentration between MODIS derived Chl-a and Sentinel-2 RI/RCA retrieved Chl-a near-shore than at off-shore.

Overall, with the improve spectral resolution of Sentinel-2 from introduce by MODIS through the downscaling approach of regression analysis, the generated maps at fine spatial scale, shows potentials of using the newly developed RI/RCA model to estimate Chl-a. We can also can accurately track the onset, development, spread and trends of algal bloom with results of newly developed maps with RI/RCA than OC3M derived Chl-a over the study area throughout the different seasons for a long period of time. However, there were no simultaneous coincident in-situ measurements to validate the results of the atmospheric correction and retrieved chlorophyll-a data, hence, there is a need for coherent field measured chlorophyll-a data to validate the performance of the derived model applied to the datasets for chlorophyll-a retrieval used in this study.

## 6.2. Recommendations

The following recommendations are drawn for future works to optimize results attained in this research;

1. The performance of ACOLITE atmospheric correction may need to be validated to evaluate if errors in retrieved products in some part may be due to the AC method, therefore, simultaneous coincident in-situ radiometric measured data is required.
2. The derived approach of algorithm development although pertains to the area studied can be applied or regionally extended to the whole Arabian Gulf to examine the model performance for the different regions in the study area. Also, it can be used to investigate the long-term trend in Chl-a variation in under-sampled areas and under eutrophic conditions for inland water bodies.
3. A follow-up of testing the downscaling approach of regression analysis used in this study for MERIS Chl-a data at 300 m with Sentinel-2 MSI and/or Landsat 8 OLI may also be more appropriate to retrieve Chl-a and monitor HABs at high spatial scale.
4. Integrating MCI with RI/RCA can improve delineation and early detection of bloom patches from non-bloom waters although results of the analysis of MCI model indicated overestimation and incapable of retrieving very low Chl-a concentration ( $\sim < 1 \text{ mg/m}^3$ ). However, both models will need a robust tuning and calibration using in-situ measured data.
5. Incorporating sea-surface temperature (SST), total suspended matter (TSM) and other environmental factors responsible of HABs development with RI/RCA and MCI in future research to separate bloom from non-bloom water condition is also recommended to avoid possible false alarms of HABs.

## REFERENCES

- Ahn, Y.-H., & Shanmugam, P. (2006). Detecting the red tide algal blooms from satellite ocean color observations in optically complex Northeast-Asia Coastal waters. *Remote Sensing of Environment*, 103(4), 419–437. <https://doi.org/10.1016/j.rse.2006.04.007>
- Al-Ansari, E. M. A. S., Rowe, G., Abdel-Moati, M. A. R., Yigiterhan, O., Al-Maslmani, I., Al-Yafei, M. A., ... Upstill-Goddard, R. (2015). Hypoxia in the central Arabian Gulf Exclusive Economic Zone (EEZ) of Qatar during summer season. *Estuarine, Coastal and Shelf Science*, 159, 60–68. <https://doi.org/10.1016/j.ecss.2015.03.022>
- Al-Naimi, N., Raitsos, D., Ben-Hamadou, R., & Soliman, Y. (2017). Evaluation of Satellite Retrievals of Chlorophyll-a in the Arabian Gulf. *Remote Sensing*, 9(3), 301. <https://doi.org/10.3390/rs9030301>
- Al-Shehhi, M. R., Gherboudj, I., & Ghedira, H. (2017). In situ spectral response of the Arabian Gulf and Sea of Oman coastal waters to bio-optical properties. *Journal of Photochemistry and Photobiology B: Biology*, 175(September), 235–243. <https://doi.org/10.1016/j.jphotobiol.2017.09.007>
- Al Kaabi, M., Zhao, J., & Ghedira, H. (2016). MODIS-Based Mapping of Secchi Disk Depth Using a Qualitative Algorithm in the Shallow Arabian Gulf. *Remote Sensing*, 8(5), 423. <https://doi.org/10.3390/rs8050423>
- Al Shehhi, M. R., Gherboudj, I., & Ghedira, H. (2014). An overview of historical harmful algae blooms outbreaks in the Arabian Seas. *Marine Pollution Bulletin*, 86(1–2), 314–324. <https://doi.org/10.1016/j.marpolbul.2014.06.048>
- Al Shehhi, M. R., Gherboudj, I., & Ghedira, H. (2017). Satellites-Based Monitoring of Harmful Algal Blooms for Sustainable Desalination. *Desalination Sustainability*, 341–366. <https://doi.org/10.1016/B978-0-12-809791-5.00009-2>
- Al Shehhi, M. R., Gherboudj, I., Jun, Z., Mezhoud, N., & Ghedira, H. (2013). Evaluating the performance of MODIS FLH ocean color algorithm in detecting the Harmful Algae Blooms in the Arabian Gulf and the Gulf of Oman. *Oceans - San Diego*, 2013, 1–7.
- Al Shehhi, M. R., Gherboudj, I., Zhao, J., & Ghedira, H. (2017). Improved atmospheric correction and chlorophyll-a remote sensing models for turbid waters in a dusty environment. *ISPRS Journal of Photogrammetry and Remote Sensing*, 133, 46–60. <https://doi.org/10.1016/j.isprsjprs.2017.09.011>
- Ambarwulan, W. (2010). *Remote sensing of tropical coastal waters: Study of the Berau Estuary, East Kalimantan, Indonesia*.

- Anderson, D. M., Glibert, P. M., & Burkholder, J. M. (2002). Harmful algal blooms and eutrophication: Nutrient sources, composition, and consequences. *Estuaries*, 25(4), 704–726.  
<https://doi.org/10.1007/BF02804901>
- Anderson, D., & Price, K. (2015). Combatting the Emerging Impacts of Harmful Algal Blooms (HABs) on Desalination Plants: Bloom Detection: Forecasting and Strategies for Impact, (January), 1–19.
- Arun Kumar, S. V. V., Babu, K. N., & Shukla, A. K. (2015). Comparative Analysis of Chlorophyll-a Distribution from SeaWiFS, MODIS-Aqua, MODIS-Terra and MERIS in the Arabian Sea. *Marine Geodesy*, 38(1), 40–57. <https://doi.org/10.1080/01490419.2014.914990>
- Atkinson, P. M. (2013). Downscaling in remote sensing. *International Journal of Applied Earth Observations and Geoinformation*, 22, 106–114. <https://doi.org/10.1016/j.jag.2012.04.012>
- Bauman, A. G., Burt, J. A., Feary, D. A., Marquis, E., & Usseglio, P. (2010). Tropical harmful algal blooms: An emerging threat to coral reef communities?  
<https://doi.org/10.1016/j.marpolbul.2010.08.015>
- Beck, R., Xu, M., Zhan, S., Liu, H., Johansen, R., Tong, S., ... Huang, Y. (2017). Comparison of Satellite Reflectance Algorithms for Estimating Phycocyanin Values and Cyanobacterial Total Biovolume in a Temperate Reservoir Using Coincident Hyperspectral Aircraft Imagery and Dense Coincident Surface Observations. *Remote Sensing*, 9(6), 538. <https://doi.org/10.3390/rs9060538>
- Berktaş, A. (2011). Environmental Approach and Influence of Red Tide to Desalination Process in the Middle East Region. *International Journal of Chemical and Environmental Engineering*.
- Binding, C. E., Greenberg, T. A., & Bukata, R. P. (2013). The MERIS Maximum Chlorophyll Index; its merits and limitations for inland water algal bloom monitoring. *Journal of Great Lakes Research*, 39(S1), 100–107. <https://doi.org/10.1016/j.jglr.2013.04.005>
- Bisquert, M., Sanchez, M. J., & Caselles, V. (2002). Evaluation of Disaggregation Methods for downscaling MODIS Land Surface Temperature to Landsat Spatial Resolution in Barrax Test Site, 4(8), 1–4.  
<https://doi.org/10.1109/JSTARS.2016.2519099>
- 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>
- Boerlage, S., & Nada, N. (2015). Algal toxin removal in seawater desalination processes. *Desalination and Water Treatment*, 55(10), 2575–2593. <https://doi.org/10.1080/19443994.2014.947785>
- Brockmann, C., Roland, Peters, M., Stelzer, K., Sabine, & Ruescas, A. (2016). EVOLUTION OF THE C2RCC NEURAL NETWORK FOR SENTINEL 2 AND 3 FOR THE RETRIEVAL OF OCEAN COLOUR PRODUCTS IN NORMAL AND EXTREME OPTICALLY COMPLEX WATERS.
- Cannizzaro, J. P., & Carder, K. L. (2006). Estimating chlorophyll a concentrations from remote-sensing reflectance in optically shallow waters. *Remote Sensing of Environment*, 101(1), 13–24.

<https://doi.org/10.1016/j.rse.2005.12.002>

- Caron, D. A., Garneau, M.-V., Seubert, E., Howard, M. D. A., Darjany, L., Schnetzer, A., ... Caron, D. A. (2009). Harmful algae and their potential impacts on desalination operations off southern California. <https://doi.org/10.1016/j.watres.2009.06.051>
- Che, X., Feng, M., Jiang, H., Song, J., & Jia, B. (2015). Downscaling MODIS surface reflectance to improve water body extraction. *Advances in Meteorology*, 2015. <https://doi.org/10.1155/2015/424291>
- Chen, J., Zhu, W., Tian, Y. Q., Yu, Q., Zheng, Y., & Huang, L. (2017). Remote estimation of colored dissolved organic matter and chlorophyll-a in Lake Huron using Sentinel-2 measurements. *J. Appl. Remote Sens*, 11(4). <https://doi.org/10.1117/1.JRS.11>
- Chu, C. P., & Kuo, Y. (2010). Detection of Red Tides in the Southwestern Florida Coastal Region Using Ocean Color Data. In *Proceedings on MTS/IEEE OCEANS 10, Seattle, Washington, USA*, . Seattle, Washington.
- Clark, J. M., Schaeffer, B. A., Darling, J. A., Urquhart, E. A., Johnston, J. M., Ignatius, A. R., ... Stumpf, R. P. (2017). Satellite monitoring of cyanobacterial harmful algal bloom frequency in recreational waters and drinking water sources. *Ecological Indicators*, 80(April), 84–95. <https://doi.org/10.1016/j.ecolind.2017.04.046>
- Comprehensive Studies Task Team of Group Coordinating Sea Disposal Monitoring (CSTT). (1997). *Comprehensive studies for the purposes of Article 6 & 8.5 of DIR 91/271 EEC, the Urban Waste Water Treatment Directive*. Edinburgh, UK.
- Conley, D. J., Paerl, H. W., Howarth, R. W., Boesch, D. F., Seitzinger, S. P., Havens, K. E., ... Likens, G. E. (2009). ECOLOGY: Controlling Eutrophication: Nitrogen and Phosphorus. *Science*, 323(5917), 1014–1015. <https://doi.org/10.1126/science.1167755>
- D’Odorico, P., Gonsamo, A., Damm, A., & Schaepman, M. E. (2013). Experimental Evaluation of Sentinel-2 Spectral Response Functions for NDVI Time-Series Continuity. *IEEE Transactions on Geoscience and Remote Sensing*, 51(3), 1336–1348. <https://doi.org/10.1109/TGRS.2012.2235447>
- D’Ortenzio, F., Marullo, S., Ragni, M., Ribera d’Alcalà, M., & Santoleri, R. (2002). Validation of empirical SeaWiFS algorithms for chlorophyll-a retrieval in the Mediterranean Sea: A case study for oligotrophic seas. *Remote Sensing of Environment*, 82(1), 79–94. [https://doi.org/10.1016/S0034-4257\(02\)00026-3](https://doi.org/10.1016/S0034-4257(02)00026-3)
- Darwish, M. A., Abdulrahim, H. K., Hassan, A. S., & Shomar, B. (2016). Reverse osmosis desalination system and algal blooms Part I: harmful algal blooms (HABs) species and toxicity. *Desalination and Water Treatment*, 57(54), 25859–25880. <https://doi.org/10.1080/19443994.2016.1159618>
- Díaz-Uriarte, R., Alvarez de Andrés, S., Mutanga, O., Ismail, R., Dgpf, D., West, A., & NI, C. (2014). Downscaling land surface temperatures from MODIS data to mesoscale resolution with Random Forest regression. *BMC Bioinformatics*, 7(Vi), 3. <https://doi.org/10.1186/1471-2105-7-3>
- Dörnhöfer, K., Göritz, A., Gege, P., Pflug, B., & Oppelt, N. (2016). Water constituents and water depth retrieval from Sentinel-2A-A first evaluation in an oligotrophic lake. *Remote Sensing*, 8(11).



<https://doi.org/10.3390/rs8110941>

- Elkadiri, R., Manche, C., Sultan, M., Al-Dousari, A., Uddin, S., Chouinard, K., & Abotalib, A. Z. (2016). Development of a Coupled Spatiotemporal Algal Bloom Model for Coastal Areas: A Remote Sensing and Data Mining-Based Approach. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 9(11), 5159–5171. <https://doi.org/10.1109/JSTARS.2016.2555898>
- ESA. (2017). Meet the Sentinel-2 Data Quality Manager - Thematic Areas - Sentinel Online. Retrieved August 17, 2017, from <https://earth.esa.int/web/sentinel/thematic-content/-/article/meet-the-sentinel-2-data-quality-manager>
- Ghanea, M., Moradi, M., & Kabiri, K. (2015). Investigation the behavior of MODIS ocean color products under the 2008 red tide in the eastern Persian Gulf. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 40(1W5), 227–232. <https://doi.org/10.5194/isprsarchives-XL-1-W5-227-2015>
- Ghanea, M., Moradi, M., & Kabiri, K. (2016). A novel method for characterizing harmful algal blooms in the Persian Gulf using MODIS measurements. *Advances in Space Research*, 58(7), 1348–1361. <https://doi.org/10.1016/j.asr.2016.06.005>
- Gholizadeh, M. H., Melesse, A. M., & Reddi, L. (2016). A Comprehensive Review on Water Quality Parameters Estimation Using Remote Sensing Techniques. *Sensors (Basel, Switzerland)*, 16(8), 1298. <https://doi.org/10.3390/s16081298>
- Gower, J., King, S., Borstad, G., & Brown, L. (2005). Detection of intense plankton blooms using the 709 nm band of the MERIS imaging spectrometer. *International Journal of Remote Sensing*, 26(9), 2005–2012. <https://doi.org/10.1080/01431160500075857>
- Gower, J., King, S., & Goncalves, P. (2008). Global monitoring of plankton blooms using MERIS MCI. *International Journal of Remote Sensing*, 29(21), 6209–6216. <https://doi.org/10.1080/01431160802178110>
- Gregg, W. W., & Casey, N. W. (2004). Global and regional evaluation of the SeaWiFS chlorophyll data set. *Remote Sensing of Environment*, 93(4), 463–479. <https://doi.org/10.1016/J.RSE.2003.12.012>
- Heisler, J., Glibert, P. M., Burkholder, J. M., Anderson, D. M., Cochlan, W., Dennison, W. C., ... Suddleson, M. (2008). Eutrophication and harmful algal blooms: A scientific consensus. <https://doi.org/10.1016/j.hal.2008.08.006>
- Hu, C., Lee, Z., & Franz, B. (2012). Chlorophyll a algorithms for oligotrophic oceans: A novel approach based on three-band reflectance difference. *Journal of Geophysical Research: Oceans*, 117(C1). <https://doi.org/10.1029/2011JC007395>
- IOCCG. (2006). IOCCG Report Number 05: Reports of the International Ocean-Colour Coordinating Group Remote Sensing of Inherent Optical Properties : Fundamentals, Tests of Algorithms, and Applications. *IOCCG Report 5*, 5(5), 126. <https://doi.org/10.1006/jmbi.1998.2073>
- Jun, Z., Marouane, T., Sheikha, A. K., Nahla, M., Zhao, J., Temimi, M., ... Mezhoud, N. (2016). Monitoring HABs in the shallow Arabian Gulf using a qualitative satellite-based index. *International*

- Journal of Remote Sensing*, 37(8), 1937–1954. <https://doi.org/10.1080/01431161.2016.1165886>
- Jun Zhao, Ghedira, H., & Temimi, M. (2014). Remote sensing of red tide in the Arabian Gulf. In 2014 *IEEE Geoscience and Remote Sensing Symposium* (pp. 3870–3873). IEEE. <https://doi.org/10.1109/IGARSS.2014.6947329>
- Kahru, M., Kudela, R. M., Anderson, C. R., Manzano-Sarabia, M., & Mitchell, B. G. (2014). Evaluation of satellite retrievals of ocean chlorophyll-a in the California current. *Remote Sensing*, 6(9), 8524–8540. <https://doi.org/10.3390/rs6098524>
- Kahru, M., Kudela, R. M., Anderson, C. R., & Mitchell, B. G. (2015). Optimized Merger of Ocean Chlorophyll Algorithms of MODIS-Aqua and VIIRS. *IEEE Geoscience and Remote Sensing Letters*, 12(11), 2282–2285. <https://doi.org/10.1109/LGRS.2015.2470250>
- Kamerosky, A., Cho, H. J., & Morris, L. (2015). Monitoring of the 2011 super algal bloom in Indian River Lagoon, FL, USA, Using MERIS. *Remote Sensing*, 7(2), 1441–1460. <https://doi.org/10.3390/rs70201441>
- Kotchenova, S. Y., Vermote, E. F., Matarrese, R., & Klemm, F. J. (2006). Validation of a vector version of the 6S radiative transfer code for atmospheric correction of satellite data. Part I: path radiance. *Applied Optics*, 45(26), 6762–74.
- Kustas, W. P., Norman, J. M., Anderson, M. C., & French, A. N. (2003). Estimating subpixel surface temperatures and energy fluxes from the vegetation index-radiometric temperature relationship. *Remote Sensing of Environment*, 85(4), 429–440. [https://doi.org/10.1016/S0034-4257\(03\)00036-1](https://doi.org/10.1016/S0034-4257(03)00036-1)
- Kutser, T., Paavel, B., Verpoorter, C., Ligi, M., Soomets, T., Toming, K., & Casal, G. (2016). Remote Sensing of Black Lakes and Using 810 nm Reflectance Peak for Retrieving Water Quality Parameters of Optically Complex Waters. *Remote Sensing*, 8(12), 497. <https://doi.org/10.3390/rs8060497>
- Liu, H., Li, Q., Shi, T., Hu, S., Wu, G., & Zhou, Q. (2017). Application of Sentinel 2 MSI Images to Retrieve Suspended Particulate Matter Concentrations in Poyang Lake. *Remote Sensing*, 9(7), 761. <https://doi.org/10.3390/rs9070761>
- Majid, S. A., Graw, M. F., Chatziefthimiou, A. D., Nguyen, H., Richer, R., Louge, M., ... Hay, A. G. (2016). Microbial characterization of Qatari Barchans and Dunes. *PLoS ONE*, 11(9), 1–22. <https://doi.org/10.1371/journal.pone.0161836>
- Mezhoud, N., Temimi, M., Zhao, J., Al Shehhi, M. R., & Ghedira, H. (2016). Analysis of the spatio-temporal variability of seawater quality in the southeastern Arabian Gulf. *Marine Pollution Bulletin*, 106(1), 127–138. <https://doi.org/10.1016/j.marpolbul.2016.03.016>
- Mishra, D. R., Schaeffer, B. A., & Keith, D. (2014). Performance evaluation of normalized difference chlorophyll index in northern Gulf of Mexico estuaries using the Hyperspectral Imager for the Coastal Ocean. *GIScience & Remote Sensing*, 51(2), 175–198. <https://doi.org/10.1080/15481603.2014.895581>
- Mishra, S., & Mishra, D. R. (2012). Normalized difference chlorophyll index: A novel model for remote estimation of chlorophyll-a concentration in turbid productive waters. *Remote Sensing of Environment*,

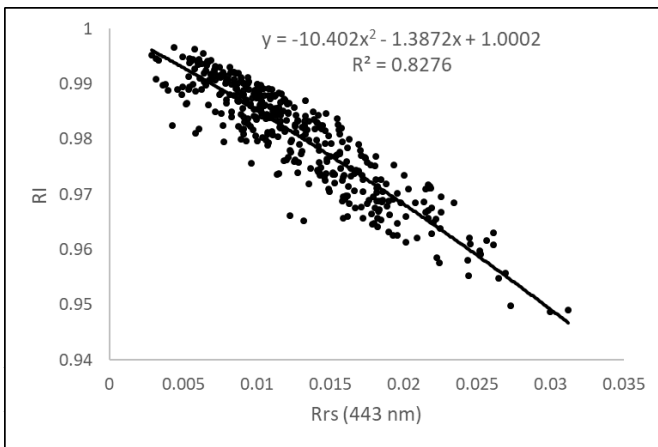
- 117(January 2014), 394–406. <https://doi.org/10.1016/j.rse.2011.10.016>
- Moore, T. S., Campbell, J. W., & Dowell, M. D. (2009). A class-based approach to characterizing and mapping the uncertainty of the MODIS ocean chlorophyll product. *Remote Sensing of Environment*, 113(11), 2424–2430. <https://doi.org/10.1016/j.rse.2009.07.016>
- Moradi, M., & Kabiri, K. (2012). Red tide detection in the Strait of Hormuz (east of the Persian Gulf) using MODIS fluorescence data. *International Journal of Remote Sensing*, 33(4), 1015–1028. <https://doi.org/10.1080/01431161.2010.545449>
- Moradi, N., Hasanlou, M., & Saadatesresht, M. (2016). Ocean color retrieval using Landsat-8 imagery in coastal case 2 waters. <https://doi.org/10.5194/isprsarchives-XLI-B8-1161-2016>
- Nezlin, N. P., Polikarpov, I. G., & Al-Yamani, F. (2007). Satellite-measured chlorophyll distribution in the Arabian Gulf: Spatial, seasonal and inter-annual variability. *International Journal of Oceans and Oceanography*, 2(1), 139–156.
- Nezlin, N. P., Polikarpov, I. G., Al-Yamani, F. Y., Subba Rao, D. V., & Ignatov, A. M. (2010). Satellite monitoring of climatic factors regulating phytoplankton variability in the Arabian (Persian) Gulf. *Journal of Marine Systems*, 82, 47–60. <https://doi.org/10.1016/j.jmarsys.2010.03.003>
- Njuki, S. M. (2016). Assessment of Irrigation Performance by Remote Sensing in the Naivasha Basin , Kenya. *ITC MSc Thesis*, 65.
- Ogashawara, I., Li, L., & Moreno-Madriñán, M. J. (2016). Slope algorithm to map algal blooms in inland waters for Landsat 8/Operational Land Imager images. *Journal of Applied Remote Sensing*, 11(1), 12005. <https://doi.org/10.1117/1.JRS.11.012005>
- Olalekan, A. A., & Malik, K. (2015). Application of Giovanni for rapid assessment of harmful algal blooms in the Arabian Gulf. *Arabian Journal of Geosciences*, 8(10), 8767–8775. <https://doi.org/10.1007/s12517-015-1826-3>
- Patra, P. P., Dubey, S. K., Trivedi, R. K., Sahu, S. K., & Rout, S. K. (2017). Estimation of chlorophyll-a concentration and trophic states in Nalban Lake of East Kolkata Wetland, India from Landsat 8 OLI data. *Spatial Information Research*, 25(1), 75–87. <https://doi.org/10.1007/s41324-016-0069-z>
- Pei, W., Yao, S., Dong, S., Knight, J. F., Xu, C., & Chen, Y. (2015). Using field spectral measurements to estimate chlorophyll-a in waterlogged areas of Huainan, China. *GIScience & Remote Sensing*, 52(6), 660–679. <https://doi.org/10.1080/15481603.2015.1082173>
- Pereira, N. E. S., Harari, J., & de Camargo, R. (2014). Correlations of Remote Sensing Chlorophyll-a Data and Results of A Numerical Model of the Tropical and South Atlantic Ocean Circulation. *Geology and Geosciences*, 3(6), 1–10. <https://doi.org/10.4172/2329-6755.1000171>
- Richlen, M. L., Morton, S. L., Jamali, E. A., Rajan, A., & Anderson, D. M. (2010). The catastrophic 2008–2009 red tide in the Arabian gulf region, with observations on the identification and phylogeny of the fish-killing dinoflagellate *Cochlodinium polykrikoides*. *Harmful Algae*, 9(2), 163–172. <https://doi.org/10.1016/j.hal.2009.08.013>
- Salama, M. S., Van Der Velde, R., Van Der Woerd, H. J., Kromkamp, J. C., Philippart, C. J. M., Joseph, A.

- T., ... Su, Z. (2012). Technical Note: Calibration and validation of geophysical observation models. *Biogeosciences*, 9(6), 2195–2201. <https://doi.org/10.5194/bg-9-2195-2012>
- Salem, S., Strand, M., Higa, H., Kim, H., Kazuhiro, K., Oki, K., & Oki, T. (2017). Evaluation of MERIS chlorophyll-a retrieval processors in a complex turbid Lake Kasumigaura through 10-year mission. *Remote Sensing*, 9(10), 1022. <https://doi.org/10.3390/RS9101022>
- Schalles, J. F. (2006). 1.3. Optical remote sensing techniques to estimate phytoplankton chlorophyll a concentrations in coastal. In *Remote Sensing of Aquatic Coastal Ecosystem Processes* (pp. 27–79). Springer, Dordrecht. [https://doi.org/10.1007/1-4020-3968-9\\_3](https://doi.org/10.1007/1-4020-3968-9_3)
- Sellner, K. G. (1997). Physiology, ecology, and toxic properties of marine cyanobacteria blooms. *Limnology and Oceanography*, 42(5part2), 1089–1104. [https://doi.org/10.4319/lo.1997.42.5\\_part\\_2.1089](https://doi.org/10.4319/lo.1997.42.5_part_2.1089)
- Shang, S. L., Dong, Q., Hu, C. M., Lin, G., Li, Y. H., & Shang, S. P. (2014). On the consistency of MODIS chlorophyll  $a$  products in the northern South China Sea. *Biogeosciences*, 11(2), 269–280. <https://doi.org/10.5194/bg-11-269-2014>
- Shanmugam, P., Ahn, Y.-H., & Ram, P. S. (2008). SeaWiFS sensing of hazardous algal blooms and their underlying mechanisms in shelf-slope waters of the Northwest Pacific during summer. *Remote Sensing of Environment*, 112(7), 3248–3270. <https://doi.org/10.1016/j.rse.2008.04.002>
- Smayda, T. J. (1997). Harmful algal blooms: Their ecophysiology and general relevance to phytoplankton blooms in the sea. *Limnology and Oceanography*, 42(5part2), 1137–1153. [https://doi.org/10.4319/lo.1997.42.5\\_part\\_2.1137](https://doi.org/10.4319/lo.1997.42.5_part_2.1137)
- Subba Rao, D. V., Al-Yamani, F., & Nageswara Rao, C. V. (1999). Eolian dust affects phytoplankton in the waters off Kuwait, the Arabian Gulf. *Naturwissenschaften*, 86(11), 525–529. <https://doi.org/10.1007/s001140050667>
- Toming, K., Kutser, T., Laas, A., Sepp, M., Paavel, B., & N?ges, T. (2016). First experiences in mapping lakewater quality parameters with sentinel-2 MSI imagery. *Remote Sensing*, 8(8), 1–14. <https://doi.org/10.3390/rs8080640>
- Vanhellemont, Q., & Ruddick, K. (2014). Turbid wakes associated with offshore wind turbines observed with Landsat 8. *Remote Sensing of Environment*, 145, 105–115. <https://doi.org/10.1016/j.rse.2014.01.009>
- Vanhellemont, Q., & Ruddick, K. (2015). Advantages of high quality SWIR bands for ocean colour processing: Examples from Landsat-8. *Remote Sensing of Environment*, 161, 89–106. <https://doi.org/10.1016/j.rse.2015.02.007>
- Vanhellemont, Q., & Ruddick, K. (2016a). Acolite for Sentinel-2: Aquatic Applications of Msi Imagery. *European Space Agency, (Special Publication) ESA SP, SP-740*(May), 9–13.
- Vanhellemont, Q., & Ruddick, K. (2016b). ACOLITE processing for Sentinel-2 and Landsat-8: atmospheric correction and aquatic applications.
- Vanhellemont, Q., Vanderzande, D., & Ruddick, K. (2016). Processing and Performance of Sentinel-2 for Aquatic Applications, (November).

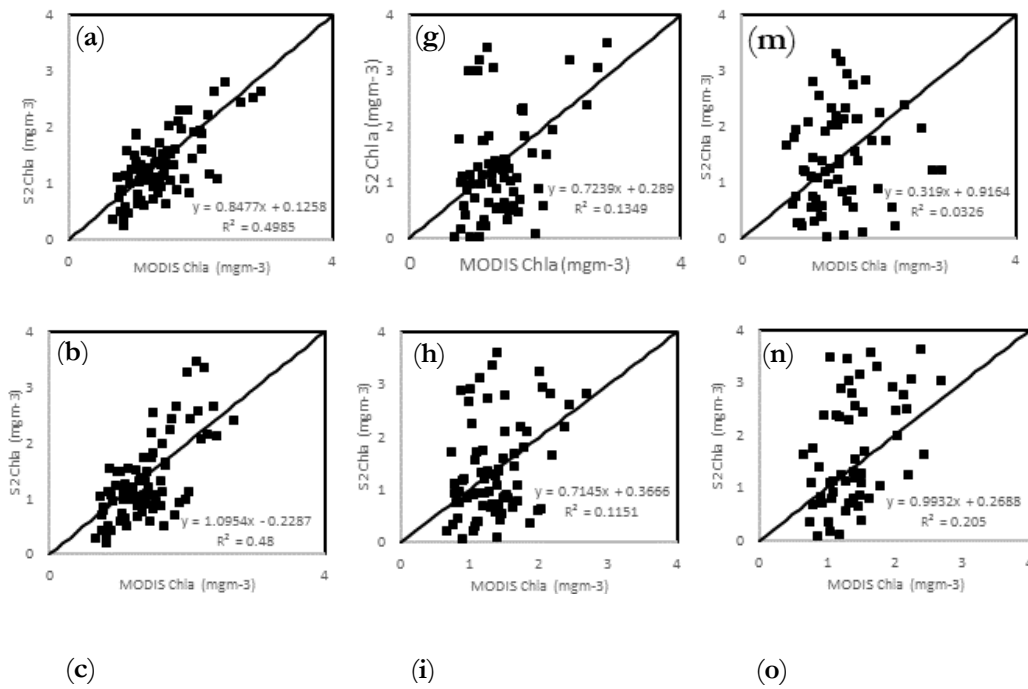
- Villacorte, L. O., Tabatabai, S. A. A., Anderson, D. M., Amy, G. L., Schippers, J. C., & Kennedy, M. D. (2015). Seawater reverse osmosis desalination and (harmful) algal blooms. *Desalination*, *360*, 61–80. <https://doi.org/10.1016/j.desal.2015.01.007>
- Villacorte, L., Tabatabai, S., Dhakal, N., Amy, G., Schippers, J., & Kennedy, M. D. (2014). Algal blooms: an emerging threat to seawater reverse osmosis desalination. *Desalination and Water Treatment*, (March 2015), 1–11. <https://doi.org/10.1080/19443994.2014.940649>
- Wang, Q., Shi, W., Atkinson, P. M., & Zhao, Y. (2015). Downscaling MODIS images with area-to-point regression kriging. *Remote Sensing of Environment*, *166*, 191–204. <https://doi.org/10.1016/j.rse.2015.06.003>
- World Health Organization (WHO). (2003). *Guidelines for safe recreational water environments, Volume 1: coastal and fresh waters*. Geneva: World Health Organization.
- Zhang, F., Li, J., Shen, Q., Zhang, B., Wu, C., Wu, Y., ... Lu, Z. (2015). Algorithms and Schemes for Chlorophyll *Estimation by Remote Sensing and Optical Classification for Turbid Lake Taihu, China*. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(1), 350–364. <https://doi.org/10.1109/JSTARS.2014.2333540>
- Zhao, J., Temimi, M., Al Azhar, M., & Ghedira, H. (2015). Satellite-Based Tracking of Oil Pollution in the Arabian Gulf and the Sea of Oman. *Canadian Journal of Remote Sensing*, *41*(2), 113–125. <https://doi.org/10.1080/07038992.2015.1042543>
- Zhao, J., Temimi, M., & Ghedira, H. (2015). Characterization of harmful algal blooms (HABs) in the Arabian Gulf and the Sea of Oman using MERIS fluorescence data. *ISPRS Journal of Photogrammetry and Remote Sensing*, *101*, 125–136. <https://doi.org/10.1016/j.isprsjprs.2014.12.010>
- Zhao, J., Temimi, M., Kitbi, S. Al, & Mezhoud, N. (2016). Monitoring HABs in the shallow Arabian Gulf using a qualitative satellite-based index. *International Journal of Remote Sensing*, *37*(8), 1937–1954. <https://doi.org/10.1080/01431161.2016.1165886>
- Zhao, W. J., Wang, G. Q., Cao, W. X., Cui, T. W., Wang, G. F., Ling, J. F., ... Hu, S. B. (2014). Assessment of SeaWiFS, MODIS, and MERIS ocean colour products in the South China Sea. *International Journal of Remote Sensing*, *35*(11–12), 4252–4274. <https://doi.org/10.1080/01431161.2014.916044>
- Zolfaghari, K., & Duguay, C. (2016). Estimation of Water Quality Parameters in Lake Erie from MERIS Using Linear Mixed Effect Models. *Remote Sensing*, *8*(6), 473. <https://doi.org/10.3390/rs8060473>

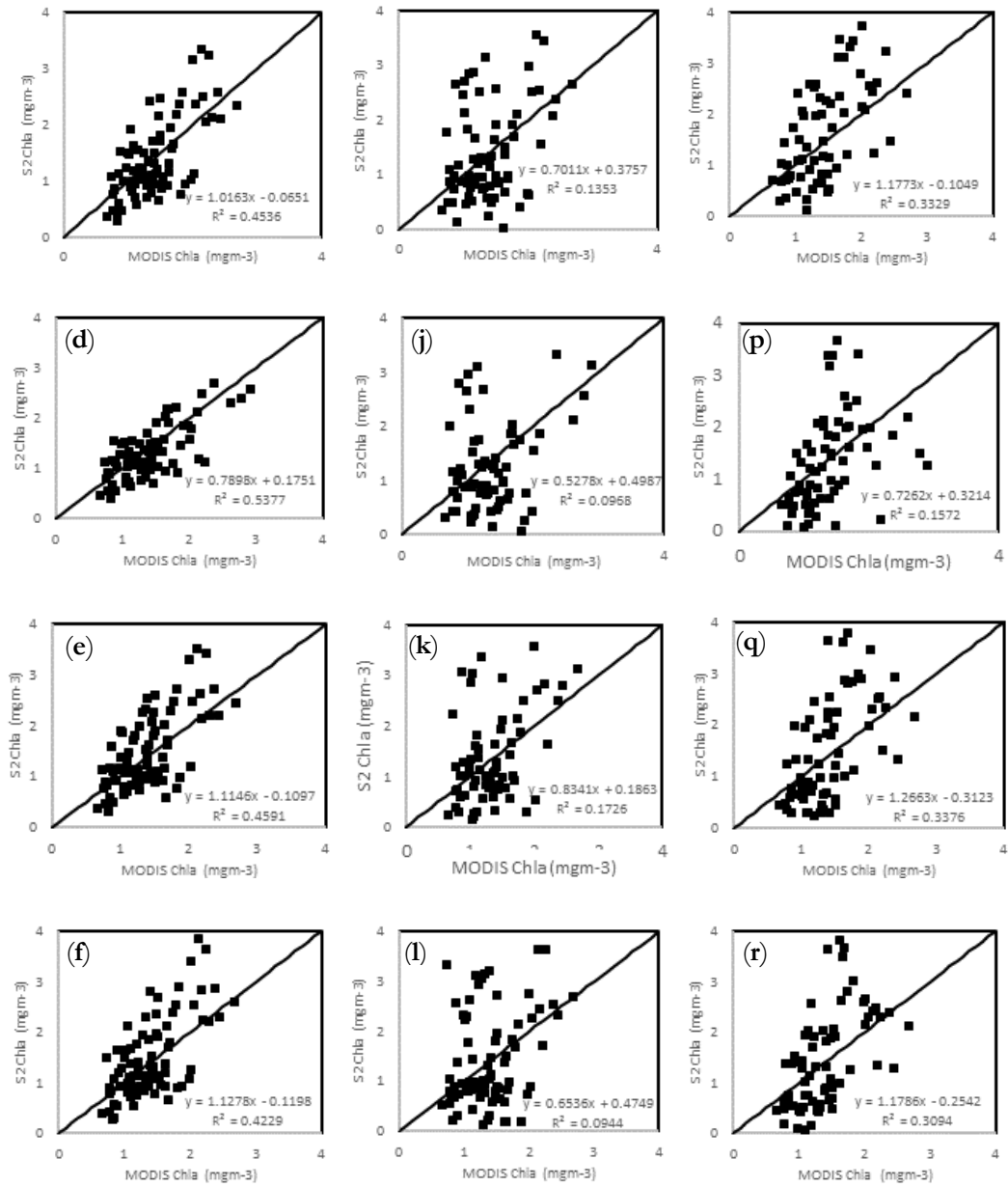
# APPENDIX

Annex A: Scatter plots of Sentinel-2 derived RI versus Rrs (443nm)



Annex B: Scatterplots of Sentinel-2 derived chlorophyll-a by RI, MCI and NDCI against MODIS L2 chlorophyll-a data: a - f by RI model; g - l by MCI model and m - r by NDCI model; a, g, m for 3 x 3; b, h, n for 5 x 5; c, i, o for 7 x 7; d, j, p for 9 x 9; e, k, q for 11 x 11 and f, l, r for 99 x 99 moving window size with Chl-a < 3mg/m<sup>3</sup>.





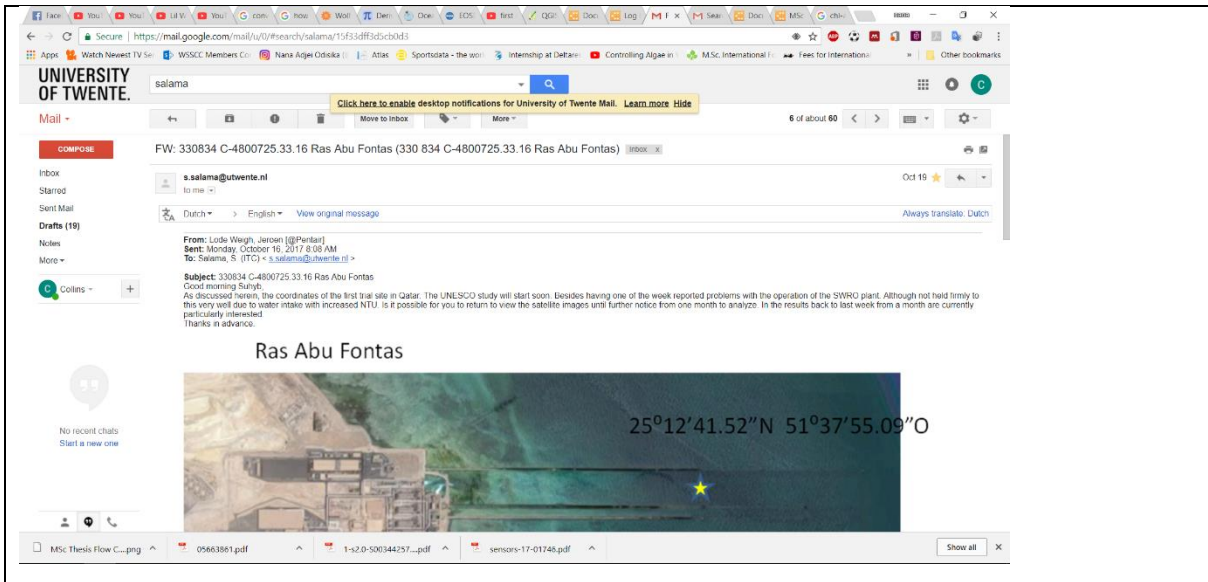
Annex C Statistical summary of scatter plots of Sentinel-2 derived chlorophyll-a concentration against MODIS L2 chlorophyll-a measurement for the validation datasets when Chl-a < 3 mg/m<sup>3</sup>

Window size	Model					
	NDCI		MCI		RI	
	RMSE (mg/m <sup>3</sup> )	MAPE (%)	RMSE (mg/m <sup>3</sup> )	MAPE (%)	RMSE (mg/m <sup>3</sup> )	MAPE (%)
3 x 3	0.903	55.24	0.885	55.69	0.423	25.79
5 x 5	0.949	64.07	0.865	52.66	0.525	30.09
7 x 7	0.779	44.76	0.804	47.68	0.505	29.62

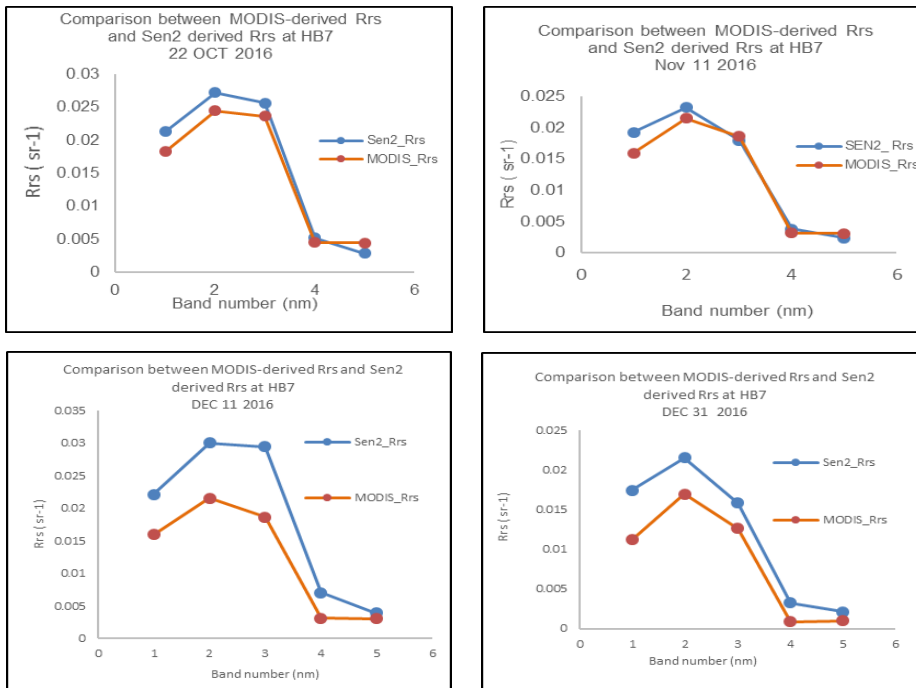
Deriving algal concentration from Sentinel-2 through a downscaling technique: A case near the intake of a desalination plant

9 x 9	0.803	45.41	0.809	48.79	0.385	22.92
11 x 11	0.808	48.28	0.805	49.89	0.544	31.73
99 x 99	0.789	47.85	0.909	53.79	0.595	34.37

Annex D: Email received from project management of possible HAB occurrence due to plant shutdown



Annex E: Comparison between MODIS-derived Rrs and Sen2 derived Rrs at data point





modis	Sen 2
Rrs_443	Rrs_443
Rrs_488	RRS_497
Rrs_645	RRS_560
Rrs_667	RRS_665
Rrs_678	RRS_704

Annex G: Graph of Average residual estimated for aggregated Sentinel-2 MSI and disaggregated MODIS L2 Chl-a maps at 90 m all 12 images

