Remote Sensing Report on Harmful Algal Blooms in Varna Lake and its Bay, and Their Link to Wastewater Spills

By Christina A. Orieschnig, Kristof Horvath

chrstina.orieschnig@icloud.com

Summary

This report provides a summary of the remote sensing analyses carried out in order to track algal blooms in Lake Varna and its bay on the Bulgarian Black Sea coast, linking their dynamics to the occurrence of waste water spills in the lake. To assess the temporal distribution, intensity, and extents of algal blooms, a time series of multispectral optical Sentinel-2 images provided by the European Space Agency was analyzed on the open-access cloud computing platform Google Earth Engine. In addition to the 12 optical bands available by default, six spectral indices were calculated for each image to better detect the extents of algal blooms: the Normalized Difference Chlorophyll Index (NDCI), the Floating Algae Index (FAI), the Adjusted Floating Algae Index (AFAI), the Algal Bloom Detection Index (ABDI), the Normalized Difference Vegetation Index (NDVI), and the Enhanced Vegetation Index (EVI). Subsequently, the extents of algal blooms was determined through a simple thresholding approach, and the average NDCI value, indicating the concentration of chlorophyll in the water, were calculated. In addition, the dynamics of these metrics were statistically compared to in-situ measurements of chlorophyll concentrations, and nutrient concentrations, as well as remotely sensed surface temperatures and precipitation events. Finally, the analysis targeted five locations of aquaculture farms in Varna lake and along its bay to determine whether algal blooms could pose a danger for the fish and mussels produced there. The results show that the dynamics of algal incidence in the area are dominated by regular annual variations, with peaks in spring and late summer. However both 2019 and 2020 show deviations from this pattern in terms of the magnitude and duration of NDCI values, as well as their spatial distribution. This could be attributed to the wastewater spill in the Eastern part of the lake. Furthermore, the results show that algal incidence is principally limited to the lake itself, probably due to the limited, narrow connection between the bake and the lake. Finally, the analysis of aquaculture farms showed that those located in the western part of the lake itself faced regular algal blooms, while those along the coast were relatively unaffected.

1. Introduction

Lake Varna is located at the Bulgarian Black Sea Coast. It has a surface of 21.5 km^2 and is connected to the Black Sea through two narrow channels at its eastern end. There are several popular beaches along the lake's shores, as well as in the bay that lies just downstream of it. In addition, there are several aquaculture farms in the area, both in the lake and along the bay's coastline.

The motivation for this study lies in a wastewater spill that occurred in the eastern part of lake Varna. A pipeline to conduct wastewater through the lake was built starting in 2007. Contrary to planning guidelines, this pipeline was constructed above ground, leaving it vulnerable to mechanical damage. Such damage occurred first in 2015, when a ship had to lower anchor under

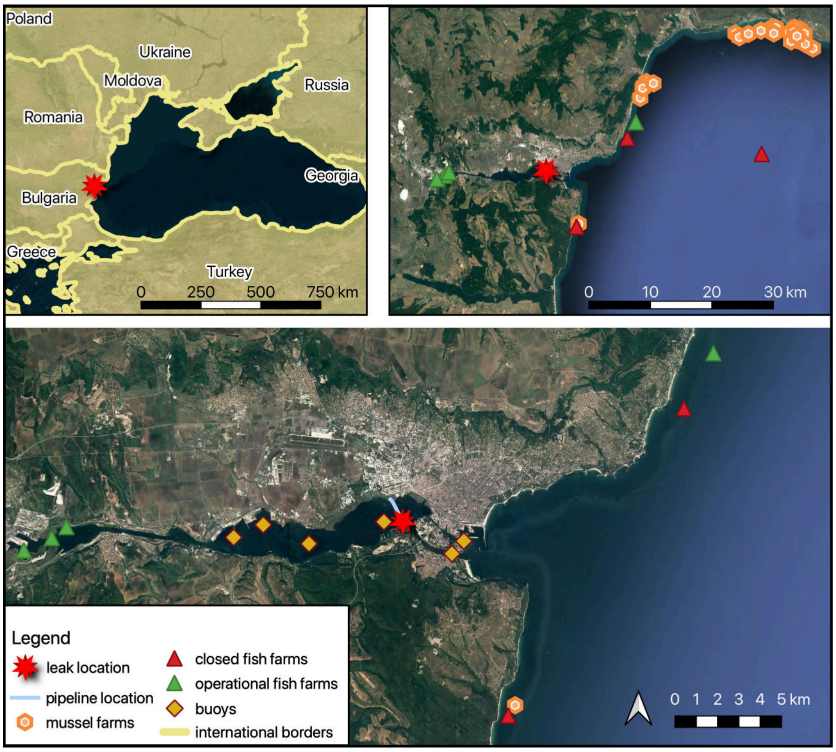


Fig. 1: Location of the area of interest, Lake Varna, on the Bulgarian Black Sea coast.

unforeseen circumstances. A similar incident occurred four years later. Structural damage leading to wastewater spilling into the bay was reported to the mayor of Varna in August 2019. The question this study seeks to answer is whether this wastewater spill has triggered or aggravated algal blooms in the area - and, if so, to what extent.

To a certain degree, algal blooms are a regular phenomenon in the region, particularly since extensive environmental degradation occurred in the 1970s and 80s. This degradation was due to the alteration of nitrogen, phosphorus and silicon fluxes in the area. Increased concentrations of N and P can be traced back to extensive fertilizer use, resulting in eutrophication (Yunev et al., 2007; Mee et al., 2005). In general, elevated concentrations of nutrients allow for the proliferation of opportunistic species of algae, leading to potentially harmful blooms. These blooms can cause damage either through the secretion of toxic substances by certain species of algae, or through the depletion of oxygen in the water, which can lead to the death of other aquatic species. However, since the turn of the millennium, a trend towards ecosystem recovery has been noted (Klisarova et al., 2019).

In lake Varna, this trend might have been interrupted by the input of nutrients, specifically N and P, from the wastewater spill. Untreated wastewater contains these nutrients in high concentrations and their release into surrounding lake water could lead to the proliferation of various algal species. Since the wastewater pipeline was damaged, bloom events as well as their adverse impacts were reported, such as mass fish death from oxygen deprivation in May 2020 (Novinite.com, 2020).

However, these bloom events have also been reported to occur on a regular basis. For example, widespread fish death was also reported in May 2018, as well as in late August every year from 2011 to 2016 (Novinite.com, 2020). Similarly, Klisarova et al. (2019) reported regular peaks in algal species numbers in the Bulgarian marine areas in May and September between 2014 and 2017, though biomass values were observed to peak in November and February.

The objective of the present study is to assess the impact of waste water spilling from the damaged pipeline at the bottom of lake Varna on the algal bloom dynamics in the lake and its bay. In particular, the impact of a possibly increased incidence of harmful algal blooms on aquaculture farms in the area will be discussed.

2. Methodology

The core of this approach is to derive a time series of algal extents and chlorophyll concentrations in lake Varna and its bay from Sentinel-2 images made available by the European Space Agency. Sentinel-2 images have a spatial resolution of up to 10 m and a return frequency of 5 days or less. They are provided by the Sentinel-2A satellite, launched in 2015, and the Sentinel-2B satellite, launched in 2017. The constellation delivers multispectral optical images with 12 bands recording reflectance values at different wavelengths (cf. Table 2 in the Appendix).

This approach has the advantage of delivering information on the spatial extents, patterns, and distributions of algal presence in all of Lake Varna and its bay. Unlike in-situ sampling, which is punctual and expensive, this enables a broad overview of the dynamics of algal blooms. However, a downside is that it is not possible to determine the exact species of algal present in the water without in-situ bio calibration data (Ansper & Alikas, 2018; Smith & Bernard, 2020).

To derive algal extents and chlorophyll concentrations, this paper makes use of several spectral indices. These are calculated by performing arithmetic operations on the different spectral bands of satellite images, which highlights reflectance differences resulting from biophysical processes on the earth's surface. One of the most illustrative examples of this application is the Normalized Difference Vegetation Index (NDVI, Rouse J.W. et al. (1974), Equ. 1).

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{1}$$

This index exploits a discrepancy of the reflectance values in the red (Red) and near infra-red (NIR) spectrum depending on the presence of vegetation. If healthy vegetation is present, far more NIR than Red light will be reflected, since the former is not needed for photosynthesis. Consequently, the more healthy vegetation is present, the higher the NDVI will be.

In addition to the NDVI, this paper uses the Enhanced Vegetation Index (EVI, Huete et al. (2002), Equ. 2), the Floating Algae Index (FAI, Hu (2009), Equ. 3), the Adjusted Floating Algae Index (AFAI, Fang et al. (2018), Equ. 4), the Algal Bloom Detection Index (ABDI, Cao et al. (2021), Equ. 5), and the Normalized Difference Chlorophyll Index (NDCI, Mishra & Mishra (2012), Equ. 6). The latter index can serve as a proxy for chlorophyll concentrations in the water (see Table 3).

$$EVI = 2.5 * \frac{NIR - Red}{NIR + 6 * Red - 7.5 * Blue + 1} \tag{2}$$

$$FAI = NIR - [Red + (SWIR - Red) * \frac{\lambda NIR - \lambda Red}{\lambda SWIR - \lambda Red}] \tag{3}$$

$$AFAI = NIR - Red + (SWIR - Red) * 0.5 \tag{4}$$

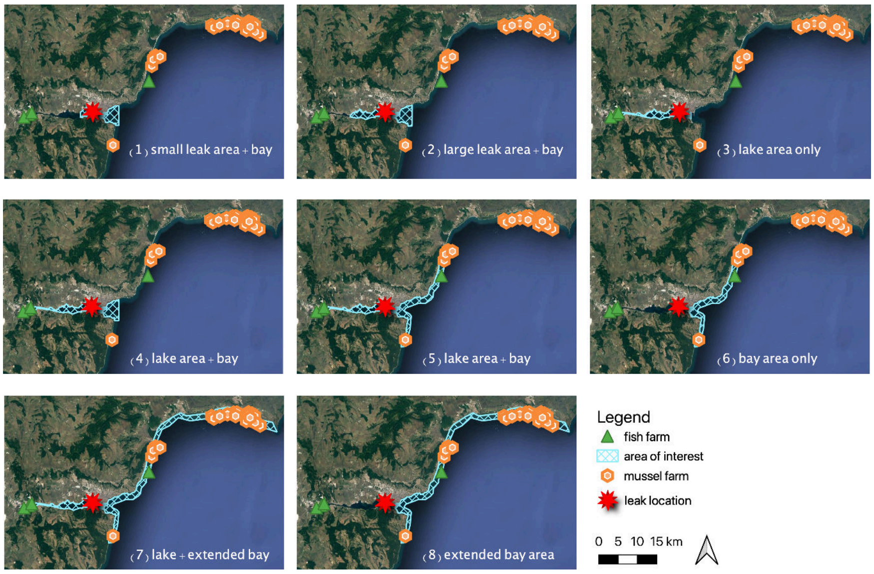


Fig. 2: Extents of areas of interest studied in this analysis.

$$ABDI = [RE2 - Red - (RE4 - Red) * \frac{\lambda RE2 - \lambda Red}{\lambda RE4 - \lambda Red}] - [Red - 0.5 * Green] \tag{5}$$

$$NDCI = \frac{NIR - Red}{NIR + Red} \tag{6}$$

As a first step in the analysis, all available Sentinel-2 images in the area of interest were filtered for cloud cover, and clouds were masked. Only those showing 90% or more of the area as cloud-free were retained. Out of these, daily mosaics were constructed. Overall, 789 images that completely or partially covered the study area were available between 2015-07-01 and 27/06/2021. Of these, 212 images were retained for the analysis and processed, which makes for an average return frequency of 10 days. Subsequently, each of the indices was calculated for every image. Then, the average value of the NDCI over the area of interest was calculated for each image in the time series, to determine an approximate average chlorophyll-a concentration. To determine algal presence, a threshold value was set for each individual index. If a pixel showed an index value above that threshold, it was classified as having algal presence. To test the sensitivity of this approach, it was repeated once with a threshold of 0 and once with a threshold of 0.01. Next, the total area of all pixels classified as exhibiting an algal presence was calculated.

This process was repeated for a total of eight overlapping areas of interest, allowing a comparison of algal dynamics at different scales, both in the lake and the bay. The areas are centered on the wastewater spill and progressively expand to include active fish farms and mussel farms (see Figure 1).

Subsequently, the time series of algal extents and NDCI values was statistically examined for trends and change points using the modified Mann-Kendall test (Hamed & Ramachandra Rao, 1998) and Buishand's range test (Buishand, 1982). These tests were performed on the average values of chlorophyll concentrations and extents of each month.

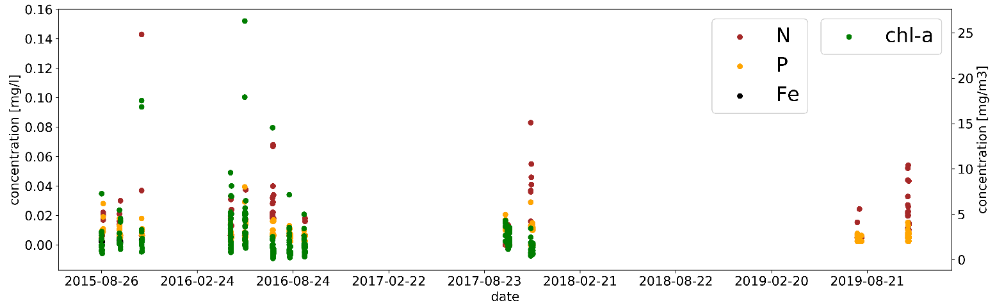


Fig. 3: Concentrations of nutrients and chlorophyll measurements at the bays in Varna bay, from SeaDataNet.

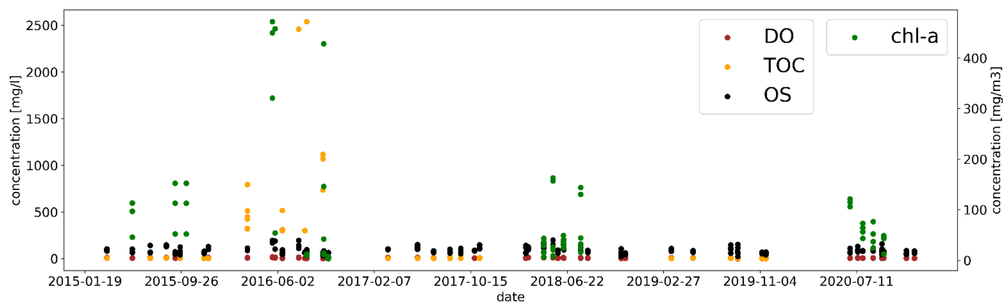


Fig. 4: Concentrations of dissolved oxygen (DO), total organic carbon (TOC), oxygen saturation (OS) and chlorophyll in the lake.

In addition, local measurements of nutrient and chlorophyll-a concentrations were used to validate the remotely sensed algal extents and chlorophyll-a concentrations. Some of this data was made available by the Black Sea Basin Directorate via SeaDataNet (seadatanet.org). Furthermore, measurements from buoys in lake Varna and Varna bay were directly extracted from the directories of the Bulgarian Oceanographic Institute. The buoys used to measure the parameters are shown in Figure 1. In particular, five dates for which both chlorophyll-a in-situ measurements and Sentinel-2 images are available are used for comparisons: 2016-04-28, 2016-07-17, 2016-08-16, 2017-10-10, 2018-06-12, and 2020-09-10 (see Figures 3 and 4). The concentrations of nutrients and chlorophyll-a measured at buoy 16 (BG2PR00155MS016) are shown separately in Figure 5 because this buoy is the closest in the lake to the leak location.

Furthermore, trends in water temperatures and precipitation events were analysed in parallel to the algal extents and NDCI-derived chlorophyll-a concentrations. This was done to exclude the possibility that sudden spikes in algal blooms are due solely to temperature aberrations or high-intensity precipitation events leading to increased fluxes of N and P to the lake and bay areas. Since in-situ data was not readily available, remotely sensed daily measurements from CHIRPS (Funk et al., 2015) and ERA-5 Copernicus Climate Change Service (C3S) (2017) were used.

Finally, once the validity of the approach had been assessed for the entire area of interest, it was zoomed in onto active aquaculture farms in the region. These are shown in Figure **??**. As this figure illustrates, there are five clusters of active farms in the area - one located in the western

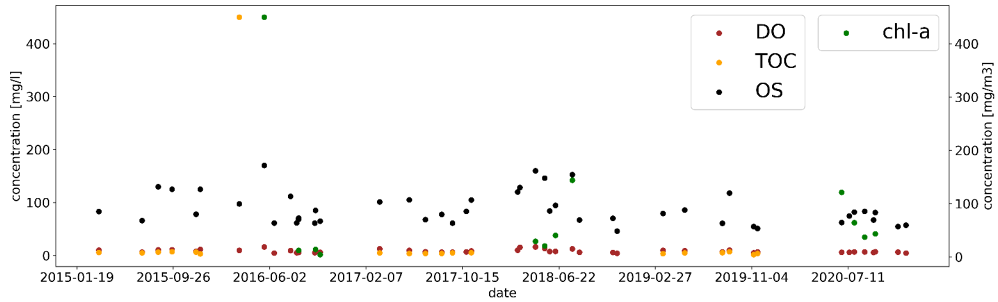


Fig. 5: Concentrations of nutrients and chlorophyll in the lake at buoy 16, closest to the leak location (870 m).

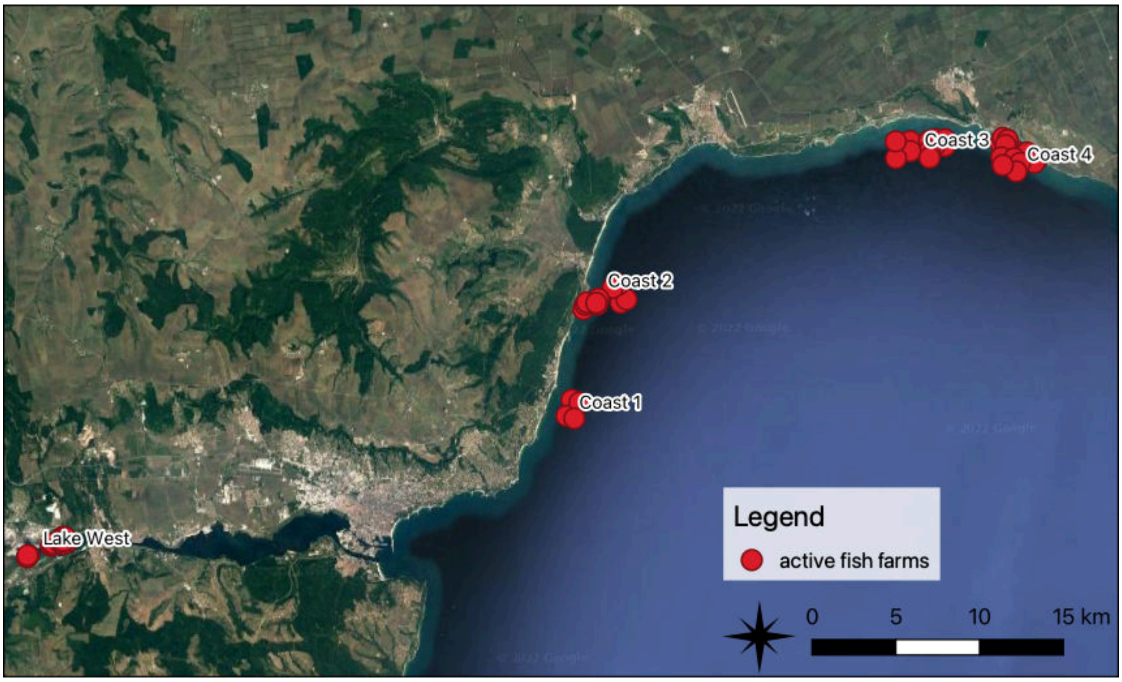


Fig. 6: Clusters of aquaculture farms in the study area.

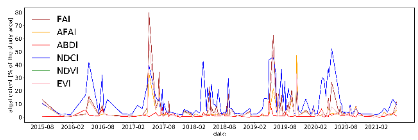
part of the lake and four along the coastline of Varna bay. For the analysis, five polygons were drawn around these clusters and the algal indices and NDCI calculated for the entire time series for each location.

3. Results

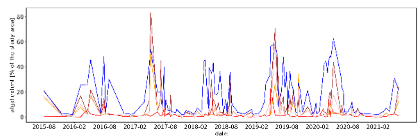
3·1. *Results of Algal Detection, Temporal, and Spatial Dynamics*

The results show parallel dynamics of chlorophyll-a concentrations and algal extents throughout the entire study area, with peaks each year in late winter and spring, as well as late summer. However, there is variation in the magnitude and density of the peaks. As Figures 7 and 8 show, the incidence of algae is much higher in the study areas that include the lake (a, b, c, d, e, g) than in those that encompass only the fully marine part of the region of interest (f, h). Overall, the study area that stretches over the entirety of the lake (c) registers the largest algal extent values. Interestingly, it also registers higher values than the study areas that are centered specifically on the leak area (a and b). Generally, the lowest algal extent values are recorded for the bay-only areas.

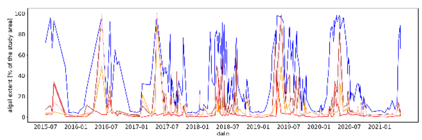
In terms of comparing the sensitivity of the threshold used to detect algal presence, it can be noted that it has relatively little impact for the lake-focused areas. For those that encompass mostly maritime areas, though, it can be noted that at the 0.00 threshold, the ABDI detects much larger algal extents in periods during which other indices fail to register algal presence. Cao et al. (2021) used the Otsu algorithm to determine an ideal threshold for the newly developed ABDI. However, this approach was beyond the scope of this analysis.



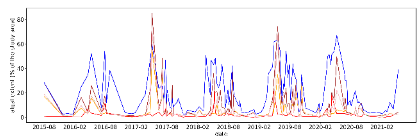
(a) over the small study area surrounding the leak



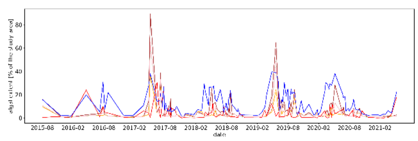
(b) over the big study area surrounding the leak



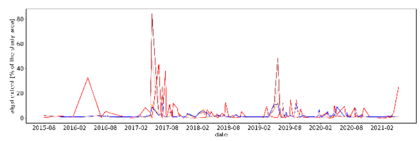
(c) over the entirety of the lake area



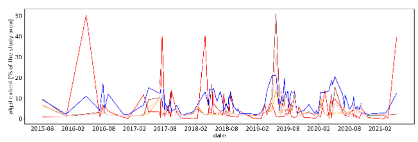
(d) over the lake area and its bay



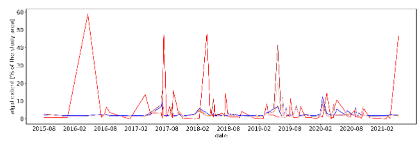
(e) over the lake and the extended bay area



(f) over the extended bay area only

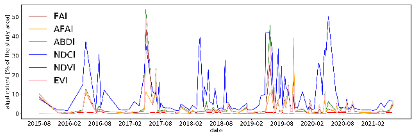


(g) over the lake and full bay area

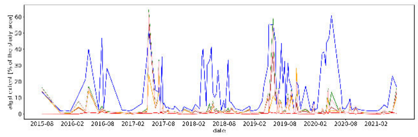


(h) over the full bay area only

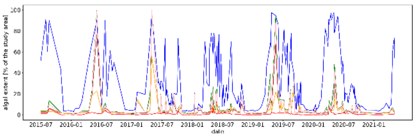
Fig. 7: Dynamics of algal extent according to all indices, 0.00-threshold



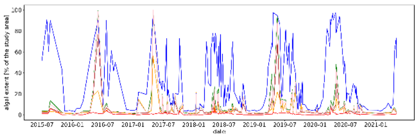
(a) over the small study area surrounding the leak



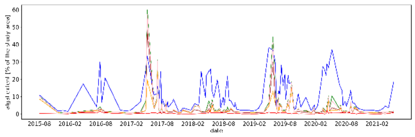
(b) over the big study area surrounding the leak



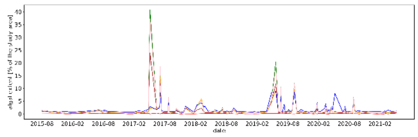
(c) over the entirety of the lake area



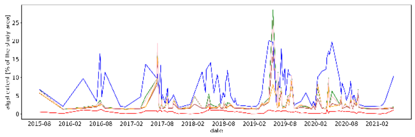
(d) over the lake area and its bay



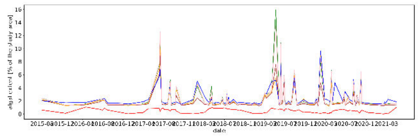
(e) over the lake and the extended bay area



(f) over the extended bay area only



(g) over the lake and full bay area

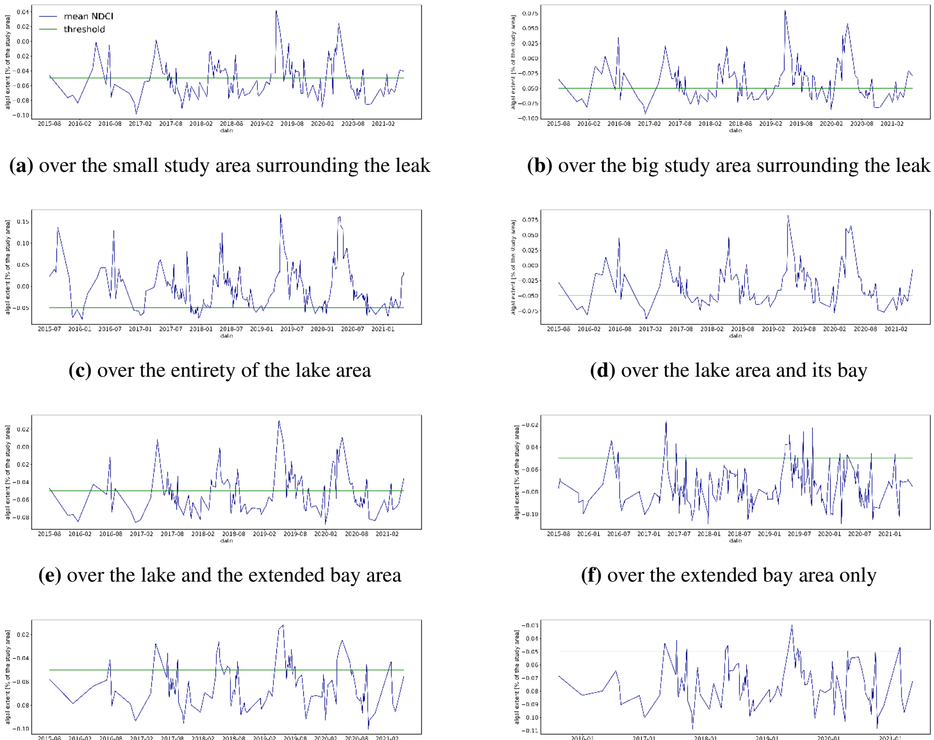


(h) over the full bay area only

Fig. 8: Dynamics of algal extent according to all indices, 0.01-threshold

The mean NDCI values, shown in Figure 9, exhibit similar tendencies. The threshold value of -0.05 (which corresponds to a chlorophyll-a concentration of ca. 10 mg/m^3[1]) is exceeded almost continuously when looking at the entirety of lake Varna (c), but only occasionally when the study area includes the bay as well (g).

[1] This is at the threshold between meso- and eutrophic lake status (Anthwal et al., 2018)



(a) over the small study area surrounding the leak

(b) over the big study area surrounding the leak

(c) over the entirety of the lake area

(d) over the lake area and its bay

(e) over the lake and the extended bay area

(f) over the extended bay area only

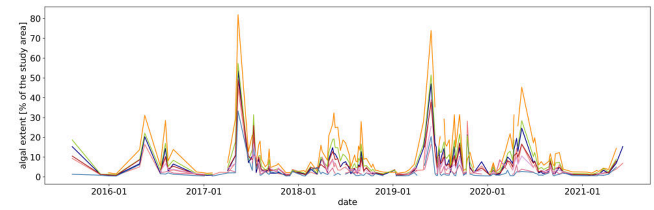
(g) over the lake and full bay area

(h) over the full bay area only

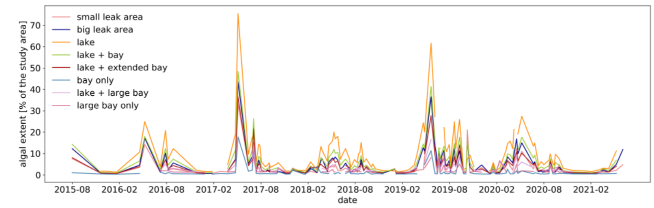
Fig. 9: Mean NDCI values per date over the areas of interest

The temporal dynamics of the algal extents detected at threshold 0.00 and 0.01, as well as the mean NDCI values can be seen in Figure 10. Averaged out over the study areas, two main peaks in algal extents become visible: one in May 2017, and one in May 2019. Smaller peaks are registered during the summer months of the other years as well, with the spring peak of 2020 being particularly broad. Looking at the NDCI, however, larger and more prolonged peaks are visible in 2019 and 2020 than in the previous years. For all parameters, the lake-only study area exhibits the highest values. Consequently, the rest of the analysis focuses on this area.

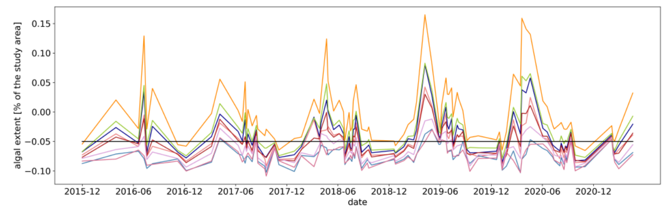
Statistically, a significant increase in NDCI value in January and February can be found, while the average September value decreased. However, no other statistically relevant trends or break points could be identified. To a certain extent, this could be due to the relatively short duration of the time series of this analysis.



(a) Mean algal extents (0.00 threshold) over the study areas.



(b) Mean algal extents (0.01 threshold) over the study areas.



(c) Mean NDCI values over the study areas.

Fig. 10: Dynamics of algal extents and NDCI values across the study areas.

Finally, Figure 11 shows the spatial distribution of NDCI values across Lake Varna on the days of the maximum average value measured. Figure 12 illustrates the FAI values for the same date. Respectively, these dates are 2016-08-06, 2017-10-18, 2018-05-18, 2019-05-03, and 2020-04-25. The images show that in all years except 2019, the intensity of the algal incidence seems to be concentrated in the western part of the lake. In 2019, the center seems to have shifted eastward, towards the leak location. The figures also illustrate the severity and intensity of the algal blooms in 2019 and 2020. In addition, they highlight a possible shortcoming of indices such as the FAI. Here, the highest values are registered in the immediate vicinity of the shore. While this may really indicate the very concentrated presence of algae, it is equally likely that aquatic vegetation such as seagrass is interfering with the detection process.

To a certain extent, this is concordant with in-situ measurements in Lake Varna. Buoy 13, located at the western end of the lake, registers higher chlorophyll values on average than buoy 16, located at the eastern end.

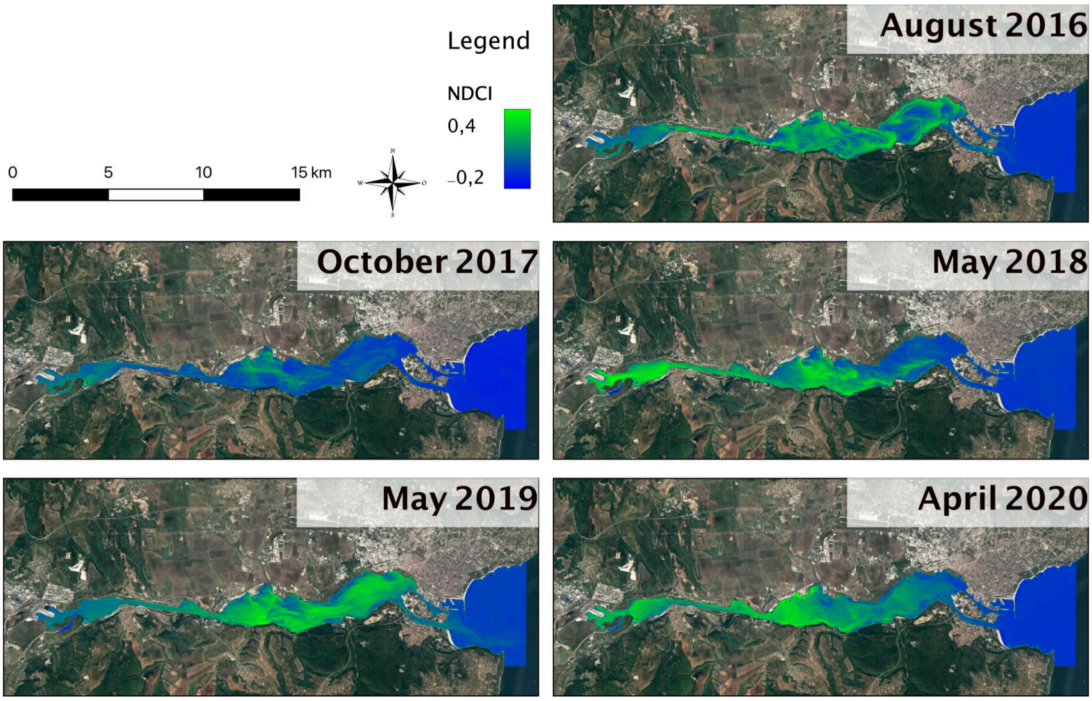


Fig. 11: Distribution of NDCI values throughout lake Varna and its bay at peak levels 2016-2020.

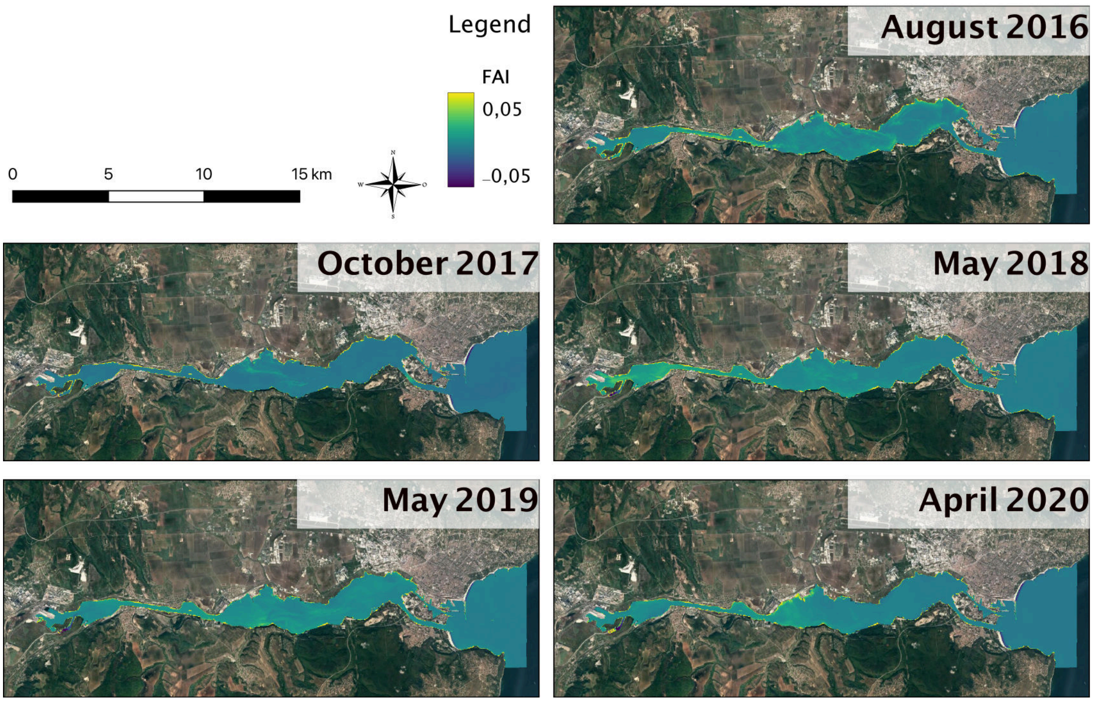


Fig. 12: FAI values for dates of peak NDVI average 2016-2020.

3·2. *Comparison with In-situ Data*

A comparison with in situ data reveals that the dynamics of algal extents and chlorophyll concentrations in the lake are not directly correlated with temperature or rainfall. Figure 14 shows that high-intensity precipitation events are not followed by increased algal densities - as they would be if algal blooms were due to increased nutrient input due to erosion of surrounding agricultural areas, for example. This is underlined by Figure 13, which plots precipitation against the deviation of the average NDCI values across the lake from the mean of -0.004, which corresponds to approximately 16 mg/m^3. It shows no correlative effects.

Similarly, no direct relation between temperature anomalies and algal growth in Lake Varna could be observed. While patterns emerge on an annual basis (see Figure 16) due to the seasonal variations in algal growth, deviations from the mean temperature are not correlated with deviations from the mean NDCI value (see Figure 15), as would be the case if increased algal growth were largely due to the effects of global warming, for example.

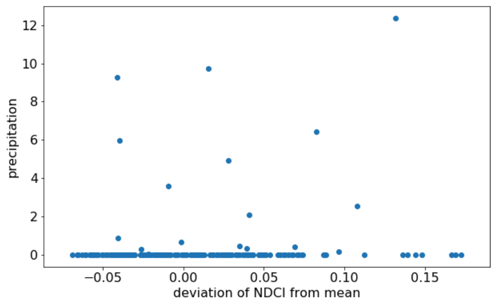


Fig. 13: Precipitation values, plotted against the deviation of NDCI values from the mean.

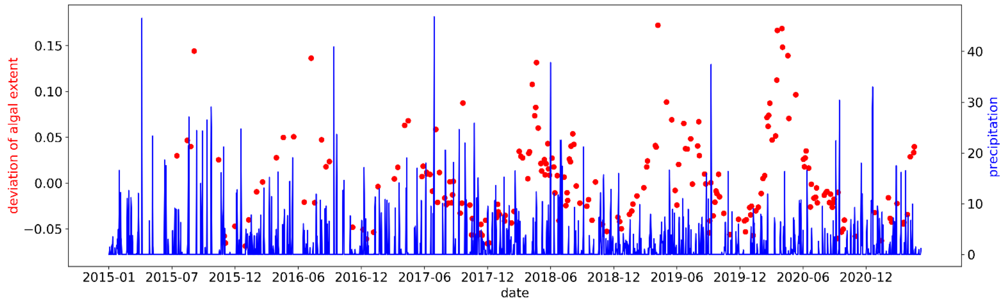


Fig. 14: Comparison of temporal dynamics of rainfall and the deviation of NDCI values from the mean.

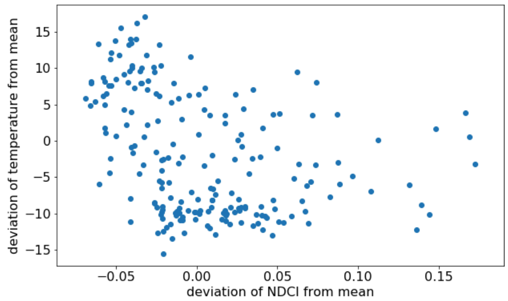


Fig. 15: Temperature values, plotted against average NDCI values.

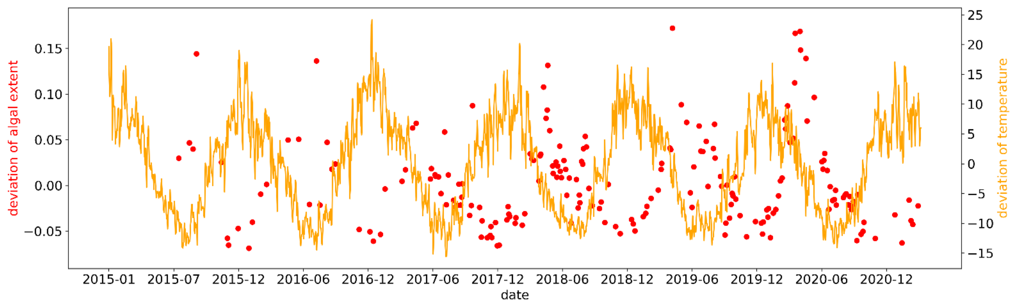


Fig. 16: Comparison of temporal dynamics of temperatures and the deviation of NDCI values from the mean.

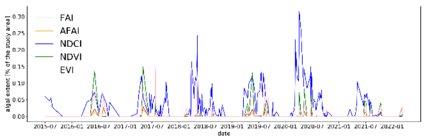
In terms of comparing remotely sensed estimations of chlorophyll-a concentrations and in-situ measurements, there is a broad correspondence, though satellite measurements tend to overestimate the concentrations slightly - by 3.20 mg/m^3 on average. This is equally reflected by the values given in literature. It can be noted that the agreement between estimated and measured concentrations is better in the lake area than in the bay, with an overestimation of just 0.88 mg/m^3, rather than 4.19 mg/m^3, as shown in Table **??**).

Date	Location	NDCI	Chlorophyll [mg/m^3] Estimated	Chlorophyll [mg/m^3] Measured	Difference	AVERAGE
2017-10-10	bay	-0,14	4,17	2,36	1,81	
2017-10-10	bay	-0,12	5,83	3,46	2,37	
2016-04-28	bay	-0,11	6,67	4,52	2,15	
2017-10-10	bay	-0,11	6,67	3,53	3,14	4,19
2016-04-28	bay	-0,09	8,35	7,05	1,30	
2016-07-17	bay	-0,07	10,05	1,29	8,76	
2017-10-10	bay	-0,05	11,75	1,93	9,82	
2016-08-16	lake	-0,08	9,2	9,22	-0,02	
2018-06-12	lake	0,01	16,9	16,23	0,67	0,88
2020-09-19	lake	-0,09	14,3	12,31	1,99	
AVERAGE		3,20				

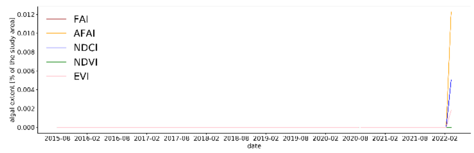
Table 1: Comparison of chlorophyll concentrations, estimated remotely and measured in-situ

3·3. *Assessment of Algal Concentrations at Active Fish Farms in Varna Lake and Bay*

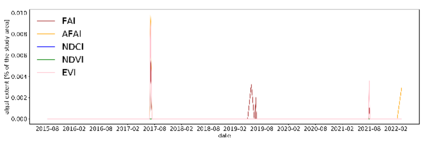
The results of the analysis at the sites of the clusters of active aquaculture farms are shown in Figure 17. Only values above 0 are shown. The figure illustrates that the cluster in the western part of the lake is heavily affected by algal proliferation and high NDCI values, with a particularly significant spike in 2020. In contrast, the coastal stations all have a few disparate peaks, though the total values are much lower (note the difference in the axis scaling).



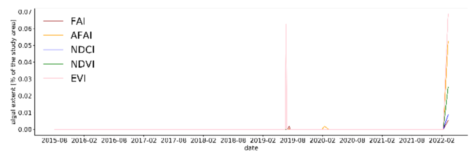
(a) Results for the aquaculture farms in the western part of the lake.



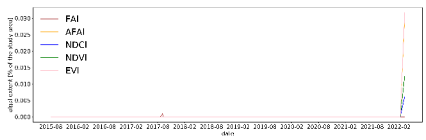
(b) Results for the aquaculture farms in the first cluster along the coast.



(c) Results for the aquaculture farms in the second cluster along the coast.



(d) Results for the aquaculture farms in the third cluster along the coast.



(e) Results for the aquaculture farms in the fourth cluster along the coast.

Fig. 17: Dynamics of algal indices and NDCI values at the active aquaculture farms.

4. Discussion

4·1. *Strengths and Weaknesses of the Methodology*

All in all, the approach outlined above seems to deliver results with satisfactory accuracy in terms of chlorophyll concentrations using the NDCI, especially in the lake area. However, at this point the limited nature of the data available for validation needs to be highlighted. While some data is available via the EU and the web directories of the Bulgarian Institute for Oceanography, these are by no means temporally consistent, and scattered across buoys located at various places in the bay and the lake. With a consistent time series of chlorophyll-a measurements, or even proxies such as oxygen saturation, it would be possible to perform a local calibration of the matrix used to convert NDCI values into chlorophyll-a concentrations. Since only a very limited number of dates during the period of interest offer both in-situ measurements and remote sensing data, such a calibration is not possible as of now. For this reason, the correspondence table given in the original publication by Mishra & Mishra (2012) was used.

In addition, a certain discontinuity between the indices used in this study must be noted. For the ABDI, an issue is the lack of availability of standardized thresholding values. Since the index was only published in March 2021 (see Cao et al. (2021)), it has not yet seen widespread application. With a limited number of use cases, and approaches such as Otsu thresholding challenging to implement, its usefulness in the study area remains in doubt. In contrast, established indices such as the FAI and AFAI are in sync with each other, indicating a certain amount of inter-reliability. However, due to the possibility of interference by natural vegetation in areas close to the shore, it appears that the NDCI provides the best spatial overview of the distribution of algal presence throughout Lake Varna and its bay.

An additional challenge of the study area deserves mention at this point. Several of the areas of interest discussed in this work present a mixture of marine and freshwater surfaces. For example, while Lake Varna's salinity is certainly higher than that of most inland freshwater lakes, it also does not exhibit the same wave or tidal activity as its bay. All of these factors render the area under investigation, once it encompasses both the lake and the bay, heterogeneous. This could be one possible explanation for the discrepancy between the agreement values between estimated and measured chlorophyll concentrations between the lake and its bay.

4·2. *Dynamics of Algal Blooms in the Area*

As for the dynamics of algal blooms and their relation to the wastewater spill in Lake Varna, it can be noted that the the predominant dynamics are due to regular annual variations, with a primary bloom occurring in spring and a secondary one in late summer to early autumn.

However, in 2019 and 2020, the annual dynamics of algal incidence exhibited several characteristics that might be attributed to the increased nutrient input from the spill. In terms of average NDCI values (see Figure 9), these years exhibited higher peaks and longer durations of elevated chlorophyll-a concentrations than previous years. In particular, the autumn peak of 2019 is pronounced and prolonged.The spring peak of 2019 is the highest in the series in terms of NDCI. While the mayor of Varna was only notified of the leak in August, it might well be that the damage was done earlier on, influencing this increased magnitude of the spring peak. In addition, it might be due to the spill at the Eastern end of take that the spatial distribution of algal incidence shifted that year. As Figure 11 shows, 2019 is the only year in which the peak NDCI values center on the Eastern part of the lake. It must also be noted, however, that it is uncertain to which extent currents in the lake - especially due to tidal influence - distribute the nutrients released in the Eastern section of the lake to other regions.

Lake Varna's connection to the sea in general adds an additional layer of complexity to this analysis. In general, the chlorophyll-a concentration and algal incidence is much higher in the lake than in the bay. This was also noted by Klisarova et al. (2019), who found the bay to have a "very good" ecological status, while the lake was classed as "very bad" according to the evaluation framework of the European Union's Water Framework Directive. This discrepancy is further apparent from the fact that the inclusion of maritime sections in the study area automatically lowers average NDCI values across all analyses. Part of the reason for these dynamics are certainly the limited pathways for hydraulic exchange between the lake and the bay. It is only through two relatively narrow channels (100 and 150 m wide respectively) at the lake's Eastern end that a connection is established with the sea. Conditions of internal currents in the lake as well as tidal effects are unclear. However, an exchange of sea water with low nutrient concentrations between the Black sea and the eastern part of the lake could also help account for the usual concentration of algal incidence in the West.

Overall, for aquaculture activities in this area, this means that the fish and mussel farms that are principally affected are located in the Western part of Lake Varna. Here, chlorophyll-a concentrations are regularly higher than in other areas of the lake, and its bay. According to the current analysis, farms located along the coast have seen only relatively minor fluctuations in algal incidence. The localised results reveal singular peaks, but few systematic variations in the marine stations. Partially, this may be due to the fact that the marine environment provides a more difficult application for the indices, for instance due to waves. However, it is more likely that the tidal influence as well as the currents along the coast dilute any nutrient input that passes through the narrow strait that connects the lake to the sea. Since the clusters of aquaculture farms are located between 14 and 45 km (aerial distance) upstream of the outlet of the lake, this dilution effect was to be expected to a certain extent.

A final point to be made is that the results indicate that neither high-intensity precipitation causing increased nutrient input, nor temperature anomalies can be linked to the variations in the pattern of algal incidence in the region in 2019 and 2020.

5. Conclusion and Outlook for Further Analyses

Overall, this analysis showed that the major defining factor of algal incidence in lake Varna are regular annual variations, with peaks in spring and late summer. However, the years 2019 and 2020 showed several pattern-breaking characteristics, such as longer durations of high NDCI values, higher NDCI peaks, and a concentration of algal incidence in the Eastern part of the lake, which may be linked to the wastewater spill in the area. It further illustrated that algal incidence are principally limited to the lake itself, which is only connected to the sea by two narrow straits.

As for the method itself, evaluation with the scant in-situ data indicates that the NDCI provides an acceptable proxy for in-situ chlorophyll-a concentrations, with an average deviation of 3 mg/m^3. However, additional local measurements could provide a firmer basis for both calibration and validation.

Furthermore, analyses focusing on algal incidence around aquaculture farms in the bay area showed that there is relatively little influence of nutrient input from the lake. This is likely due to the narrow nature of the strait that connects the lake to the sea, as well as the diluting influence of currents and tides.

Code and Data Availability

The Earth Engine analyses are freely available here: https://code.earthengine.google.com/4f79bca5d5728046b84b9c920061cff1.

Additional files, as well as the Python scripts used for the statistical analysis can be found here: https://drive.google.com/drive/folders/1bBsJbRbyO3QpW4llHKMRUlE8NS0eQKmb?usp=sharing.

REFERENCES

ANSPER, A. & ALIKAS, K. (2018). Retrieval of Chlorophyll a from Sentinel-2 MSI Data for the European Union Water Framework Directive Reporting Purposes. *Remote Sensing* **11**, 64.

ANTHWAL, T., CHANDOLA, A. & THAPLIYAL, M. P. (2018). Performance Analysis of ANN Model for Estimation of Trophic Status Index of Lakes. *IAES International Journal of Artificial Intelligence (IJ-AI)* **7**, 1.

BUISHAND, T. (1982). Some methods for testing the homogeneity of rainfall records. *Journal of Hydrology* **58**, 11–27.

CAO, M., QING, S., JIN, E., HAO, Y. & ZHAO, W. (2021). A spectral index for the detection of algal blooms using Sentinel-2 Multispectral Instrument (MSI) imagery: a case study of Hulun Lake, China. *International Journal of Remote Sensing* **42**, 4510–4531.

COPERNICUS CLIMATE CHANGE SERVICE (C3S) (2017). ERA5: Fifth generation of E CMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS).

FANG, C., SONG, K., LI, L., WEN, Z., LIU, G., DU, J., SHANG, Y. & ZHAO, Y. (2018). Spatial variability and temporal dynamics of HABs in Northeast China. *Ecological Indicators* **90**, 280–294.

FUNK, C., PETERSON, P., LANDSFELD, M., PEDREROS, D., VERDIN, J., SHUKLA, S., HUSAK, G., ROWLAND, J., HARRISON, L., HOELL, A. & MICHAELSEN, J. (2015). The climate hazards infrared precipitation with stations—a new environmental record for monitoring extremes. *Scientific Data* **2**, 150066.

HAMED, K. H. & RAMACHANDRA RAO, A. (1998). A modified Mann-Kendall trend test for autocorrelated data. *Journal of Hydrology* **204**, 182–196.

HU, C. (2009). A novel ocean color index to detect floating algae in the global oceans. *Remote Sensing of Environment* **113**, 2118–2129.

HUETE, A., DIDAN, K., MIURA, T., RODRIGUEZ, E., GAO, X. & FERREIRA, L. (2002). Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* **83**, 195–213.

KLISAROVA, D., GERDZHIKOV, D., KOSTADINOVA, D. & DERMENDZHIEVA, G. (2019). Phytoplankton quantitative development and species diversity in the Bulgarian Black Sea waters during 2014-2017. Tech. rep.

MEE, L. D., FRIEDRICH, J. & GOMOIU, M. T. (2005). Restoring the Black Sea. Tech. Rep. 2.

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**, 394–406.

NOVINITE.COM (2020). Lake Varna: Tons of Dead Fish at the Shore, What is the Reason? - Novinite.com - Sofia News Agency.

ROUSE J.W., J., HAAS, R., SCHELL, J. & DEERING, D. (1974). Monitoring Vegetation Systems in the Great Plains with ERTS. In *Third Earth Resources Technology Satellite-1 Symposium- Volume I: Technical Presentations NASA SP-351*, S. C. Freden, E. P. Mercanti & M. A. Becker, eds., vol. 30103017. NASA, pp. 301–217.

SMITH, M. E. & BERNARD, S. (2020). Satellite Ocean Color Based Harmful Algal Bloom Indicators for Aquaculture Decision Support in the Southern Benguela. *Frontiers in Marine Science* **7**, 61.

YUNEV, O. A., CARSTENSEN, J., MONCHEVA, S., KHALIULIN, A., ÆRTEBJERG, G. & NIXON, S. (2007). Nutrient and phytoplankton trends on the western Black Sea shelf in response to cultural eutrophication and climate changes. *Estuarine, Coastal and Shelf Science* **74**, 63–76.

APPENDIX

Table 2: Sentinel-2 bands, their resolutions, and wavelengths.

Sentinel-2 SR Band	Description	Native Resolution	Wavelength
B1	Aerosols	60 meters	443.9nm (S2A) / 442.3nm (S2B)
B2	Blue	10 meters	496.6nm (S2A) / 492.1nm (S2B)
B3	Green	10 meters	560nm (S2A) / 559nm (S2B)
B4	Red	10 meters	664.5nm (S2A) / 665nm (S2B)
B5	Red Edge 1 (RE1)	20 meters	703.9nm (S2A) / 703.8nm (S2B)
B6	Red Edge 2 (RE2)	20 meters	740.2nm (S2A) / 739.1nm (S2B)
B7	Red Edge 3 (RE3)	20 meters	782.5nm (S2A) / 779.7nm (S2B)
B8	Near Infrared (NIR)	10 meters	835.1nm (S2A) / 833nm (S2B)
B8A	Red Edge 4 (RE4)	20 meters	864.8nm (S2A) / 864nm (S2B)
B9	Water vapor	60 meters	945nm (S2A) / 943.2nm (S2B)
B11	Short-wave Infrared 1 (SWIR 1)	20 meters	1613.7nm (S2A) / 1610.4nm (S2B)
B12	Short-wave Infrared (SWIR 2)	20 meters	2202.4nm (S2A) / 2185.7nm (S2B)

NDCI Range	Chlorophyll-a Range [mg/m^3]
< - 0.1	< 7.5
-0.1 to 0	7.5 - 16
0 to 0.1	16 - 25
0.1 to 0.2	25 - 33
0.2 - 0.4	33 - 50
0.4 to 0.5	>50
0.5 to 1	severe bloom

Table 3: Correspondence between NDCI ranges and Chlorophyll-a concentrations.