ZONATION OF SUBALPINE LAKES BASED ON REMOTELY SENSED WATER QUALITY PARAMETERS

C. Gerosa 1*, M. Bresciani 2, G. Luciani 3, C. A. Biraghi 1, D. Carrion 1, M. Rogora 4, M. A. Brovelli 1,3

1 Dept. of Civil and Environmental Engineering, Politecnico di Milano – Lecco Campus, Via Gaetano previali 1/c, 23900 Lecco, Italy (chiara.gerosa, giulia.luciani, carloandrea.biraghi, daniela.carrion, maria.brovelli@polimi.it)
2 Istituto per il Rilevamento Elettromagnetico dell’Ambiente, CNR-IREA, via Bassini 15, 20133 Milano – bresciani.m@irea.cnr.it
3 Istituto per il Rilevamento Elettromagnetico dell’Ambiente, CNR-IREA, via Bassini 15, 20133 Milano – maria.brovelli@polimi.it
4 Istituto di Ricerca sulle Acque, CNR-IRSA, Largo Tonolli 5028922 - Verbania Pallanza (VB) – michela.rogora@irsa.cnr.it

Commission III, WG III/10

KEY WORDS: Zonation, Water Quality, Remote Sensing, Lakes, Clustering, Water Monitoring.

ABSTRACT:

SIMILE is an INTERREG Italy-Switzerland project that aims to preserve water quality of the subalpine lakes Como, Lugano and Maggiore (Northern Italy), through an integrated innovative monitoring system. For this purpose, satellite images are processed to map and monitor Chlorophyll-a (CHL-a), Total Suspended Solids (TSM) and Lake Water Surface Temperature (LWST). This study combines these remotely sensed water quality parameters (WQP) maps, produced for the SIMILE project during 2019-2020, to propose and discuss a zonation approach that can support the monitoring of the study lakes through the analysis of spatial and temporal dynamics of the selected parameters. The approach consists in performing a cluster analysis on a combined sample of WQP maps, on a monthly basis, for each lake; then the different lake clusters are compared over time, through time series analysis of the WQP patterns. Finally, the clusters patterns are aggregated over time to map the lakes’ areas that have experienced higher or lower WQP values during 2019-2020. The results show a high spatial variability for the lakes under study, both during the different seasons and years; a North-South gradient has been identified for all WQP patterns, requiring for further investigation.

1. INTRODUCTION

1.1 Lake water quality monitoring to support decision-making in the SIMILE project

Lakes are fundamental socio-economic resources as well as home to complex and fragile ecosystems. They provide essential ecosystem services (Schallenberg et al., 2013) as the freshwater supply for civil, industrial, agricultural purpose but they are also important resources from a recreational and touristic point of view. They can also serve to mitigate climate change (Schallenberg et al., 2013). However, to maintain ecosystems services provided by lakes, it is crucial to preserve their water quality by phenomena as acidification, eutrophication and other disturbances due to an excessive anthropic pressure and global warming (Solimini et al., 2006, Vincent, 2009). To this aim, it is crucial to perform a correct management and protection of lake water resources; to support decision-making, a regular and affordable monitoring of water quality is needed. Therefore, a frequent and comprehensive monitoring system is essential, leading also to a better understanding of lake processes and temporal dynamics.

This is the strategy of SIMILE (Integrated monitoring system for knowledge, protection and valorisation of the subalpine lakes and their ecosystems), an INTERREG Italy-Switzerland project, started in 2019 (Brovelli et al., 2019). The project focuses on the subalpine lakes Como, Lugano and Maggiore (Figure 1) whose catchments cross Italy and Switzerland, requiring therefore joint coordinated management policies. For this reason, SIMILE involves partners from different countries and sectors cooperating to preserve the water quality of the lakes under study. Partners are from academic (Politecnico di Milano – Lecco Campus; Fondazione Politecnico; SUPSI - University of Applied Sciences and Arts of Southern Switzerland), research (Water Research Institute - National Research Council) and public institutions (Lombardy Region, Italy; Ticino Canton, Switzerland) sectors. Their first objective is to design an innovative integrated monitoring system, which represents an evolution with respect to the traditional campaigns for lake water quality, to support decision and policy making. Particularly, SIMILE designed monitoring system takes advantage of the information deriving from different sources as satellite imagery, in situ sensors for high frequency data collection and Citizen Science methods (Brovelli et al., 2019). This study will be developed starting from the remotely sensed data collected by SIMILE project since 2019.

Figure 1. The study area: lakes Maggiore, Lugano and Como, situated in the insubric region, between Italy and Switzerland, and their main characteristics (from Salmaso et al., 2012).

Table: Lake Surface (km²) Volume (km³) Max depth (m) Mean Depth (m) Altitude (m a.s.l.) Trophic state

<table>
<thead>
<tr>
<th>Lake</th>
<th>Surface (km²)</th>
<th>Volume (km³)</th>
<th>Max Depth (m)</th>
<th>Mean Depth (m)</th>
<th>Altitude (m a.s.l.)</th>
<th>Trophic state</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maggiore</td>
<td>212</td>
<td>17</td>
<td>370</td>
<td>177</td>
<td>191</td>
<td>oligo-mesotrophic</td>
</tr>
<tr>
<td>Lugano</td>
<td>28</td>
<td>4.69</td>
<td>288</td>
<td>171</td>
<td>271</td>
<td>meso-oligotrophic</td>
</tr>
<tr>
<td>Como</td>
<td>145</td>
<td>22.5</td>
<td>410</td>
<td>154</td>
<td>198</td>
<td>mesotrophic</td>
</tr>
</tbody>
</table>

* Corresponding author

This contribution has been peer-reviewed.
https://doi.org/10.5194/isprs-archives-XLIII-B3-2021-551-2021 | © Author(s) 2021. CC BY 4.0 License. 551
1.2 SIMILE Remotely sensed WQPs

Remote sensing methods represent an opportunity for lakes monitoring: differently from in situ traditional measurements, they can simultaneously monitor large areas, with a relevant temporal coverage, catching the spatial and temporal variability of optically active Water Quality Parameters (WQPs) (Giardino et al., 2015). Particularly, due to last years’ continuous improving of satellite sensors design and corresponding resolutions, inland water quality studies and monitoring applications based on optical satellite images have notably increased (Topp et al., 2020; Bresciani et al., 2020; Tyler et al., 2016). Earth Observation has become a recognized and validated integrative tool for monitoring lakes, able to extend the traditional punctual point of view of sampling campaigns to a synoptic view of lake water quality status.

SIMILE exploits satellite imagery to frequently and freely monitor the following main WQPs (Brovelli et al., 2019) that are important descriptors of water quality status:

- Surface concentration of Chlorophyll-a (CHL-a);
- Surface concentration of Total Suspended Solids (TSS);
- Lake Surface Water Temperature (LWST).

For this scope, the ESA Sentinel-3 A and B OLCI images have been processed since 2019 to map CHL-a and TSM concentrations, offering a daily revisit time and spatial resolution of 300m (Free et al., 2021). Imagery from Thermal Infrared Sensor TIRS on board NASA Landsat 8 satellite has also been used to monitor LSWT, providing a higher spatial resolution of 30m with 16 days revisit time. SIMILE also exploits ESA Sentinel-2 A and B MSI images to possibly investigate, on a higher level of detail (10-20m), some detected anomalies or unexpected concentrations (Luciani et al., 2021).

The processing of images is based on free and opensource algorithms (Luciani et al., 2021), validated for the specific application to inland aquatic environments (Soomets et al., 2020, Free et al., 2021) and applied within the free software SNAP distributed by ESA (Zuhlike et al., 2015). Sentinel-3 imagery is processed through the Case 2 Regional Coast Colour C2RCC (Brockmann et al., 2016) that, based on a neural network, performs radiometric and atmospheric correction, and automatically retrieves CHL-a and TSM concentrations. LWST is instead computed with the Barsi method (Barsi, 2015). For both Sentinel-3 and Landsat-8 imagery processing includes semi-automatic mask application on disturbed pixels (e.g. for clouds presence).

All WQPs maps produced within SIMILE are available on a dedicated Web Map Service (Toro Herrera et al., ISPRS 2021) to all interested users as a tool for decision-making and lake water quality protection.

1.3 The purpose of a zonation of lakes

Under SIMILE purpose, the frequent production of remotely sensed lakes WQPs maps is meant also to analyse and better understand water quality trends over space and time of the three lakes under study (Brovelli et al., 2019). Particularly, this paper aims at combining CHL-a, TSM and LSWT maps produced for SIMILE, during 2019-2020, to propose and discuss a zonation for lakes Como, Lugano and Maggiore on the basis of the selected water quality parameters, to improve the knowledge of lakes spatial and temporal variability and complexity (Ryberg, 2006). Particularly, zonation could allow for prioritization, focusing the efforts in the areas that most likely require interventions to improve water quality.

In addition, future studies could further exploit lakes zonation results based on water quality patterns in a combination with lakes geomorphology and hydrodynamics to better understand and model the lakes functioning.

2. DATASET DESCRIPTION

2.1 SIMILE remotely sensed WQPs for years 2019-2020

This study starts from the WQPs maps produced by the SIMILE project for the years 2019-2020, a large dataset of remotely sensed CHL-a, TSM and LWST maps. Since these WQPs are retrieved from different sources of satellite imagery, there are some differences among themselves that need to be addressed before the processing. The main differences are shown in Table 1 and are linked to the limits of the satellite sensors exploited by SIMILE. CHL-a and TSM maps share the same characteristics, indeed they are detected from Sentinel-3 A and B images. From 2019 to 2020 nearly two hundred CHL-a and TSM maps have been produced with a spatial resolution of 300m, at least on a weekly basis. LWST maps, on the other hand, are less in number (due to the longer revisit time of Landsat 8 TIRS) and have a spatial resolution of 30m (Table 1). The minimum LWST map availability is on a monthly basis but it can be improved by taking advantage also of LWST maps that cover just some portions of lakes Como and Maggiore.

<table>
<thead>
<tr>
<th>WQPs 2019-2020</th>
<th>Number of WQPs produced maps</th>
<th>Temporal resolution of WQPs maps</th>
<th>Spatial resolution of WQP map</th>
<th>Spatial coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHL-a map</td>
<td>198</td>
<td>At least weekly</td>
<td>300m</td>
<td>Whole study area</td>
</tr>
<tr>
<td>TSM map</td>
<td>198</td>
<td>At least weekly</td>
<td>300m</td>
<td>Whole study area</td>
</tr>
<tr>
<td>LWST map</td>
<td>61</td>
<td>At least monthly</td>
<td>30m</td>
<td>Whole study area- portions of lakes Como or Maggiore</td>
</tr>
</tbody>
</table>

Table 1. The remotely sensed WQPs maps dataset for years 2019-2020.

The temporal resolution and the final number of WQPs produced maps is related to the weather conditions during the image acquisition: since the approach adopted in SIMILE exploits optical satellite imagery, it must handle with cloud coverage that can shadow the sensors sight also for many days. Satellite images were processed when the cloud coverage did not cover the entire lake areas. As a consequence, the temporal resolution of this dataset is lower than the satellite sensors revisit time, linked to the effective temporal resolution of them (Hestir et al., 2015). Moreover, cloud coverage is not the only reason why information may be less frequent over a certain area during time: often some pixels covering lakes can experience some radiometric disturbs (e.g. glint, adjacency effects) or could result in mixed pixels (next to the shore) (Sagan et al., 2020).
These pixels would result in a wrong WQP detection and therefore they are semi-automatically masked during the processing, resulting in some areas of lakes with no data. The WQPs maps therefore do not cover all lakes areas with the same frequency. Thus, the following study and procedure for zonation must handle with no data pixels as well as periods that can lack data more than others. This issue has been dealt by considering data aggregation over time.

### 2.2 Pre-processing of WQPs maps

Before performing the cluster analysis, the input WQPs dataset has been explored and filtered in order to avoid future wrong results due to out-of-range values. Their occurrence is low; however, it could influence the cluster analysis results. For LWST maps, the statistics of each map have been considered to detect values which were too high or too low with respect to the other ones in the map, possibly due to the presence of disturbed pixels in the original satellite image. The applied processing is shown in Figure 2 for LWST map referred to date 06-01-2019. The frequency of occurrence of LWST values in each map has been investigated using Matlab software. In a second step, the out-of-range detected values have been explored on the map to interpret the reason of the anomaly with respect to geographical location (e.g. lower surface temperature values could be admissible when detected in the area affected by inflowing waters) or the image characteristics (e.g. lower temperature due to cloud shadows). If also this step confirmed an unrealistic LWST distribution, the detected out-of-range pixels were filtered out, when too low, or masked with a suitable maximum value, when too high. The objective of the whole filtering was to remove the pixels that could have brought to a wrong result, without losing too much information.

Also, CHL-a and TSM maps were analysed. Based on the literature (Giardino et al., 2014) and on information on the characteristics of the lakes under study (see Fig. 3, Fig.4) pixels with values clearly higher than expected were considered as out-of-range and set to 15 mg/m³ for CHL and 10 mg/m³ for TSM. These mentioned values were chosen as extreme limits, out of WQPs common ranges, to avoid filtering out admissible detected information.

Figure 3 shows the statistics of the WQPs for the full dataset. WQPs mean trends for all lakes under study are also shown in Figure 4 for the years 2019-2020.

![Figure 2. Out-of-range data filtering, by means of numerical and visual analysis, for the LWST map of 06-01-2019.](image)

![Figure 3. Statistics for each WQP for each year under study.](image)

![Figure 4. WQPs mean values trends for 2019-2020 for the full dataset.](image)

### 2.3 Combined WQPs maps

In the end, the corrected CHL-a, TSM and LWST maps were combined considering their corresponding dates. Nearly 30 combinations were found for each year, at least on a monthly basis. In order to combine the maps the following criteria have been considered:
Maps providing a more complete coverage of the single lake areas (with few missing pixels) have been considered.

Since LWST behaves more predictably and changes slowly over time, differently from the other parameters, it was decided to combine LWST maps with CHL-a and TSM also in the case the date of LWST map was slightly different (few days before or after CHL-a and TSM referred date).

The final combined maps for 2019-2020 years describe the spatial distribution of WQPs in the lakes’ areas through time. Not all combinations are available for all lakes under study, due to the limited spatial coverage of some LWST maps.

In addition, the spatial resolution of LWST maps was resampled at 300m to be consistent with the spatial resolution of CHL-a and TSM maps. The WQPs values were finally normalized with respect to each WQP maximum value, as in literature (Akbar et al., 2011), to prepare WQPs combined maps for the cluster analysis. Indeed, by normalizing WQPs values it is possible to consider all WQPs with the same importance.

3. METHODOLOGY AND RESULTS

3.1 Cluster analysis over t-map

Lakes zonation was obtained starting from a clustering methodology, (Akbar et al.,2011), based on the previously described combination of CHL-a, TSM and LSWT parameters for maps with corresponding dates.

Each lake was processed separately, in order not to miss its specific spatial variability of water quality parameters through the year. Cluster analysis was applied on every combined sample of CHL-a, TSM and LSWT maps referred to a reference date. A virtual raster was obtained considering the three WQPs as bands of one t-map referred to that single reference date.

Clustering was performed through K-means algorithm (Lloyd, 1957) in QGIS. The purpose was to automatically group the pixels of each Day-image in a suitable number of clusters, by means of numerical similarities among normalized WQPs values for that day. In K-means algorithm, a cluster centroid is obtained according to the initial value of the pixel selection process. Then, pixels are recursively assigned to a cluster based on the Euclidean distance between their values and the centroid of the cluster that changes at each step as soon as the cluster grows. Some other algorithms, such as the ISODATA, were considered, however K-means provided the most significant results. Also in literature (Ragno et al., 2007, Areerachakul & Sanguansintukul, 2010), K-means has proven to be effective toanalyse the large and complex WQPs data with several parameters corresponding to different units and ranges.

Some tests aimed to find out the most suitable number of clusters that could meaningfully group the different water quality patterns and detect the different areas’ behaviours in an easily interpretable way were performed (see the example on Day-image of 03-09-2019 in Figure 5). The most suitable number of clusters that has been selected is 5.

The influence of the TSM parameter on the zonation was explored: the clustering was performed both including it or not with CHL-a and LSWT values (Figure 6).
3.2 The time-series of clusters over the lakes

The resulting clusters have been compared across the 2019-2020 mapped days to investigate temporal and spatial trends, variability and significance of the clusters. The K-means algorithm groups pixels only on the basis of similar distribution of WQPs values. To be able to compare the maps through time it was necessary to label the clusters in an interpretable way, according to the WQPs values ranges included in each cluster. For each cluster in a map, the mean values for CHL-a, TSM, LWST have been computed. An RGB composition was used to represent the clusters: LWST values were assigned to the Red channel, CHL-a to the Green channel and TSM to Blue. In this way, the colour composition provides information about the higher or lower values corresponding to each channel and, consequently, to each WQP. Figure 7 shows a time series for Lake Maggiore, with the clusters represented with this RGB composition. This helps to detect the similar patterns through the seasons of the two years. The map legend has been produced highlighting the presence in each cluster of high or low values for the WQPs. It is possible to recognise some major spatial patterns that are present in the lake, in particular a North-South gradient. Also, some colours tend to match between over the seasons. This cluster series has allowed to show the behaviour of the three WQPs into the different clusters referred to the dates under study. Particularly, looking at the frequency of occurrence of the different colours in the time series, it is possible to understand also the correlations through time among these WQPs in each lake: for example, the low occurrence of yellow coloured cluster for lake Maggiore can be interpreted as a low probability to have CHL-a and LWST high values while TSM is low; on the other hand, the high occurrence of red or light blue clusters shows that there is not much correspondence between higher LWST and a higher presence of CHL-a and TSM.

Cluster time series were computed also by combining the three WQPs patterns of each cluster, in order to represent the lakes areas that showed higher or lower combinations of WQPs through the year. Results are shown in Figure 8 for all the lakes under study, comparing 2019 and 2020. This visualization allows to label the clusters according to a generalized water quality pattern, that helps for interpretation. However, the two time-series in Figure 7 and Figure 8 should be analysed together to understand the combinations of the WQPs values that bring to a higher or lower generalized water quality pattern. Visualization of Figure 8 confirms a North-South gradient not only for Lake Maggiore but also for the other lakes.

3.3 Aggregation over time

The time series of clusters was useful to aggregate the data over time. Mean values had been computed within each cluster for each WQP: to the pixels of each clustered t-map were assigned the mean WQPs values of the cluster they were grouped in. Then, these WQPs mean values were averaged for each pixel over time to aggregate the information coming from the clusters along the year. The resulting final averaged mean WQPs values for each pixel were grouped in 5 new clusters by considering the standard deviation of WQPs for each lake under study and normalizing with respect to the average value of WQP for each lake and for each year under study. In this way, it was possible to detect the areas of lakes whose pixels were grouped more frequently in classes with higher WQPs, showing a lower water quality. Classes are numbered from 1 to 5, with 5 meaning for higher WQPs. An example for these results is shown in Figure 9 for Lakes Como and Maggiore, where this approach has allowed to detect for example the different TSM and CHL-a distributions between the three main branches of Lake Como.

Figure 7. Time series clusters for Lake Maggiore for years 2019-2020. Position of lake cluster maps is due to the date they are referred to. Clusters are represented with RGB composition linked to clusters mean WQPs values (Red channel represents LWST, Green CHL-a, Blue TSM).
Figure 8. Time series clusters for lakes Maggiore, Como and Lugano for years 2019-2020. Position of lakes cluster maps is due to the date they are referred to. Particularly, lake Lugano lacks comprehensive information for the first semester of 2020. The legend is linked to the generalized combined mean WQPs values within each cluster.
In addition, also some generalized WQPs patterns for 5 generalized clusters were defined, as in literature (Akbar et al., 2010), computed by averaging together the three normalized WQPs. This allowed to single out lakes areas that could have shown in 2019-2020 the higher or lower WQPs. This analysis showed a high variability for the whole combination of WQPs.

4. DISCUSSION AND CONCLUSIONS

The proposed zonation has allowed to single out lakes areas that showed a similar behaviour during 2019-2020, allowing for a comparison also between the two years under study.

The time-series of clusters have shown high variability during the seasons and years. However, it is possible to perceive some major areas that behave differently during the year, as the three different branches of Lake Como or the two distinct basins of Lake Lugano (North and South basin). Also, in Lake Maggiore a quite evident North-South gradient was observed, as well as some differences in the major gulls (e.g. the Borromeo Gulf), confirming some spatial differences which have been described through monitoring data (Rogora et al., 2019).

As it is possible to perceive from Figure 8, Lake Lugano clusters time series showed a higher occurrence of missing data over its area but also with respect to the periods under study: while Lake Como and Lake Maggiore WQPs t-maps show a higher temporal resolution, due to the major overpass of Landsat 8 TIRS on their areas (due to swath overlapping), Lake Lugano was present only in the images that detected the whole study area, with a consequent lower effective temporal resolution. Moreover, due to medium-low spatial resolution of Sentinel 3 OLCI satellite imagery it is more difficult to map a smaller lake as Lugano, due to mixed pixels problems and adjacency effects that cause many pixels to be masked. In a future study, Lake Lugano zonation could be assessed by remotely sensed products with a higher spatial resolution, to be able to better catch its spatial variability.

ACKNOWLEDGEMENTS

This study is part of SIMILE (Informative System for the Integrated Monitoring of Insublic Lakes and their Ecosystems, ID: 523544) and as such it is supported by the European Commission in the frame of Interreg Italy-Switzerland 2014-2021 program.

REFERENCES


