

Water quality estimation in Greek lakes from Landsat 8 multispectral satellite data

E. Kontopoulou*, P. Kolokoussis and K. Karantzalos

Remote Sensing Laboratory, National Technical University of Athens, Greece

* e-mail: evi.kontopoulou@gmail.com

Abstract: The European Water Framework Directive (WFD) requires that a Good Environmental Status is retained for all water bodies along their constant monitoring. However, the frequent (e.g., weekly, monthly) monitoring of numerous and extensive inland water bodies places several logistic and economic challenges. Remote sensing is considered one of the most cost-effective methods for providing the spatial and temporal environmental data that are necessary for the effective monitoring and implementation of the WFD. In this study the goal was to establish concrete relations between satellite observations and key quality parameters over several permanent lake monitoring stations in Greece. In particular, water quality indicators such as Chlorophyll-a and turbidity through Secchi disk depth were estimated through the analysis of Landsat 8 satellite and in-situ data. The employed multispectral images and in-situ measurements were selected to have a maximum temporal offset of 15 days, while relations were established based on regression models. The experimental validation indicated correlations with R^2 values of up to 0.75 for Chl-a ($n>160$) and 0.70 for turbidity ($n>190$). These promising results along with a further multi-temporal validation can be employed for the development of operational, remote sensing monitoring systems for the accurate and timely estimation of key water quality indicators.

Key words: Remote sensing, Water Framework Directive, Earth observation, regression, chlorophyll, turbidity

1. INTRODUCTION

Water is a natural resource of utmost importance and a precondition for human, animal and plant life cycles. The adoption of the Water Framework Directive (WFD) in 2000, establishes a legal framework to guarantee sufficient quantities of good quality waters across Europe. In order to obtain a coherent and comprehensive overview of water status, frequent surveillance and a well organised monitoring program is required.

The rife methodology to examine the quality of water consists of in-situ sampling and the consequent chemical analysis, method that leads to accurate estimations for several water quality parameters. However, the results correspond to the exact location/region from which the sample was taken but not to the total extend of the water body. Moreover, it is an expensive method due to the equipment used and the frequent sampling required for monitoring crucial phenomena.

In this study the possibility of estimating key monitoring indicators for lakes, using Remote Sensing and specifically Landsat 8 high resolution satellite data has been examined. More precisely, this study has tried to evaluate water quality indicators documented in relevant literature as well as to determine new indicators, capable of providing significant correlations between high resolution satellite data and in-situ ground truth data, for two key monitoring indicators, Chlorophyll-a and Secchi disk depth. The literature is particularly focused on these parameters, since they are the main indicators of eutrophication.

Satellite remote sensing is already internationally used for water quality estimation. However, because of the optical complexity of natural inland waters, the creation of a widely used algorithm is difficult and an ongoing research is necessary (Bukata et al., 1985; Carder et al., 1989; Theologou et al., 2016).

The most widespread Chlorophyll-a retrieval algorithms used for the estimation of Chlorophyll-a concentration employ the blue, the green, the yellow, the red and the infrared spectral regions. The

selection of the most appropriate bands to be used depends on the water body category. For inland water bodies the preferred band ratios for Chlorophyll-a concentration include blue and green bands (440-550 nm), because chlorophyll absorption is maximum at this spectral region. For coastal waters the above ratios are not sensitive enough and the use of the red (620-700nm) and infrared spectral bands ($>700\text{nm}$) are also required (Le et al., 2013; Theologou et al., 2016).

For the turbidity estimation in relevance to the Secchi disk depth measurements the use of ratios that include red reflectance are widely used. For multispectral sensors like Landsat 8, there are promising results for Secchi disk depth correlation even with the red band itself (Hellweger et al., 2007). The reflection at this bandwidth increases when the turbidity and Secchi disk depth decrease. Additionally, the red to blue band ratio provides good results when used for a lake (Kloiber et al., 2002).

2. MATERIALS AND METHODS

The methodological steps (Figure 1) followed for establishing correlations between high resolution satellite data and in-situ ground truth data were the following:

- Collection of in-situ measurements
- Acquisition of high resolution Landsat 8 satellite data
- Correlation of high resolution satellite data and in-situ ground truth data
- Determination of adequate water quality indicators via regression analysis
- Production of water quality maps showing Chl-a and Turbidity spatial distribution

Since up to now a single open database with all in-situ measurements per inland monitoring station does not exist, in-situ data from 11 representative lakes in Greece have been gathered. In particular, Chlorophyll-a and Secchi disk depth measurements were collected from the Greek Biotope/Wetland Centre and the Management Body of Ecodevelopment Area of Karla-Mavrovouni-Kefalovriso-Velestino for the years 2013, 2014 and 2015 (Table 1).

Atmospherically corrected satellite data were downloaded via the 'United States Geological Survey' website. From the available Landsat 8 images, those that included the 11 lakes of the study area and had a maximum 15 days time-offset between the in-situ sampling date and the acquisition of the satellite image (Table 2), were used.

Methodology	
Study Area	11 Greek Lakes <ul style="list-style-type: none"> • 6 natural: Amvrakia, Vegoritida, Volvi, etc • 5 artificial: Polifitou, Tauropos, etc
In-situ Data	<ul style="list-style-type: none"> • Chl-a & • Secchi disk depth <p style="text-align: right;">(Table 1)</p>
Satellite Data	Multispectral Landsat 8 data <ul style="list-style-type: none"> • Atmospherically corrected • Maximum 15 days offset <p style="text-align: right;">(Table 2)</p>
Established Correlations	Linear Regression: <ul style="list-style-type: none"> • 155 indices for Chl-a • 105 for Secchi disk <p style="text-align: right;">(Table 3)</p>
Model Selection	Selection of the regression models with the highest correlations rates
Produce Geospatial Maps	Production of water quality estimation maps (Figure 2)

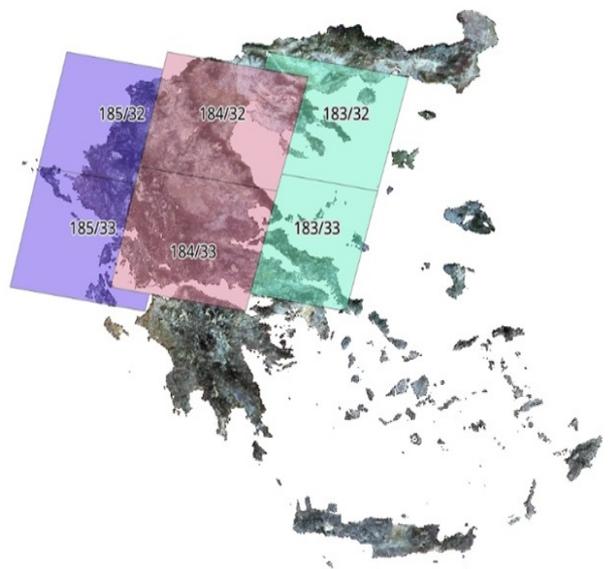


Figure 1. Methodological steps followed and Path/Row Map for Landsat 8 images used.

Table 1. Number of in-situ measurements for the quality parameters / lake

Measurements / Quality Parameter	Chl-a	Secchi Disk Depth	Total Measurements
Amvrakia	8	8	24
Vegoritida	8	8	23
Volvi	8	7	23
Megali Prespa	16	16	47
Trichonida	8	8	24
Yliki	5	5	15
Polifitou	8	8	23
Tauropos	8	8	24
Kastraki	7	7	21
Kremasta	8	8	23
Karla	16	33	49
Total	100	124	304

In order to extract the reflectance values from the selected images at the sampling coordinates, a Python script has been employed. Outliers due to extensive cloud cover and certain inaccurate reflectance calculations during atmospheric correction were excluded.

The possible consistent relationships between the water quality parameters and the satellite reflectance values of Landsat 8 images were examined via regression techniques, and more specifically via simple linear regression. This method was selected in accordance with the empirical models of the literature. Additionally, the correlation between the Chl-a concentration logarithms and the satellite reflectance values was also examined via linear regression equations. Finally, the best regression models were selected and water quality estimation maps were produced for Chl-a and Secchi disk depth spatial distribution.

Due to the large volume of the data, the automation of the process was essential and has been achieved with the implementation of regression analysis through Matlab scripts. The methodological steps mentioned above as well as a path/row map of the Landsat 8 images used are presented in Figure 1.

Table 2. Path, Row, Number of Landsat 8 Surface Reflectance images used / lake

Quality Parameters	Landsat Data		
	WRS Path	WRS Row	Number of images
Amvrakia	184/185	33	29
Vegoritida	184/185	32	26
Volvi	183/184	32	26
Megali Prespa	185	32	15
Trichonida	184	33	12
Yliki	183	33	8
Polifitou	184	32	13
Tauropos	184	33	16
Kastraki	184/185	33	25
Kremasta	184/185	33	31
Karla 1	184	32/33	55
Karla 2	183	33	

3. EXPERIMENTAL RESULTS AND VALIDATION

From the 155 satellite image indices used for Chl-a and $\ln(\text{Chl-a})$ and the 105 indices used for Secchi disk depth, the best fitting indices between the water quality in-situ measurements and the high resolution Landsat 8 data are presented in Table 3.

It is worth mentioning that by employing more complex models through for example neural networks higher regression rates may be achieved. However, simple linear, exponential and logarithmic models have been employed in this work, similar to relevant models of the literature, in order to mainly examine the degree of correlation between in-situ observations and high resolution Landsat 8 data for different time windows, number of observations, water quality parameter ranges

and satellite image indices.

Table 3. Regression results for various time windows (dt), number of observations (n), water quality parameter ranges and satellite data indices (b1 to b7 are the atmospherically corrected Landsat 8 bands).

Parameter	R ²	R ²	R ²
Chl-a	dt = 0-15	dt = 0-5	dt = 0-3
	n = 168	n = 54	n = 39
	0.45-118.88 µg/l	0.45-118.88 µg/l	0.45-72.00 µg/l
b4-b7	0.74	0.77	0.60
exp(b4÷b2)	0.66	0.71	0.77
b3+b4-b5	0.64	0.62	0.52
exp(b4÷(b2+b7))	0.59	0.72	0.77
exp(b4÷(b2+b5))	0.59	0.63	0.80
Chl-a	dt = 0-15	dt=0-3	
	n=54	n=13	
	9.00-118.88 µg/l	12.52-72.00 µg/l	
b4-b7	0.63	0.69	-
b3+b4-b6	0.60	0.55	-
b3-b7	0.59	0.63	-
b3-b6	0.57	0.65	-
b2-b6	0.46	0.67	-
Ln (Chl-a)	dt = 0-15	dt = 0-5	dt = 0-3
	n = 168	n = 54	n = 39
	0.45-118.88 µg/l	0.45-118.88 µg/l	0.45-72.00 µg/l
(b2-b4)÷(b3-b6)	0.78	0.77	0.80
b2-b4	0.76	0.72	0.71
b4÷b2	0.75	0.77	0.79
Secchi disk depth	dt = 0-15	dt = 0-5	
	n = 193	n = 65	
	0.01-14.00 m	0.01-1.40m	
b2÷(b1+b2+b3)	0.70	0.70	-
log(b2÷b3)	0.63	0.63	-
ln((b2-b7)÷(b3-b6))	0.46	0.65	-
log(b2÷b3)	0.63	0.63	-
b2÷b3	0.63	0.63	-
b3÷b2	0.60	0.60	-

4. DISCUSSION

The experimental results demonstrated quite promising results and were generally in accordance with the literature. In particular, high R² values (up to 0.74) have been achieved for the estimation of Chl-a concentration for a concentration range between 0.45 to 118.88 µg/l and 168 observations. A better relation has been achieved with a linear regression of the Chl-a concentration logarithms and the satellite reflectance values (max R²=0.78, n=168).

Moreover, for a narrower three day time offset between the in-situ sampling date and the acquisition of the satellite image, the correlation rate of Chl-a and ln(Chl-a) with the satellite reflectance values exceeded 0.80 for 39 observations.

The regression models which have been suggested by the relevant literature for Chl-a estimation (Zhang and Han, 2015; Kabbara et al., 2008; Han and Jordan, 2005; Giardino et al., 2007; Hellweger et al., 2004; Theologou et al., 2015; Theologou et al., 2016) have achieved correlations with R² ranging from 0.67 to 0.93. However, these results were based on less number of observations and with a relatively narrower time window between the sampling date and the acquisition of the satellite images. It is worth mentioning that in most relevant studies as well as in this work, the models that presented the higher correlation rates for Chl-a always included the red spectral band. Most of the relevant studies did find good relation between Chl-a and a combination of the red with the blue and/or green spectral bands, while in this work relatively high correlations were also established between the red spectral band and the infrared.

Concerning the turbidity, as measured with Secchi disk depth, the experimental results demonstrated a correlation of 0.70 for 193 observations with a depth range from 0.01 to 14 m and

for a 15-day time window. In another experiment, similar R^2 rates were established for 65 observations and a 3-day time offset between the sampling date and the acquisition of the satellite image, but for smaller depth range (0.01-1.40 m). In the relevant literature (Kabbara et al., 2008; Kratzer et al., 2008; Zhang, 2005; Hellweger et al., 2004) the suggested models accomplished correlations with R^2 values ranging from 0.54 to 0.85 but for less observations and relatively narrower time windows. The models with the highest correlation rates, both in this study and the literature, used indices which included the blue, green and red spectral bands and sometimes bands from the infrared spectral region. It is worth mentioning that the best correlation in this study for Secchi disk depth has been achieved by combining the reflectance values of the coastal band of Landsat 8 with the reflectance values of the blue and green spectral bands: $b2/(b1+b2+b3)$.

Overall, the experimental results of this study indicate that high resolution, multispectral satellite data can be operationally employed for remotely monitoring key water quality parameters. Such remote sensing techniques have the advantage of being cost-effective, and have the ability to monitor consistently large areas even over not easily accessible water bodies. In addition, the high resolution remote sensing data can densely map the spatial distribution of the phenomena and provide information for the quality throughout the water body and not only at the sampling points (Figure 2).

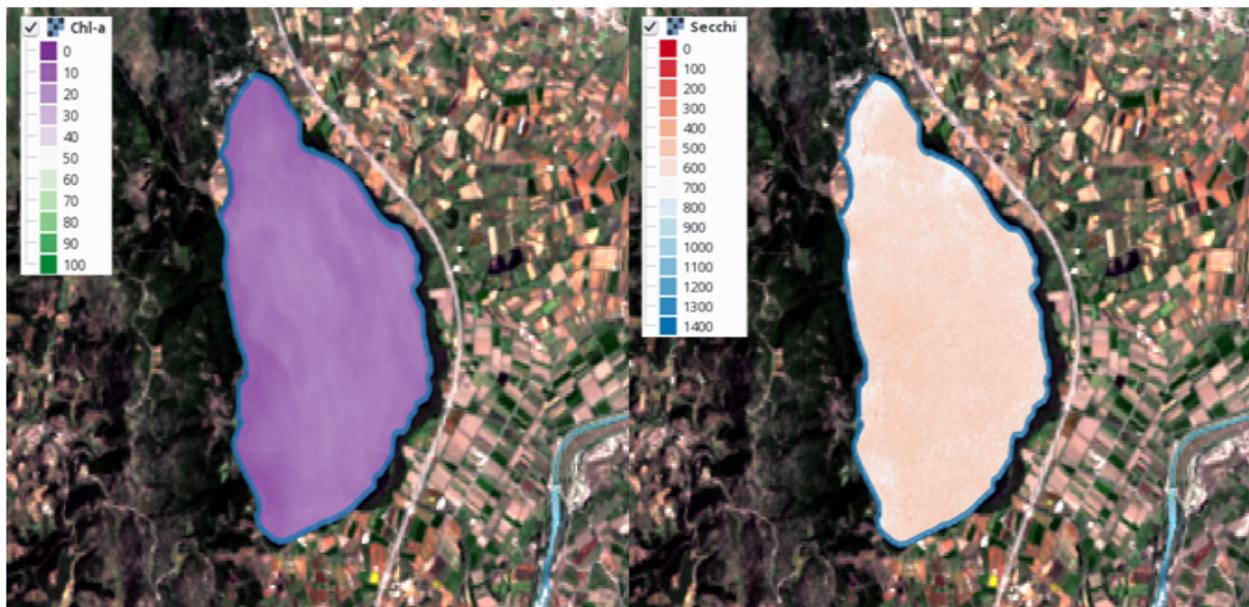


Figure 2. Water quality maps for the lake Amvrakia based on the established Chl-a ($b4-b7$) and turbidity ($b2/(b1+b2+b3)$) models.

5. CONCLUSIONS

In this study correlations were established between satellite and in-situ data for 11 representative lakes of Greece. The quite promising results regarding the estimation of Chl-a and turbidity indicate that operational earth observation systems can be developed through frequent satellite image acquisitions and thorough validation with dense spatially and temporarily in-situ data. The provision of the responsible authorities to carry out in-situ sampling as close as possible to the satellite image acquisition could make the evaluation more effective. Improving the calibration of the satellite sensors would also be a new asset for the accomplishment of better water quality parameter estimations. New perspectives for remote sensing technology arise from the upgrade of the satellites' spatial and spectral resolution, as the European Space Agency (ESA) has launched the new earth observation satellites Sentinel II and Sentinel III, and have already broadened the horizons in this field.

REFERENCES

- Bukata, P., Bruton, E., Jerome, H., 1985. Application of direct measurements to the estimation of lake water quality indicators. *Environment Canada Inland Waters Directorate*, 140: 35.
- Carder, L., Steward, G., Harvey, G., Ortner, P., 1989. Marine humic and fulvic acids: their effects on remote sensing of ocean chlorophyll. *Journal of Limnology and Oceanography*, 34: 68-81.
- Giardino, C., Brando, V., Dekker, A., Strömbeck, N., Candiani, G., 2007. Assessment of water quality in Lake Garda (Italy) using Hyperion. *Remote Sensing of Environment*, 109(2): 183-195.
- Han, L., Jordan, K., 2005. Estimating and mapping chlorophyll a concentration in Pensacola Bay, Florida using Landsat ETM data. *International Journal of Remote Sensing*, 26 (23): 5245-5254.
- Hellweger, C., Schlosser, L., Lall, U., Weisell, K., 2004. Use of satellite imagery for water quality studies in New York Harbor. *Estuarine, Coastal and Shelf Science*, 61(3): 437-448.
- Kabbara, N., Benkheil, J., Awad, M., Barale, V., 2008. Monitoring water quality in the coastal area of Tripoli (Lebanon) using high-resolution satellite data. *ISPRS Journal of Photogrammetry & Remote Sensing*, 63, 488-495.
- Kloiber, M. S., Brezonik, L. P., Olmanson, G. L., Bauer, E. M., 2002. A procedure for regional lake water clarity assessment using Landsat multispectral data. *Journal of Remote Sensing of Environment*, 82: 38-47.
- Kratzer, S., Brockmann, C., Moore, G., 2008. Using MERIS full resolution data to monitor coastal waters — A case study from Himmerfjärden, a fjord-like bay in the northwestern Baltic Sea, *Remote Sensing of Environment*, 112(5): 2284-2300.
- Le, C., Hu, C., Cannizzaro, J., Lee, Z., 2013. Evaluation of chlorophyll-a remote sensing algorithms for an optically complex estuary. *Remote Sensing of Environment*, 129: 75-89.
- Theologou I., Kagalou I., Papadopoulou M., Karantzas K., 2016. Multitemporal Mapping of Chlorophyll- α in Lake Karla from High Resolution Multispectral Satellite data. *Environmental Processes*, 3(3): 681-691.
- Theologou, I., Patelaki, M. & Karantzas K., 2015. Can single empirical algorithms accurately predict inland shallow water quality status from high resolution, multi-sensor, multi-temporal satellite data? *ISPRS Archives*, G. Schreier, P. E. Skrovseth, and H. Staudenrausch (eds.) *International Symposium on Remote Sensing of Environment*, XL-7/W3: 1511-1516
- Zhang Y., 2005. Surface water quality estimation using remote sensing in the Gulf of Finland and the Finnish Archipelago Sea, Ph.D. thesis, Lab. Space Phys., Helsinki Univ. Technol, 55: 82 p.