AGRICULTURAL UNIVERSITY OF ATHENS
SCHOOL OF ENVIRONMENT & AGRICULTURAL ENGINEERING
DEPARTMENT OF NATURAL RESOURCES DEVELOPMENT & AGRICULTURAL ENGINEERING
LABORATORY OF SOIL SCIENCE & AGRICULTURAL CHEMISTRY

PhD Thesis

Water quality assessment in Greek lakes by using remote sensing and statistical modelling

Vassiliki V. Markogianni

Supervisor
Dr. Dionissios P. Kalivas, Professor AUA

Supervising committee
Dr. Dionissios P. Kalivas, Professor AUA
Dr. Elias Dimitriou, Research Director HCMR
Dr. George P. Petropoulos, Assistant Professor HUA

ATHENS
2023

AGRICULTURAL UNIVERSITY OF ATHENS
DEPARTMENT OF NATURAL RESOURCES DEVELOPMENT & AGRICULTURAL ENGINEERING
LABORATORY OF SOIL SCIENCE & AGRICULTURAL CHEMISTRY

PhD Thesis

Water quality assessment in Greek lakes by using remote sensing and statistical modelling

Εκτίμηση της ποιότητας των υδάτων των Ελληνικών λιμνών μέσω τηλεπισκόπησης και στατιστικής μοντελοποίησης

Vassiliki V. Markogianni

Examining committee
Dr. Dionissios P. Kalivas, Professor AUA (supervisor)
Dr. Elias Dimitriou, Research Director HCMR
Dr. George P. Petropoulos, Assistant Professor HUA
Dr. George Papadopoulos, Associate Professor AUA
Dr. Emmanouil Psomiadis, Assistant Professor AUA
Dr. Konstantinos Soulis, Assistant Professor AUA
Dr. Ioannis Karaouzas, Associate Researcher HCMR

Water quality assessment in Greek lakes by using remote sensing and statistical modelling

Department of Natural Resources Development & Agricultural Engineering
Laboratory of Soil Science & Agricultural Chemistry

ABSTRACT

The aim of the present PhD thesis is the achievement of continuous monitoring and assessment of water quality (WQ) and trophic state of Greek lakes by exploiting the implementation of the Water Framework Directive (WFD) in Greece in synergy with satellite RS, providing parallel support to sustainable water resources management at a national scale. Continuous WQ monitoring is the most crucial aspect for lake management. Therefore, the methodological framework developed herein has as an ultimate goal the generation of lake WQ quantitative models while the practical use of this approach was developed and evaluated in a total of 50 lake water bodies (natural and artificial) from 2013–2018, constituting the National Lake Network Monitoring of Greece in the context of the WFD. Concerning the utilized Earth Observation (EO) data, images from Landsat 7 Enhanced Thematic Mapper Plus (ETM+) and Landsat 8 Operational Land Imager (OLI) sensors have been combined with co-orbital WQ *in-situ* measurements (Chl-*a*, Secchi Disk depths and Total phosphorus-TP- concentrations) with the main objective of delivering robust WQ assessment models.

From the statistical point of view, principal component analysis (PCA) was performed to explore Greek lakes' interrelationships among their Chl-*a* values and certain criteria, e.g. their characteristics (artificial/natural), WFD typology, climatic type (according to the Köppen-Geiger climate classification), season of water samplings and the date difference between sampling and satellite overpass. PCA highlighted the lake characteristics (natural/artificial) and WFD typology as the variables that mostly contribute to the variance of Chl-*a* concentration; thus, numerous stepwise multiple regression analyses (MLRs) among different groups of cases, formed by the PCA criteria, were implemented with basic aim the generation of different remote sensing-derived Chl-*a* algorithms for different types of lakes. Moreover, correlation analysis among *in-situ* co-orbital WQ data was conducted to explore and detect their inter-relationships. Subsequently, based on correlation analysis's results, further stepwise MLRs employing available *in-situ* TP and Secchi depth datasets were further implemented to establish optimal quantitative models. Eventually, trophic status classification was conducted herein, calculating Carlson's Trophic State Index (TSI) of each lake, initially throughout all lakes and then oriented toward natural-only and artificial-only lakes. The proposed scheme resulted in the development of

models separately for natural (R = 0.78) and artificial (R = 0.76) lakes, while the model developed without criteria proved weaker (R = 0.65) in comparison to the other ones examined. MLRs among Landsat data and Secchi depths resulted in 3 optimal models concerning the assessment of Secchi depth of all lakes ($Secchi_{general}$; R = 0.78; RMSE = 2 m), natural ($Secchi_{natural}$; R = 0.95; RMSE = 1.87 m) and artificial ($Secchi_{artificial}$; R = 0.62; RMSE = 1.36 m), with reliable accuracy. Study findings showed that TP-related MLR analyses failed to deliver a statistically acceptable model for the reservoirs; nevertheless, they delivered a robust $TP_{general}$ model for all lakes (R = 0.71; RMSE = 0.008 mg/L) and a $TP_{natural}$ model for natural lakes (R = 0.93; RMSE = 0.018 mg/L). Subsequently, regarding the TSI results, the higher deviation of satellite-derived TSI values in relation to *in-situ* ones was detected in reservoirs and shallower lakes (mean depth < 5 m), indicating noticeable divergences among natural and artificial waterbodies.

Next significant key questions that were answered are initially whether Landsat-based empirical WQ algorithms can be efficiently applied to Sentinel 2 images and then whether the combined use of multi-sensor data improves those algorithms' prediction accuracy. Hence, Sentinel 2 images of 2018 with concurrent dates with those of field measurements were utilized to facilitate a WQ models' efficiency evaluation and comparison with the respective Landsat's validation results. Concerning the results, in particular for general models of all WQ elements (Chl-*a*, Secchi depth and TP), all models were more efficient and accurate when were accompanied by Landsat images while no improvement was observed by using multi sensor images. Chl-*a* and TP models (for natural lakes) presented lower values of error metrics when employing Sentinel 2 images (RMSE Chl-*a*=16.4 μg/l vs 21.5 μg/l; RMSE TP=0.03 mg/l vs 0.031 mg/l) and only $Secchi_{natural}$ model performed better with Landsat data (2.8 m vs 2.9 m). Concerning artificial lakes, performance of Chl-*a* model was better by exploiting Landsat data (RMSE= 3.7 μg/l vs 7.7 μg/l of Sentinel 2) while Secchi model achieved slightly better efficiency with Sentinel 2 images (RMSE= 1.5 m vs 1.6 m of Landsat). The largely worse performance of Chl-*a* models compared to rest of WQ elements (median MAPE values ranged from 42 % to 58%, Secchi depth from 24% to 44% and TP from 22% to 38%), emphasized once again the complexity that mapping of Chl-*a* in Case 2 waters (coastal and/or inland waters) hides.

Today, open source Cloud Computing platforms have emerged as a valuable tool for geospatial analysis of image data from various satellites while the Google Earth Engine (GEE) platform is the most widespread in the scientific field of satellite RS. Newest launches of various satellites in combination with the GEE platform, facilitate in a great extent national-scale lake monitoring. Next step was to test the transferability and performance of hereby-developed WQ algorithms when employing Landsat (7 +ETM/8 OLI) and Sentinel 2 surface reflectance (SR) values embedded in GEE and subjected to different

atmospheric correction (AC) methods from those used as they were developed. More particularly, GEE-Landsat and -Sentinel 2 SR of year 2018 was retrieved from the WFD lake sampling stations while were atmospherically corrected by the methods LaSRC (Landsat 8 OLI), LEDAPS (Landsat 7 ETM+) and Sen2Cor (Sentinel 2). Those SR values (GEE) were matched with WQ *in-situ* data of 2018 within ±7 days (from sampling date) of satellite overpasses, while the same pairs were created with SR derived from manually downloaded and pre-processed, with AC DOS1 method, respective images. Empirical WQ models of Chl-*a*, Secchi depth and TP (for all and separately for natural and artificial lakes), were applied twice employing both types of SR (DOS1- and rest in GEE- corrected). Furthermore, double application of WQ models was conducted separately for Landsat (7 ETM+/OLI) and Sentinel 2 data. Double application of WQ models resulted in double quantifications of each studied WQ element in each sampling station while those double WQ values were inserted in a linear regression analysis. Yielded linear equations (corrected WQ models), for each sensor, were accompanied by strong associations (R^2 ranging from 0.68 to 0.98). Initial and corrected sensor-specific WQ models were validated based on available *in-situ* WQ datasets of 2019 and 2020. Sensor-specific correction of WQ models was proven essential for some of them while RMSE values ranged for Chl-*a* from 11.68 μg/l (Landsat) to 14.88 μg/l (Sentinel 2), for Secchi depth from 2.02 m (Landsat) to 2.57 m (Sentinel 2) and TP from 0.14 mg/l (Landsat) to 0.09 mg/l (Sentinel 2), values that confirmed the stability and transferability of hereby WQ models even when employ differently-from-DOS1 method corrected SR.

One more ambiguous question that has been examined is whether WQ universal models are efficient for WQ monitoring of oligotrophic Case-2 waters. The classification of waters in Case 1 (oceanic) and Case 2 (coastal regions, rivers, and lakes), is characterized by great importance; Case 1 waters are determined by phytoplankton and co-varying substances, while Case 2 waters are more complex concerning their composition and optical properties. Oligotrophic lakes are classified as Case-2 rather than Case-1 waters since they typically receive significant levels of terrigenous input and their water clarity is primarily controlled by the concentration of Dissolved Organic Carbon (DOC). In purview of the above, lake WQ models were applied to Landsat 8 OLI images, with available *in-situ* WQ data, illustrating two (2) Greek oligotrophic waterbodies, Trichonis and Amvrakia lakes. Conclusively, their application was ineffective: Chl-$a_{general}$ model yielded values of RMSE=1.9 μg/l, NRMSE=1.6 and median MAPE=256.8 %, Chl-$a_{natural}$ model yielded values of RMSE=1.8 μg/l, NRMSE=1.5 and median MAPE=176.6 % while results of all models of Secchi Disk and Total Phosphorus were statistically insignificant. Based on the previous approach, an effort has been made to

develop special designed WQ algorithms in Trichonis lake based on a dense sampling network (22 stations). Subsequently, the most statistically promising quantitative models – accrued from statistical elaboration of 2014 data- were applied to another satellite image of 2013 (with available *in-situ* WQ data). Results from the validation process showed a relatively variable statistical relationship between the *in-situ* data and reflectances (R logchl-*a*: 0.4, R NH_4^+: 0.7, R Chl-*a*: 0.5, R CDOM at 420 nm: 0.3). Hereby findings were concurrent with other studies in international literature, indicating that estimations for oligotrophic are less accurate than eutrophic and mesotrophic lakes, owing to the lack of suspended particles that are detectable by satellite sensors. Conclusively, background information required, suitable spectral bands and essential circumstances are described for the most optimal designation of the WQ monitoring in oligotrophic waterbodies.

Scientific area: Remote sensing, GIS and spatial analysis for inland water quality monitoring

Keywords: lake WQ, PCA, MLR analysis, trophic status, Case-2, WFD, Landsat, Sentinel 2, GEE, oligotrophic

Εκτίμηση της ποιότητας των υδάτων των ελληνικών λιμνών μέσω τηλεπισκόπησης και στατιστικής μοντελοποίησης

Τμήμα Αξιοποίησης Φυσικών Πόρων & Γεωργικής Μηχανικής
Εργαστήριο Εδαφολογίας & Γεωργικής Χημείας

ΠΕΡΙΛΗΨΗ

Στόχος της παρούσας διδακτορικής διατριβής είναι η επίτευξη συνεχούς παρακολούθησης και εκτίμησης της ποιότητας των υδάτων και της τροφικής κατάστασης των ελληνικών λιμνών αξιοποιώντας την εφαρμογή της Οδηγίας Πλαίσιο για τα Ύδατα (ΟΠΥ) στην Ελλάδα σε συνέργεια με τη δορυφορική Τηλεπισκόπηση, παρέχοντας παράλληλη υποστήριξη στη βιώσιμη διαχείριση των υδάτινων πόρων σε εθνική κλίμακα. Η συνεχής παρακολούθηση της ποιότητας υδάτων είναι το πιο κρίσιμο χαρακτηριστικό της ολοκληρωμένης διαχείρισης λιμναίων οικοσυστημάτων. Ως εκ τούτου, το μεθοδολογικό πλαίσιο που αναπτύχθηκε έχει ως απώτερο στόχο τη δημιουργία μοντέλων ποσοτικής εκτίμησης της ποιότητας υδάτων λιμνών ενώ η πρακτική εφαρμογή της μεθοδολογίας αξιολογήθηκε σε συνολικά 50 λίμνες (φυσικές και τεχνητές) την περίοδο 2013-2018, αποτελώντας το Εθνικό Δίκτυο Παρακολούθησης Λιμνών της Ελλάδας στο πλαίσιο της ΟΠΥ. Όσον αφορά στα δεδομένα δορυφορικής τηλεπισκόπησης, εικόνες από τους αισθητήρες Landsat 7 Enhanced Thematic Mapper Plus (ETM+) και Landsat 8 Operational Land Imager (OLI) συνδυάστηκαν με *in-situ* μετρήσεις ποιοτικών παραμέτρων (Χλωροφύλλη-*α* -Chl-*a*-, βάθος Secchi Disk, συγκέντρωση Ολικού Φωσφόρου,TP- Total Phosphorus) με κύριο στόχο την επίτευξη ανάπτυξης μοντέλων εκτίμησης της ποιότητας και της τροφικής κατάστασης υδάτων. Από στατιστικής άποψης, πραγματοποιήθηκε ανάλυση κύριων συνιστωσών (PCA-Principal Component Analysis) για να διερευνηθούν οι συσχετίσεις των συγκεντρώσεων της Chl-*a* των ελληνικών λιμνών με άλλα βασικά τους κριτήρια, όπως το εάν είναι φυσικές ή τεχνητές, την τυπολογία (ΟΠΥ), τον κλιματικό τύπο σύμφωνα με το σύστημα ταξινόμησης Köppen-Geiger, την εποχή δειγματοληψιών νερού και την ημερολογιακή διαφορά μεταξύ των *in-situ* δειγματοληψιών και του περάσματος του δορυφόρου. Η ανάλυση κύριων συνιστωσών υπέδειξε το χαρακτηριστικό φυσική/τεχνητή και την τυπολογία (ΟΠΥ) τους ως τις παραμέτρους που κυρίως συμβάλλουν στη διακύμανση της συγκέντρωσης της Chl-*a*.

Στη συνέχεια, διενεργήθηκαν αναλύσεις πολλαπλής παλινδρόμησης, με τη μέθοδο stepwise, μεταξύ διαφορετικών ομάδων δεδομένων, οι οποίες διαμορφώθηκαν με βάση τα κριτήρια της PCA. Βασικός στόχος ήταν η ανάπτυξη διαφορετικών αλγορίθμων εκτίμησης της συγκέντρωσης Chl-*a*,

προσανατολισμένοι σε διαφορετικούς τύπους λιμνών. Έπειτα, διεξήχθη ανάλυση συσχέτισης (correlation analysis) μεταξύ των *in-situ* μετρήσεων των μελετούμενων παραμέτρων ποιότητας με σκοπό τη διερεύνηση και ανίχνευση των συσχετίσεών τους. Με βάση τα αποτελέσματα της ανάλυσης συσχέτισης, διενεργήθηκαν περαιτέρω αναλύσεις παλινδρόμησης, χρησιμοποιώντας τις *in-situ* τιμές βάθους Secchi disk και TP για τη δημιουργία βέλτιστων μοντέλων όλων των κατηγοριών (όλες, φυσικές και τεχνητές λίμνες). Στη συνέχεια, διεξήχθη η ταξινόμηση της τροφικής κατάστασης των μελετούμενων λιμνών, υπολογίζοντας τον Δείκτη Τροφικής Κατάστασης (TSI) του Carlson για κάθε λίμνη, αρχικά για όλες και στη συνέχεια ξεχωριστά για τις φυσικές και τις τεχνητές. Η προτεινόμενη μεθοδολογία είχε ως αποτέλεσμα την ανάπτυξη ξεχωριστών μοντέλων για τις φυσικές (R = 0.78) και τεχνητές (R= 0.76) λίμνες, ενώ το γενικό μοντέλο που αναπτύχθηκε χωρίς κριτήρια αποδείχθηκε ασθενέστερο (R = 0.65) σε σύγκριση με τα υπόλοιπα. Οι αναλύσεις παλινδρόμησης μεταξύ των δεδομένων Landsat και βάθους Secchi οδήγησαν σε 3 μοντέλα: για όλες τις λίμνες ($Secchi_{general}$; R = 0.78; RMSE = 2 m), για τις φυσικές ($Secchi_{natural}$; R = 0.95; RMSE = 1.87 m) και για τις τεχνητές ($Secchi_{artificial}$, R = 0.62, RMSE = 1.36 m), με αξιόπιστη ακρίβεια. Τα ευρήματα έδειξαν ότι οι αναλύσεις παλινδρόμησης που σχετίζονται με τον Ολικό Φώσφορο απέτυχαν να αναπτύξουν ένα στατιστικά αποδεκτό μοντέλο για τις τεχνητές λίμνες. Ωστόσο, ανέδειξαν ένα ισχυρό για όλες τις λίμνες $TP_{general}$ (R = 0.71, RMSE = 0.008 mg/L) και για τις φυσικές $TP_{natural}$ (R = 0.93, RMSE = 0.018 mg/L).

Στη συνέχεια, όσον αφορά στα αποτελέσματα του δείκτη τροφικής κατάστασης TSI, η μεγαλύτερη απόκλιση των προβλεπόμενων τιμών σε σχέση με τις αντίστοιχες *in-situ* ανιχνεύθηκε σε τεχνητές και πιο ρηχές λίμνες (μέσο βάθος < 5 m), επιβεβαιώνοντας την ύπαρξη αξιοσημείωτων διαφοροποιήσεων μεταξύ φυσικών και τεχνητών υδάτινων σωμάτων. Τα επόμενα σημαντικά ερωτήματα τα οποία απάντησε η παρούσα διδακτορική διατριβή είναι αρχικά εάν οι προαναφερθέντες εμπειρικοί αλγόριθμοι παρακολούθησης της ποιότητας λιμνών που αναπτύχθηκαν με εικόνες Landsat μπορούν να εφαρμοστούν επιτυχώς σε εικόνες Sentinel 2 και στη συνέχεια εάν η συνδυασμένη χρήση εικόνων πολλαπλών αισθητήρων δύναται να βελτιώσει την ακρίβεια ποσοτικοποίησής τους. Προς αυτή την κατεύθυνση, χρησιμοποιήθηκαν εικόνες Sentinel 2 MSI της χρονιάς 2018 με ημερομηνίες ταυτόχρονες με αυτές των μετρήσεων πεδίου με σκοπό την αξιολόγηση της απόδοσης των εμπειρικών μοντέλων και τη σύγκριση των αποτελεσμάτων επικύρωσής τους με τα αντίστοιχα των δορυφόρων Landsat. Αναφορικά με τα αποτελέσματα, τα γενικά μοντέλα όλων των ποιοτικών παραμέτρων (Chl-*a*, βάθος Secchi Disk και TP), με εφαρμογή σε όλες τις λίμνες ήταν πιο αποτελεσματικά και ακριβή όταν εφαρμόστηκαν σε εικόνες Landsat, ενώ δεν παρατηρήθηκε βελτίωση

ύστερα από τη χρήση εικόνων πολλαπλών αισθητήρων. Τα μοντέλα Chl-*a* και TP παρουσίασαν μικρότερες τιμές σφάλματος όταν εφαρμόστηκαν σε εικόνες Sentinel 2 (RMSE Chl-*a*=16.4 μg/l vs 21.5 μg/l; RMSE TP=0.03 mg/l vs 0.031 mg/l) και μόνο το μοντέλο $Secchi_{natural}$ είχε καλύτερη απόδοση με δεδομένα Landsat (2.8 m vs 2.9 m).

Όσον αφορά στους αλγόριθμους των τεχνητών λιμνών, το μοντέλο της Chl-*a* απέδωσε καλύτερα ύστερα από τη χρήση Landsat εικόνων (RMSE= 3.7 μg/l vs 7.7 μg/l; Sentinel 2), ενώ το μοντέλο εκτίμησης βάθους Secchi Disk πέτυχε ελαφρώς καλύτερη απόδοση χρησιμοποιώντας εικόνες Sentinel 2 (RMSE= 1.5 m vs 1.6 m; Landsat). Τη χειρότερη απόδοση ανάμεσα στις υπό μελέτη ποιοτικές παραμέτρους παρουσίασαν τα μοντέλα της Chl-*a* (οι διάμεσες MAPE τιμές κυμάνθηκαν από 42 % έως 58% ενώ του βάθους Secchi από 24% έως 44% και του Ολικού Φωσφόρου από 22% έως 38%), γεγονός που επιβεβαιώνει για ακόμα μία φορά την πολυπλοκότητα που κρύβει η χαρτογράφηση της Chl-*a* στα ύδατα της Περίπτωσης 2 (παράκτια, εσωτερικά ύδατα).

Σήμερα, οι πλατφόρμες υπολογιστικής νέφους ανοιχτού κώδικα έχουν αναδειχθεί σε πολύτιμα εργαλεία γεωχωρικής ανάλυσης δεδομένων εικόνας διαφόρων δορυφόρων, ενώ η πλατφόρμα Google Earth Engine (GEE) είναι η πιο διαδεδομένη στο επιστημονικό πεδίο της δορυφορικής τηλεπισκόπησης. Οι πρόσφατες εκτοξεύσεις διαφόρων δορυφόρων σε συνδυασμό με την πλατφόρμα GEE, διευκολύνουν σε μεγάλο βαθμό την παρακολούθηση των λιμνών σε εθνική κλίμακα. Με βάση τα παραπάνω, το επόμενο βήμα ήταν η εξακρίβωση της απόδοσης των ποιοτικών μοντέλων όταν χρησιμοποιούν τιμές ανάκλασης οι οποίες έχουν υποβληθεί σε διαφορετικές μεθόδους ατμοσφαιρικής διόρθωσης από αυτές όταν αναπτύχθηκαν. Πιο συγκεκριμένα, τιμές ανάκλασης από εικόνες Landsat και Sentinel 2 του έτους 2018, αποκτήθηκαν μέσω της πλατφόρμας GEE στους σταθμούς δειγματοληψίας λιμνών της ΟΠΥ ενώ ήταν ατμοσφαιρικά διορθωμένες με τις μεθόδους LaSRC (Landsat 8 OLI), LEDAPS (Landsat 7 ETM+) και Sen2Cor (Sentinel 2). Οι εν λόγω τιμές ανάκλασης (GEE) αντιστοιχίστηκαν με τα επιτόπια δεδομένα ποιότητας υδάτων του 2018 με διαφορά ημερομηνιών ±7 ημερών (από την ημερομηνία δειγματοληψίας), ενώ τα ίδια ζεύγη δημιουργήθηκαν με τις τιμές ανάκλασης προερχόμενες από τις αντίστοιχες εικόνες, οι οποίες αποκτήθηκαν και προ-επεξεργάστηκαν χειροκίνητα, με τη μέθοδο της ατμοσφαιρικής διόρθωσης DOS1. Τα εμπειρικά μοντέλα Chl-*a*, βάθους Secchi Disk και TP (για όλες και ξεχωριστά για φυσικές και τεχνητές λίμνες), εφαρμόστηκαν δύο φορές χρησιμοποιώντας και τα δύο είδη ανάκλασης, την DOS1-διορθωμένη ανάκλαση και εκείνη που αποκτήθηκε από εικόνες ενσωματωμένες στη GEE πλατφόρμα. Επιπρόσθετα, η διπλή εφαρμογή των εμπειρικών μοντέλων εκτίμησης της ποιότητας υδάτων

πραγματοποιήθηκε ξεχωριστα για τους δορυφόρους Landsat (7 ETM+/OLI) και Sentinel 2. Η διπλή εφαρμογή των ποιοτικών μοντέλων είχε ως αποτέλεσμα τη διπλή εκτίμηση των τιμών της εκάστοτε ποιοτικής παραμέτρου στον εκάστοτε σταθμό δειγματοληψίας και στη συνέχεια αυτές οι διπλές τιμές (προερχόμενες από DOS1 και λοιπές ατμοσφαιρικές μεθόδους) εισήχθησαν σε ανάλυση γραμμικής παλινδρόμησης. Οι γραμμικές εξισώσεις που προέκυψαν μεμονωμένα για κάθε αισθητήρα (Landsat, Sentinel 2) υπέδειξαν τιμές υψηλής συσχέτισης (R^2 από 0.68 έως 0.98). Η διόρθωση των εμπειρικών μοντέλων ανάλογα με τον αισθητήρα πραγματοποιήθηκε με βάση τα διαθέσιμα *in-situ* δεδομένα ποιότητας του 2019 και του 2020 ενώ αποδείχθηκε απαραίτητη για ορισμένα από αυτά, με τις τιμές RMSE να κυμαίνονται για τη Chl-*a* από 11.68 μg/l (Landsat) έως 14.88 μg/l (Sentinel 2), για το βάθος Secchi Disk από 2.02 m (Landsat) έως 2.57 m (Sentinel 2) και TP από 0.14 mg/l (Landsat) έως 0.09 mg/l (Sentinel 2), τιμές που επιβεβαίωσαν τη σταθερότητα και τη δυνατότητα μεταφοράς των εμπειρικών μοντέλων σε εικόνες διαφορετικά ατμοσφαρικά διορθωμένες, στην πλατφόρμα GEE.

Ένα ακόμη περίπλοκο ερώτημα που έχει εξεταστεί είναι εάν τα μοντέλα εκτίμησης ποιότητας υδάτων (Chl-*a*, βάθους Secchi Disk και Ολικού Φωσφόρου) δύναται να είναι αποτελεσματικά για την παρακολούθηση της ποιότητας των ολιγοτροφικών υδάτων. Η ταξινόμηση των υδάτων στην Περίπτωση 1 (ωκεάνια) και στην Περίπτωση 2 (παράκτιες περιοχές, ποτάμια και λίμνες), είναι ιδιαίτερα σημαντική. Τα νερά της Περίπτωσης 1 προσδιορίζονται με βάση το φυτοπλαγκτόν και λοιπές ουσίες, ενώ τα νερά της Περίπτωσης 2 είναι πιο πολύπλοκα όσον αφορά στη σύσταση και τις οπτικές τους ιδιότητες. Οι ολιγοτροφικές λίμνες αποτελούν μια ξεχωριστή κατηγορία των περίπλοκων οπτικά υδάτων της Περίπτωσης 2 και δε ταξινομούνται στην Περίπτωση 1 δεδομένου ότι συνήθως λαμβάνουν σημαντικές εισροές ιζημάτων και η διαύγεια των υδάτων τους ελέγχεται συνήθως από τις συγκεντρώσεις του διαλυμένου οργανικού άνθρακα (DOC-Dissolved Organic Carbon) και όχι από το φυτοπλαγκτόν. Με βάση τα προαναφερθέντα, τα μοντέλα εκτίμησης ποιότητας υδάτων λιμνών (Chl-*a*, βάθους Secchi Disk και TP), εφαρμόστηκαν σε εικόνες Landsat 8 OLI που απεικονίζουν δύο (2) ελληνικά ολιγοτροφικά υδάτινα σώματα, τις λίμνες Τριχωνίδα και Αμβρακία. Συμπερασματικά, η εφαρμογή των μοντέλων σε αυτές τις λίμνες ήταν αναποτελεσματική· το γενικό μοντέλο υπολογισμού της Χλωροφύλλης-α απέδωσε τιμές RMSE=1.9 μg/l, NRMSE=1.6 και median MAPE=256.8, το μοντέλο υπολογισμού Χλωροφύλλης-α στις φυσικές λίμνες απέδωσε: RMSE=1.8 μg/l, NRMSE=1.5 και median MAPE=176.6 ενώ τα αποτελέσματα των λοιπών μοντέλων Ολικού Φωσφόρου και βάθους Secchi Disk δεν ήταν στατιστικά σημαντικά.

Λαμβάνοντας υπόψιν τα προαναφερθέντα αποτελέσματα, έγινε μια προσπάθεια ανάπτυξης μοντέλων εκτίμησης της ποιότητας υδάτων στη λίμνη Τριχωνίδα με βάση ένα πυκνό δίκτυο δειγματοληψίας (22 σταθμοί). Στη συνέχεια, τα βέλτιστα μοντέλα (που αναπτύχθηκαν με βάση τα δορυδορικά και επιτόπια δεδομένα του 2014) εφαρμόστηκαν στη δορυφορική εικόνα του 2013 και η επικύρωση των αποτελεσμάτων πραγματοποιήθηκε χρησιμοποιώντας τα αντίστοιχα *in-situ* δεδομένα του 2013. Τα αποτελέσματα από τη διαδικασία επικύρωσης έδειξαν μια σχετικά διαφορετική στατιστική σχέση μεταξύ των *in-situ* δεδομένων και των ανακλάσεων (R logchl-*a*: 0.4, R NH_4^+: 0.7, R Chl-*a*: 0.5, R CDOM στα 420 nm: 0.3). Ως εκ τούτου, τα ευρήματα ταυτίζονται με με αυτά άλλων μελετών στη διεθνή βιβλιογραφία, υποδεικνύοντας ότι η παρακολούθηση της ποιότητας υδάτων στις ολιγοτροφικές λίμνες είναι λιγότερο ακριβής σε σχέση με τις ευτροφικές και μεσοτροφικές, λόγω της απουσίας αιωρούμενων σωματιδίων τα οποία είναι ανιχνεύσιμα από τους δορυφορικούς αισθητήρες. Επιπλέον περιγράφονται οι απαιτούμενες πληροφορίες υποβάθρου, οι κατάλληλες φασματικές ζώνες και οι βασικές συνθήκες για τον βέλτιστο σχεδιασμό των μεθόδων παρακολούθησης ποιότητας υδάτων σε ολιγοτροφικά υδάτινα σώματα.

Επιστημονική περιοχή: Τηλεπισκόπηση, GIS και χωρική ανάλυση για την παρακολούθηση της ποιότητας των εσωτερικών υδάτων

Λέξεις-κλειδιά: ποιότητα υδάτων λιμνών, PCA, ανάλυση MLR, τροφική κατάσταση, Περίπτωση-2, ΟΠΥ, Landsat, Sentinel 2, GEE, ολιγοτροφικό

ΕΠΙΤΡΟΠΕΣ ΔΙΔΑΚΤΟΡΙΚΗΣ ΔΙΑΤΡΙΒΗΣ

Επιβλέπων καθηγητής

Δρ. Διονύσιος Καλύβας, Καθηγητής
Τμήμα Αξιοποίησης Φυσικών Πόρων και Γεωργικής Μηχανικής
Γεωπονικό Πανεπιστήμιο Αθηνών

Μέλη συμβουλευτικής επιτροπής

Δρ. Ηλίας Δημητρίου, Ερευνητής Α΄
Ινστιτούτο Θαλάσσιων Βιολογικών Πόρων και Εσωτερικών Υδάτων
Ελληνικό Κέντρο Θαλάσσιων Ερευνών

Δρ. Πετρόπουλος Γεώργιος Π., Επίκουρος Καθηγητής
Τμήμα Γεωγραφίας
Χαροκόπειο Πανεπιστήμιο

Μέλη εξεταστικής επιτροπής

Δρ. Γεώργιος Παπαδόπουλος, Αναπληρωτής καθηγητής
Τμήμα Αγροτικής Οικονομίας και Ανάπτυξης
Γεωπονικό Πανεπιστήμιο Αθηνών

Δρ. Εμμανουήλ Ψωμιάδης, Επίκουρος Καθηγητής
Τμήμα Αξιοποίησης Φυσικών Πόρων και Γεωργικής Μηχανικής
Γεωπονικό Πανεπιστήμιο Αθηνών

Δρ. Κωνσταντίνος Σούλης, Επίκουρος Καθηγητής
Τμήμα Αξιοποίησης Φυσικών Πόρων και Γεωργικής Μηχανικής
Γεωπονικό Πανεπιστήμιο Αθηνών

Δρ. Ιωάννης Καραούζας, Ερευνητής Γ'
Ινστιτούτο Θαλάσσιων Βιολογικών Πόρων και Εσωτερικών Υδάτων
Ελληνικό Κέντρο Θαλάσσιων Ερευνών

The approval of the present thesis by the Examining Committee and the Department does not presuppose the acceptance of the author's views **(*Law* 5343/1932, *Article* 202).**

*Η έγκριση της παρούσας διατριβής από την Εξεταστική Επιτροπή και το Τμήμα Αξιοποίησης Φυσικών Πόρων και Γεωργικής Μηχανικής του Γεωπονικού Πανεπιστημίου Αθηνών δεν προϋποθέτει και την αποδοχή των απόψεων του συγγραφέα (**Νόμος** 5343/1932, **Άρθρο** 202, παρ. 2).*

ΥΠΕΥΘΥΝΗ ΔΗΛΩΣΗ

Δηλώνω ότι είμαι η αποκλειστική συγγραφέας της παρούσας Διδακτορικής Διατριβής με τίτλο: «Water Quality Assessment in Greek Lakes by Using Remote Sensing and Statistical Modelling», η οποία είναι πρωτότυπη και δεν έχει αντιγραφεί (μέρος της ή ολόκληρη) από άλλη δημοσιευμένη ή αδημοσίευτη πνευματική εργασία. Βεβαιώνω ότι έχω τηρήσει τους κανόνες περί λογοκλοπής του τμήματος Αξιοποίησης Φυσικών Πόρων και Γεωργικής Μηχανικής του Γεωπονικού Πανεπιστημίου Αθηνών. Επιπρόσθετα, βεβαιώνω ότι έχω τηρήσει απαρέγκλιτα όσα ο νόμος ορίζει περί πνευματικής ιδιοκτησίας και έχω συμμορφωθεί πλήρως με τα προβλεπόμενα στο νόμο περί προστασίας προσωπικών δεδομένων και τις αρχές Ακαδημαϊκής Δεοντολογίας.

DECLARATION

I declare that I am the exclusive author of this PhD Thesis entitled: "Water Quality Assessment in Greek Lakes by Using Remote Sensing and Statistical Modelling", which is original and has not been copied (in part or in whole) from other published or unpublished intellectual work. I certify that I have complied with the plagiarism rules of the Department of Natural Resources Utilization and Agricultural Engineering of the Agricultural University of Athens. In addition, I confirm that I have fully complied with the provisions of the law on intellectual property rights, personal data protection and the principles of Academic Ethics.

ΕΥΧΑΡΙΣΤΙΕΣ

Σε αυτό το σημείο θα ήθελα να εκφράσω τις ευχαριστίες και τη βαθειά ευγνωμοσύνη μου σε όλους τους ανθρώπους που στάθηκαν πλάι μου και συνέβαλλαν με ποικίλους τρόπους στην ολοκλήρωση της παρούσας διδακτορικής διατριβής. Η περίοδος εκπόνησης της διατριβής μου αποδείχθηκε κάποιες φορές "επίπονα" συναρπαστική όχι μόνο γιατί πολύ σημαντικά της τμήματα ολοκληρώθηκαν κατά τη διάρκεια της έξαρσης της πανδημίας του κορονοϊού (COVID-19) -με ό,τι αυτό συνεπάγεται- αλλά και γιατί συνέπεσε με αξιοσημείωτα γεγονότα της προσωπικής και οικογενειακής μου ζωής. Η ολοκλήρωση της διατριβής μου αποτελεί λοιπόν την πιο σημαντική στιγμή της ζωής μου μετά την απόκτηση του παιδιού μου.

Αρχικά, νιώθω την ανάγκη να ευχαριστήσω τον επιβλέποντα Καθηγητή του Γεωπονικού Πανεπιστημίου Αθηνών, Δρ. Διονύσιο Καλύβα, ο οποίος με εμπιστεύτηκε και με καθοδήγησε με ενθουσιασμό, φιλοδοξία και υπομονή κατά τη διάρκεια όλων των διαφορετικών σταδίων εκπόνησης αυτής της διατριβής. Αισθάνομαι ιδιαίτερα τυχερή που με αφορμή τη φοίτησή μου στο μεταπτυχιακό πρόγραμμα που συντονίζει με τίτλο «Εφαρμογές Γεωπληροφορικής στους Φυσικούς Πόρους», διασταυρώθηκαν οι δρόμοι μας καθώς είμαι σίγουρη ότι απέκτησα ακόμα ένα σύμμαχο στη μελλοντική μου επιστημονική πορεία (και όχι μόνο).

Τις θερμές μου ευχαριστίες θα ήθελα επίσης να εκφράσω και στο προσωπικό της ερευνητικής μονάδας Γεωγραφικών Πληροφοριακών Συστημάτων (ΕΜΓΠΣ) του Γεωπονικού Πανεπιστημίου Αθηνών καθώς υπήρξαν πάντα αρωγοί σε οποιοδήποτε εμπόδιο προέκυπτε είτε σε επιστημονικό είτε σε επίπεδο τεχνογνωσίας χρήσης ειδικών λογισμικών.

Στη συνέχεια θα ήθελα να ευχαριστήσω τον Δρ. Ηλία Δημητρίου, Ερευνητή Α' του Ελληνικού Κέντρου Θαλάσσιων Ερευνών και ένα από τα

μέλη της συμβουλευτικής επιτροπής του διδακτορικού μου. Αισθάνομαι ότι δεν υπάρχουν λόγια για να περιγράψω την ευγνωμοσύνη και την εκτίμησή μου καθώς ο Δρ. Ηλίας Δημητρίου αποτελεί τον πραγματικό μου μέντορα και είναι εκείνος ο άνθρωπος που διεύρυνε τους επιστημονικούς μου ορίζοντες και ουσιαστικά με «δίδαξε» πώς να αγαπώ και να αξιοποιώ την επιστήμη για το κοινό καλό. Ήταν πάντα δίπλα μου να με καθοδηγήσει και να με βοηθήσει να ξεπεράσω οποιοδήποτε σκόπελο εμπόδιζε την εξέλιξη της διατριβής μου, χαρίζοντάς μου παράλληλα το «πολύτιμο» δώρο της κριτικής σκέψης.

Ιδιαίτερες ευχαριστίες οφείλω στον Δρ. Γεώργιο Πετρόπουλο - Επίκουρο Καθηγητή του Χαροκόπειου Πανεπιστημίου- για την αμέριστη υποστήριξη, καθοδήγηση και υπομονή που υπέδειξε καθόλη τη διάρκεια εκπόνησης της διδακτορικής μου διατριβής. Επίσης τον ευχαριστώ θερμά για τα ιδιαιτέρως πολύτιμα σχόλιά του κατά τη διάρκεια προετοιμασίας των επιστημονικών δημοσιεύσεων που προέκυψαν από την παρούσα διατριβή καθώς συνετέλεσαν στην ταχύτερη ολοκλήρωση της διαδικασίας της κρίσης τους. Πέρα όμως από αυτό, ο Δρ. Γεώργιος Πετρόπουλος ήταν ο άνθρωπος που πρώτος μου δίδαξε το πεδίο της δορυφορικής τηλεπισκόπησης -ως καθηγητής στο μεταπτυχιακό πρόγραμμα που παρακολούθησα- και εκ των πραγμάτων με έφερε πρώτη φορά σε επαφή με το αντικείμενο στο οποίο έμελλε αργότερα να αφιερώσω το μεγαλύτερο μέρος των επαγγελματικών μου σκέψεων και ανησυχιών.

Επιπρόσθετα, θα ήθελα να ευχαριστήσω τους Δρ. Γεώργιο Παπαδόπουλο - Αναπληρωτή Καθηγητή-, Δρ. Εμμανουήλ Ψωμιάδη - Επίκουρο Καθηγητή- και Δρ. Κωνσταντίνο Σούλη-Επίκουρο Καθηγητή- του Γεωπονικού Πανεπιστημίου Αθηνών και τον Δρ. Ιωάννη Καραούζα - Ερευνητή Γ'- του Ελληνικού Κέντρου Θαλάσσιων Ερευνών, αρχικά για την αποδοχή τους να συμμετάσχουν ως μέλη της εξεταστικής επιτροπής της

διαδακτορικής μου διατριβής αλλά και για το χρόνο που αφιέρωσαν για την ανάγνωσή της και την παροχή των ωφέλιμων σχόλιών τους.

Επιπρόσθετα, θα ήθελα να εκφράσω τις ευχαριστίες μου στο προσωπικό του Ελληνικού Κέντρου Βιοτόπων-Υγροτόπων (ΕΚΒΥ) του Μουσείου Γουλανδρή Φυσικής Ιστορίας για τον αποδεδειγμένο επαγγελματισμό τους σχετικά με τη συλλογή και διανομή των άρτια επεξεργασμένων *in-situ* δεδομένων ποιότητας υδάτων. Η παρακολούθηση των 50 λιμνών λειτουργεί αδιάλειπτα από το ΕΚΒΥ από το 2012 και εντάσσεται στο Εθνικό Δίκτυο Παρακολούθησης της ποιότητας και της ποσότητας των υδάτων που συγκροτήθηκε με Κοινή Υπουργική Απόφαση (ΚΥΑ 140384/2011) στο πλαίσιο εφαρμογής της Οδηγίας για τα Ύδατα.

Σε αυτό το σημείο δε θα μπορούσα να μην ευχαριστήσω και να μην εκφράσω την αμέριστη εκτίμησή μου για όλους τους συν-Αδέρφους και καλούς/ές φίλους/ες που έχω αποκτήσει τόσα χρόνια στον Τομέα Εσωτερικών Υδάτων (αλλά και στα λοιπά Ινστιτούτα) του ΕΛΚΕΘΕ. Η καθημερινή μας επαφή τα τελευταία τουλάχιστον 10 χρόνια συνεισέφερε όχι μόνο στην εξέλιξη της επαγγελματικής μου πορείας αλλά και στην απόκτηση μιας ζεστής ανθρώπινης αλληλεπίδρασης, η οποία ήταν πάντα αρωγός στη ψυχολογική μου ευημερία και στην ολοκλήρωση του εν λόγω επιστημονικού εγχειρήματος.

Απεριόριστα ευγνώμων και τυχερή αισθάνομαι και για τους καλούς/ές και πιστούς/ές μου φίλους/ες που με συντροφεύουν στη ζωή για πάνω από 20 χρόνια στηρίζοντας με σε όλες τις χαρές και λύπες˙ χωρίς αυτούς/ές, η εξέλιξή μου θα ήταν εντελώς διαφορετική.

Η μεγαλύτερη ευγνωμοσύνη μου ανήκει δικαιωματικά στην οικογένειά μου, η οποία με στηρίζει και με ενδυναμώνει αδιάκοπα από τότε που θυμάμαι τον εαυτό μου. Μητέρα και Αντιγόνη σας ευχαριστώ ιδιαίτερα για τη συνεχή ενθάρρυνση, συναισθηματική υποστήριξη και την άνευ όρια αγάπη σας.

ACKNOWLEGMENTS

At this point, I would like to express my thanks and deep gratitude to all the people who stood by me and, in their way, contributed to the completion of this doctoral thesis. The period of my PhD dissertation proved to be, at times, "painfully" exciting, not merely because essential parts of it were completed during the coronavirus pandemic outbreak (COVID-19) - with all that it entailed- but also because it coincided with significant events in my personal and family life. Completing my thesis is, therefore, the most important moment in my life since the birth of my child.

First and foremost, I am extremely grateful to the supervising Professor of the Agricultural University of Athens, Dr. Dionyssios Kalivas, who trusted me and guided me with ongoing enthusiasm, ambition and patience throughout the different stages of my thesis preparation. I feel particularly lucky and grateful that my studies in the master's program "Geoinformatics in Natural Resources", which he coordinates, allowed our paths to cross and I am confident that I gained a supporter in my future scientific career (and not only).

My warm thanks further extend to the staff of the Geographical Information Systems (GIS) Research Unit at the Agricultural University of Athens, as they have always been helpful in tackling any obstacle that arose, either at scientific level or in the use of special softwares.

Next, I would like to thank Dr. Elias Dimitriou, Research Director at the Hellenic Center for Marine Research and member of my PhD Supervisory Committee. I feel that words are not enough to describe my gratitude and deep appreciation. Dr. Dimitriou is my true mentor, as he is the one who broadened my scientific horizons and essentially "taught" me how to love and use science for the common good. He was always by my side, to guide me

and help me overcome any obstacles that hindered the development of my thesis, while gifting me with "precious" critical thinking skills.

I would like to extend special thanks to Dr. George Petropoulos - Assistant Professor at Harokopeio University - for his unwavering support, guidance and patience throughout the preparation of my doctoral thesis. I also sincerely thank him for his extremely valuable comments during the preparation of all the scientific publications resulting from this thesis, as they contributed to the faster completion of their review process. Beyond these contributions, Dr. Petropoulos was the one who first introduced me to the field of satellite remote sensing - as a professor in the postgraduate program I attended. He practically brought me into contact, for the first time, with the subject matter to which I would later devote most of my professional thoughts and concerns.

Additionally, I would like to thank Dr. Georgios Papadopoulos (Associate Professor), Dr. Emmanouil Psomiadis (Assistant Professor), and Dr. Konstantinos Soulis (Assistant Professor) at the Agricultural University of Athens, as well as Dr. Ioannis Karaouzas (Associate Researcher) at the Hellenic Centre for Marine Research, initially for their acceptance to participate as members on the Examination Committee of my doctoral thesis but also for the time they devoted to read my thesis and their useful comments and feedback.

In addition, I would like to extend my thanks to the staff of the Greek Biotope/Wetland Centre (or EKBY by its Greek initials), branch of the Goulandris Natural History Museum, for their proven professionalism in collecting and distributing well-processed *in-situ* water quality data. The monitoring of the 50 lakes has been operating by EKBY continuously since 2012, and is part of the National Water Quality and Quantity Monitoring Network established by Joint Ministerial Decision (KYA 140384/2011) in the context of the implementation of the Water Framework Directive.

At this stage, I could not but thank and express my undivided appreciation for all the fellow colleagues and good friends that I have acquired over the years at the Institute of Marine Biological Resources and Inland Waters (but also in the other Institutes) at HCMR. Our daily contact over the last 10 years at least, has contributed not only to the development of my professional career but also to the creation of a warm, personal interaction, which has always been helpful to my psychological well-being and the completion of this scientific undertaking.

I feel infinitely grateful and lucky for my good and faithful friends, who have accompanied me in life for over 20 years, supporting me through all the joys and sorrows; without them, my development would have been completely different.

My greatest gratitude, however, rightfully goes to my family, who have supported and encouraged me unceasingly for as long as I can remember. Mother and Antigoni, thank you very much for your constant encouragement, emotional support and unlimited love.

"Our future lies with today's kids and tomorrow's space exploration"

Sally Ride

(Astronaut and Physicist, the first American woman in space)

Στον γιο μου Οδυσσέα και

στη μνήμη του πολυαγαπημένου μου πατέρα Βασίλη

καθώς και οι δύο μαζί, ήταν, είναι και πάντα θα είναι

οι φωτεινοί φάροι της ζωής μου

Table of Contents

I. LIST OF ABBREVIATIONS AND ACRONYMS

Abbreviation	Meaning
ALI	Advanced Land Imager
ALOS	Advanced Land Observation Satellite
ANN	Artificial Neural Networks
AOP	Apparent Optical properties
ASTER	Advanced Spaceborne Thermal Emission and Reflection Radiometer
avg	average
CDOM	Colored Dissolved Organic Matter
Chl-*a*	Chlorophyll-*a*
CI	Cyanobacteria Index
CI	Condition Iindices
DN	Digital Number
DOC	Dissolved Organic Carbon
DOM	Dissolved Organic Matter
DOS1	Dark Object Subtraction 1
EC	European Commission
EKBY	Goulandris Natural History Museum, Greek Biotope/Wetland Centre
EO	Earh Observation
EPA	(U.S.) Environmental Protection Agency
ESA	European Space Agency
ETM+	Enhanced Thematic Mapper Plus
EVI	Enhanced Vegetation Index
FLH	Fluorescence Line Height
GA	Genetic Algorithms
GCP	Ground Control Points
GEE	Google Earth Engine
GIS	Geographical Information Systems
GNDVI	Green Normalized Difference Vegetation Index
HCMR	Hellenic Centre for the Marine Research
IDW	Inverse Distance Weighted
IOP	Inherent Optical Properties
KML	Keyhole Markup Language
L8	Landsat 8
LaSRC	Land Surface Reflectance Code
LEDAPS	Ecosystem Disturbance Adaptive Processing System
LOQ	Limit of Quantitation
LULC	Land Use/Land Cover

MAPE	Mean Absolute Percentage Error
MCI	Maximum Chlorophyll Index
ME	Mean Error
MERIS	Medium Resolution Imaging Spectrometer
ML	Machine learning
MLRs	Multiple Regressions
MNDWI	Modified Normalized Difference Water Index
MODIS	Moderate Resolution Imaging Spectroradiometer
MSFD	Marine Strategy Framework Directive
MSI	Multispectral Instrument
MSS	Multi-Spectral Scanner
N	Nitrogen
NASA	National Aeronautics and Space Administration
NDCI	Normalized Difference Chlorophyll Index
NDVI	Normalized Difference Vegetation Index
NDWI	Normalized Difference Water Index
NIR	Near Infrared
NRMSE	Normalized Root Mean Square Error
NRVI	Normalized Vegetation Index
OACs	Optically Active Components
OLI	Operational Land Imager (Landsat 8)
OLI-2	Operational Land Imager-2 (Landsat 9)
OWTs	Optical Water Types
P	Phosphorus
PC	Phycocyanin
PCA	Principal Component Analysis
Q-GIS	Quantum Geographic Information System
RF	Random Forest
rho	Pearson correlation coefficient
RMSE	Root Mean Square Error
RS	Remote Sensing
SABI	Surface Algal Bloom Index
SCL	Scan Line Corrector
SCP	Semi-automatic Classification Plugin
SD	Secchi Depth
SDD	Secchi Disk Depth
SDT	Secchi Disk Transparency
Sen2Cor	Sentinel 2 Correction (algorithm)
SPM	Suspended Particulate Matter
SPOT	Satellite Pour l' Observation de la Terre
SR	Surface Reflectance

SVM	Support Vector Machine
SWIR	Short-wave infrared
TIR	Thermal Infrared Sensor
TM	Thematic Mapper
TN	Total Nitrogen
TOA	Top of Atmosphere (reflectance)
TP	Total Phosphorus
TSI	(Carlson's) Trophic State Index
TSM	Total Suspended Matter
TSS	Total Suspended Solids
USGS	United State Geological Survey
VIF	Variance Inflation Factor
WFD	Water Framework Directive
WQ	Water Quality

II. GLOSSARY

Case 1 waters: Waters which optical properties are mainly determined by phytoplankton (oceanic).

Case 2 waters: The optically complex waters. The optical signal is dominated by phytoplankton but also by particulate inorganic matter and CDOM (coastal regions, rivers, and lakes).

Lacustrine waters: Freshwater lakes

Water clarity: A measure of underwater visibility, influenced by turbidity and color.

Lake trophic status/state: Biological condition dependent on various factors (nitrogen, phosphorus, pH, turbidity, color etc).

In-situ (water sampling): In the natural/original position/place (water sampling).

Multi-sensor images: Images from multiple satellite sensors.

Absorption wavelength: The wavelength at which a water sample absorbs light depending on the ion or molecule of component's composition.

Spectral Reflectance: The spectral fraction of light reflected by a surface.

Spectral bands: Specific portions of the electromagnetic spectrum of eflected light.

Backscatter: The scattering of radiation in a direction opposite to that of the incident radiation due to reflection from particles of the medium traversed.

Transboundary lakes: Lakes shared by two or more countries.

Eutrophication: Process of increased productivity of a lake (as it ages) which is greatly accelerated by human activities (increase in nutrients), resulting in an increase in biological production.

Surveillance monitoring stations (WFD): Member States must monitor at least for a period of a year for parameters indicative of all biological, hydromorphological and general physico-chemical quality elements.

Operational monitoring stations (WFD): Member States are required to monitor for those biological and hydromorphological quality elements most sensitive to the pressures to which the body or bodies are subject. Operational monitoring must use parameters relevant to the assessment of the effects of the pressures placing the body at risk.

***Spatial resolution*:** Refers to the size of the smallest feature that can be detected by a satellite sensor or displayed in a satellite image, is represented in pixels and noted as how many meters that pixel represents.

***Radiometric resolution*:** The amount of information in each pixel, that is the number of bits representing the energy recorded.

***Spectral resolution*:** The number and size of bands in the electromagnetic spectrum that a remote sensing platform can capture.

***Temporal resolution*:** Refers to the frequency at which imagery is recorded for a particular area.

***Solar zenith angle*:** The angle measured from the local zenith and the line of sight of the sun.

Optically shallow waters (OSWs): The bottom signals can be reflected in the water-leaving radiance and remote sensing reflectance (Rrs(l)) signatures. For clean waters, OSWs are those with depths <20 m, and turbid waters with depths 1–3 m.

Spectral indices: Mathematical equations (combinations) employing two or more spectral bands (wavelengths) of an image per pixel.

Image pre-processing: Radiometric, atmospheric and geometric corrections of raw remotely sensed image data.

Band stacking: The process of combining multiple bands into a single image file.

Time window: Date difference between field sampling and satellite overpass.

Path/Row: Row refers to the latitudinal center line of a frame of imagery. As the satellite moves along its path, the satellite instruments are continuously scanning the terrain below. These will be squares centered on the orbital path, but tilted clockwise when viewed on the UTM projection used for the distributed data.

Resampling procedure: Technique of transforming an image by recalculating its pixel values to be fitted to another image.

Focal statistics tool: Performs an operation that calculates a statistic (mean, maximum, or sum) for all input cells within a set of overlapping neighborhoods and within each neighborhood.

Multicollinearity: The occurrence of high intercorrelations among two or more independent variables in a multiple regression model.

Water specular reflection: The term used to describe 'mirror-like' reflection, from the surface of water (angle of reflection equals angle of incidence).

Spatial interpolation: Predicts values for cells in a raster from a limited number of sample data points. It can be used to predict unknown values for any geographic point data, such as chemical concentrations.

WQ Optically Active Components: Phytoplankton (Chl-*a*), Secchi Disk Depth, Temperature, CDOM, TOC (Total Organic Carbon), TSM (Total Suspended Matters, Turbidity, Sea Surface salinity and Electrical Conductivity. Components that interact with light and change the energy spectrum of reflected solar radiation from waterbodies.

III. PhD in numbers

PhD in numbers…

Satellite images
Landsat (7 ETM+ and 8 OLI): **320**
Sentinel 2: **44**

***In-situ* WQ measurements**
Chlorophyll-*a*: **1000**
Secchi depths: **1343**
Total Phosphorus: **726**

Sampling stations (WFD): 53
Surveillance: **27**
Operational: **26**

Study area: 50 lakes
Natural: **24**
Reservoirs: **26**
Transboundary: **3**

Publications: 4
Journal papers: **3**
Conference papers: **1**

IV. PhD OUTLINE

This PhD Thesis is divided into one (1) **Introductory section** and four (4) **main chapters** (Figure 1).

General Introduction describes the background knowledge required and the state-of-the-art methodologies applied, concerning the main herein attempted task; the monitoring of lake water quality (WQ) and trophic status through satellite remote sensing (RS). Special reference is made to: (1) the great lakes' significance for all living organisms, the necessity of continuous lake WQ monitoring for their sustainable management due to ongoing human pressures and by extension to lakes' greatest threat; the so-called eutrophication, (2) traditional and latest scientific trends of lake WQ monitoring, (3) recent developments in geoinformation technologies and the contribution of RS in WQ monitoring, (4) Water Framework Directive (WFD) instructions and requirements concerning the monitoring of lakes' WQ, (5) Earh Observation (EO) data, WQ key indicators and the contribution of Landsat mission to efficient lake WQ monitoring, (6) traditionally used methodologies for lake WQ monitoring through satellite RS, (7) description of optically and non-optically active WQ constituents measured by RS accompanied by the most utilized methodologies, (8) the high contribution, accrued from the combination of various multi-spectral sensors, to acquirement of high-frequency lake WQ time series and performance of multi-temporal analyses, (9) the development of Big Eath Data Cloud Processing Platforms and in particular the significance of Google Earth Engine (GEE) cloud-based platform to large-scale lake WQ assessment and long-time-series analyses, (10) background information about the waters' distinction in Case-1 and Case-2 and further research on WQ monitoring of a distinct category of optically complex Case-2 waters, oligotrophic lakes; obstacles and weaknesses are discussed concerning the achievement of a higher accuracy. Furthermore, this chapter presents the aim and objectives of the current thesis.

Chapter 2 is entitled "Towards the modelling of Greek lakes WQ using RS technology" and presents the methodological framework established herein for the development of WQ models (Chl-*a*, Secchi depth and Total Phosphorus), applied in 50 lakes constituting the National Lake Network Monitoring of Greece (WFD). This chapter also examines and discusses the efficiency of hereby Landsat-developed WQ models: (1) when applied to Sentinel 2 images, and (2) when employing multi-sensor image data. The significance of this chapter lies initially in the fact that lake WQ elements have been determined with high accuracy (RMSE Chl-*a* values ranging from 1.53 μg/l to 4.6 μg/l; RMSE Secchi values from 0.89 m to 1.7 m; RMSE TP values

from 0.008 mg/l to 0.03 mg/l) throughout the studied lakes. Given the complexity that characterizes the mapping of WQ elements in Case 2 waters (coastal, lakes, rivers) in combination with the wide study area while covering a broad range of limnological conditions, hereby chapter contributes essentially to sustainable water resources management at a country level. Towards strengthening the facilitation of WQ and/or trophic status monitoring across Greek lakes, employment of Sentinel 2 and multi-sensor image data (Landsat 7 ETM; Landsat 8; Sentinel 2 MSI) permits the integration among existing and historical missions while contributing to long-term time series data collection. In this way, the delivery of detailed spatial variability of WQ and trophic status over Greek lakes in fine spatial resolution (10-30 m) further grants the monitoring of lake eutrophication and its spatio-temporal changes; information that is fundamentally valuable in terms of lake management in the framework of national environmental policy.

Chapter 3 with title "Atmospheric correction analysis of lake WQ models by employing surface reflectance (SR) embedded in GEE platform" emphasizes on the harmonization among SR values subjected to different atmospheric correction (AC) methods. The harmonization is based on the further development of corrected sensor-specific (Landsat/ Sentinel 2) WQ models, accommodating inherent spectral and pre-processing differences. Eventually, efficient performance of WQ models employing GEE-derived SR values (datasets of 2 validation years yielded mean RMSE values of Chl-*a*: 20 μg/l-Landsat; 13.4 μg/l Sentinel 2, Secchi depths: 2.1 m –Landsat; 2.8 m- Sentinel 2 and TP: 0.15 mg/l-Landsat; 0.11 mg/l- Sentinel 2) confirmed their spatio-temporal stability despite the AC method applied and satellite sensor used. The significance of this chapter lies on the fact that even though retrieval of WQ parameters requires precise AC, hereby developed models managed to perform well exploiting the massive GEE warehouse of data while exempting from the pre-processing procedure. In this way, GEE facilitates the long-term, near real-time, national-scale lake WQ and trophic status monitoring by mapping long-term WQ trends in less time and fine spatial resolution.

Chapter 4 is entitled "Operational development of techniques for characterising WQ of oligotrophic Case-2 waters" and emphasizes on a distinct category of optically complex Case-2 waters; oligotrophic waterbodies. This chapter investigates hereby WQ models' performance when applied in oligotrophic Trichonis and Amvrakia lakes, and underlies the necessity of the development of special oligotrophic algorithms for a more efficient and comprehensive lake management. Afterwards, an attempt to independently model WQ of Trichonis lake is described, by assessing certain WQ elements through satellite RS and evaluating special models' accuracy. This chapter confirms, in absolute agreement with the relevant literature, the

underachievement of universal WQ models at the lowest chlorophyll concentrations (oligotrophic waterbodies) and at cases where the optical contribution is non-algal. Furthermore, several indications concerning the appropriate satellite bands, the existence of more and narrower wavelengths in specific ranges of electromagnetic spectrum and the refinement of AC processors are provided dedicated to this sub-category of Case-2 waters. Aforementioned indications, accompanied by the preceding and thorough research on phytoplankton community composition, Chl-*a* distributions, particles and CDOM, constitute the background knowledge required and the trigger for a more accurate WQ monitoring of oligotrophic waterbodies along Greece. Based on the above, this research facilitates the refinement of lake WQ monitoring in Greece by laying the foundation stone of further discrimination of Case-2 Greek waterbodies into discrete optical water types (OWT).

Chapter 5 summarizes the most fundamental conclusions, innovation and limitations emerged in the context of this thesis. Additionally, the significance of the current thesis is presented which substantially lies in the generation of national lake WQ models which:

a) have been developed and applied in 50 different Greek lake systems of varied chemistry, limnological conditions and trophic level while were sampled during different seasons
b) accommodate the spectral composition differences among Landsat (7 ETM+, 8 OLI) and Sentinel 2 sensors
c) also accommodate the differences emerging from the different atmospheric correction methods applied in manually-elaborated and in GEE-embedded reflectance values
d) were proven reliable for the systematic assessment of Chl-*a* and TP concentrations and Secchi Disk depths with high accuracy across Greek lakes while providing spatial WQ and trophic status variability in fine resolution (10-30 m)

Based on those characteristics, hereby-delivered WQ models substantially facilitate the monitoring of lake WQ in Greece through satellite RS as they have the potential to constitute a part of a wider national lake management plan and early warning system through the timely identification of pollution events and by extension the promptly performance of sustainably efficient solutions.

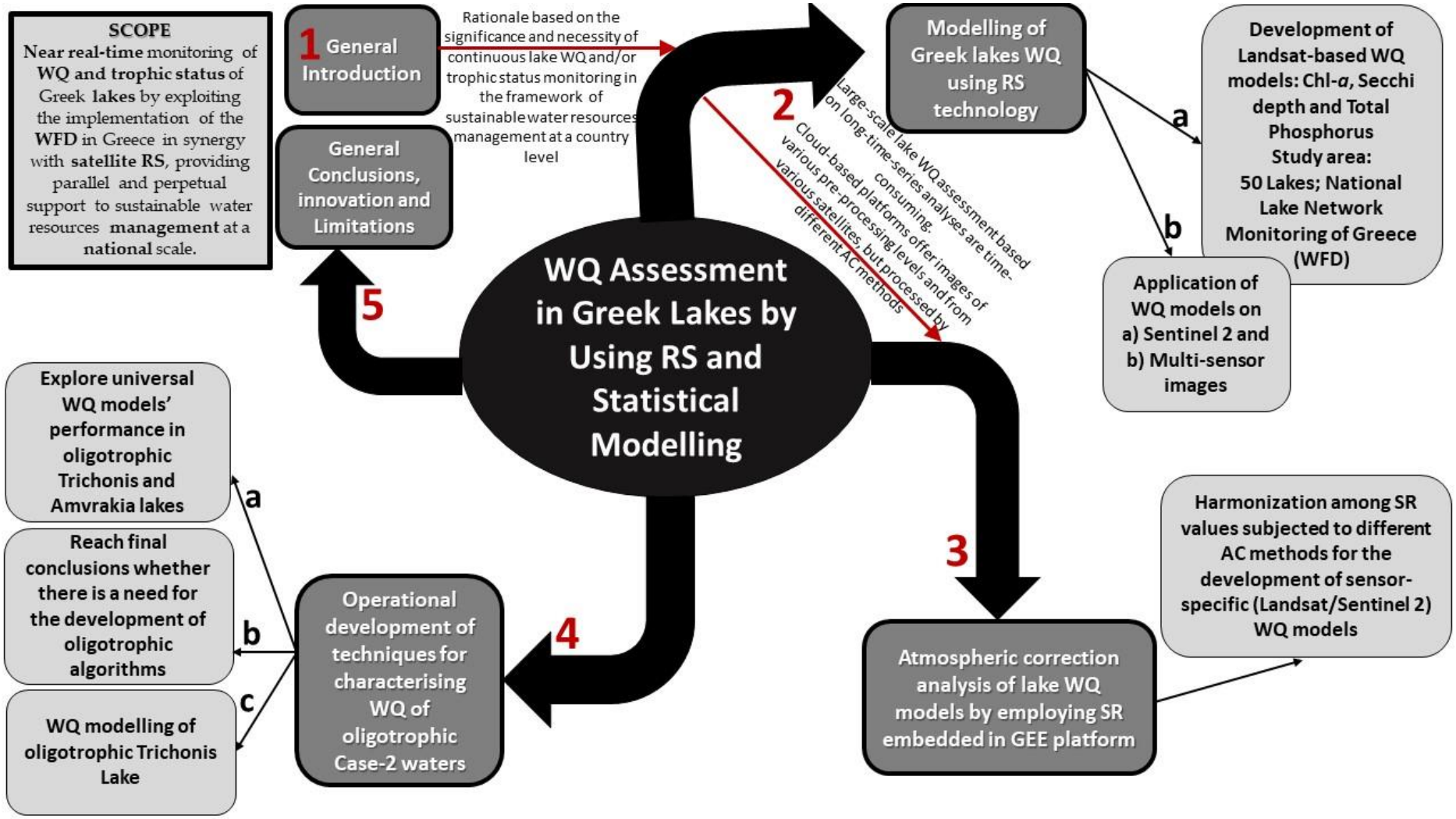


Figure 1. Flowchart presenting the main chapters of the current thesis.

V. ΕΚΤΕΝΗΣ ΠΕΡΙΛΗΨΗ

Δεδομένης της μεγάλης σημασίας που έχουν οι λίμνες για το περιβάλλον και τη ζωή του ανθρώπου, η συνεχής παρακολούθηση της ποιότητας των υδάτων τους μέσω της εφαρμογής της Οδηγίας Πλαίσιο για τα Νερά (ΟΠΥ-WFD) θεωρείται απαραίτητη για την βιώσιμη διαχείρισή τους. Η επιτόπια (*in-situ*) παρακολούθηση της ποιότητας νερού -μέσω δειγματοληψιών- των λιμνών σε συνδυασμό με τη δορυφορική τηλεπισκόπηση (Remote Sensing) αντιπροσωπεύει παγκοσμίως την τελευταία επιστημονική τάση σε πολλά προγράμματα παρακολούθησης ποιότητας υδάτων. Ως εκ τούτου, το ευρύ μεθοδολογικό πλαίσιο που αναπτύχθηκε στην παρούσα διδακτορική διατριβή έχει ως βασικό στόχο τη δημιουργία μοντέλων εκτίμησης ποιότητας υδάτων λιμνών υποστηρίζοντας παράλληλα τη βιώσιμη διαχείριση των υδάτινων πόρων σε εθνική κλίμακα. Η πρακτική εφαρμογή αυτής της μεθοδολογίας αναπτύχθηκε και αξιολογήθηκε συνολικά σε 50 λίμνες (φυσικές και τεχνητές) την περίοδο 2013–2018, αποτελώντας το Εθνικό Δίκτυο Παρακολούθησης Λιμνών της Ελλάδας στο πλαίσιο της ΟΠΥ. Όσον αφορά στα δεδομένα δορυφορικής τηλεπισκόπησης, εικόνες από τους αισθητήρες Landsat 7 Enhanced Thematic Mapper Plus (ETM+) και Landsat 8 Operational Land Imager (OLI) συνδυάστηκαν με *in-situ* μετρήσεις ποιοτικών παραμέτρων με κύριο στόχο την επίτευξη ανάπτυξης μοντέλων εκτίμησης της ποιότητας και της τροφικής κατάστασης υδάτων. Οι ποιοτικές παράμετροι που μελετήθηκαν είναι η συγκέντρωση της Χλωροφύλλης-α (Chl-*a*), το βάθος Secchi Disk, η συγκέντρωση Ολικού Φωσφόρου (TP- Total Phosphorus) και κατ'επέκταση ο Δείκτης Τροφικής Κατάστασης (TSI- Trophic State Index).

Αρχικά, η εκτίμηση της συγκέντρωσης της Chl-*a* σε ύδατα της Περίπτωσης-2 (Case-2 waters; παράκτιες περιοχές, ποτάμια και λίμνες) είναι καίριας σημασίας, καθώς αυτή η παράμετρος αποτελεί σημαντική ένδειξη της ακεραιότητας του οικοσυστήματος. Από στατιστικής άποψης, πραγματοποιήθηκε ανάλυση κύριων συνιστωσών (PCA-Principal Component Analysis) για να διερευνηθούν οι συσχετίσεις των συγκεντρώσεων της Chl-*a* των ελληνικών λιμνών με άλλα βασικά τους κριτήρια, όπως το εάν είναι φυσικές ή τεχνητές, την τυπολογία (ΟΠΥ), τον κλιματικό τύπο σύμφωνα με το σύστημα ταξινόμησης Köppen-Geiger, την εποχή δειγματοληψιών νερού και την ημερολογιακή διαφορά μεταξύ των *in-situ* δειγματοληψιών και του περάσματος του δορυφόρου. **Η ανάλυση κύριων συνιστωσών υπέδειξε το χαρακτηριστικό φυσική/τεχνητή και την τυπολογία (ΟΠΥ) τους ως τις παραμέτρους που κυρίως συμβάλλουν στη διακύμανση της συγκέντρωσης της Chl-*a*.** Στη συνέχεια, διενεργήθηκαν αναλύσεις πολλαπλής παλινδρόμησης, με τη μέθοδο stepwise, μεταξύ διαφορετικών ομάδων δεδομένων, οι οποίες

διαμορφώθηκαν με βάση τα κριτήρια της PCA. Βασικός στόχος ήταν η ανάπτυξη διαφορετικών αλγορίθμων εκτίμησης της συγκέντρωσης Chl-*a*, προσανατολισμένοι σε διαφορετικούς τύπους λιμνών. Πρόσθετες αναλύσεις παλινδρόμησης εφαρμόστηκαν και σε σύνολα δεδομένων για τα οποία δε λήφθηκαν υπόψη συγκεκριμένα κριτήρια και αφορούν τις περιπτώσεις όταν δεν υπάρχουν διαθέσιμες πληροφορίες σχετικά με τα χαρακτηριστικά των μελετούμενων λιμνών.

Στη συνέχεια, διεξήχθη ανάλυση συσχέτισης (correlation analysis) μεταξύ των *in-situ* μετρήσεων των μελετούμενων παραμέτρων ποιότητας (Chl-*a*, βάθη Secchi disk και TP) με σκοπό τη διερεύνηση και ανίχνευση των συσχετίσεών τους. Με βάση τα αποτελέσματα της ανάλυσης συσχέτισης, διενεργήθηκαν περαιτέρω αναλύσεις παλινδρόμησης, χρησιμοποιώντας τις *in-situ* τιμές βάθους Secchi disk και TP για τη δημιουργία βέλτιστων μοντέλων όλων των κατηγοριών (όλες, φυσικές και τεχνητές λίμνες). Έπειτα, διεξήχθη η ταξινόμηση της τροφικής κατάστασης των μελετούμενων λιμνών, υπολογίζοντας τον Δείκτη Τροφικής Κατάστασης (TSI) του Carlson για κάθε λίμνη, αρχικά για όλες και στη συνέχεια ξεχωριστά για τις φυσικές και τις τεχνητές. Οι υπολογισμοί των διαφορετικών δεικτών TSI (γενικός, φυσικός, τεχνητός) υπολογίστηκαν με βάση τα αντίστοιχα μοντέλα ποιότητας (Χλωροφύλλη-*α*, βάθος Secchi disk και Ολικός φώσφορος).

Συνολικά, **τα αποτελέσματα που αφορούν στην εκτίμηση των συγκεντρώσεων της Χλωροφύλλης-α και πιο συγκεκριμένα του λογαριθμικού μετασχηματισμού τους (logChl-*a*), απέδειξαν την καταλληλότητα των δεδομένων των δορυφόρων Landsat (7 ETM+ και 8 OLI).** Η προτεινόμενη μεθοδολογία είχε ως αποτέλεσμα **την ανάπτυξη ξεχωριστών μοντέλων για τις φυσικές (R = 0.78) και τεχνητές (R= 0.76) λίμνες, ενώ το γενικό μοντέλο που αναπτύχθηκε χωρίς κριτήρια αποδείχθηκε ασθενέστερο (R = 0.65) σε σύγκριση με τα υπόλοιπα.** Τα αποτελέσματα του **πίνακα συσχέτισης (correlation matrix) μεταξύ των *in-situ* δεδομένων Χλωροφύλλης-α, βαθών Secchi disk και TP υπέδειξαν υψηλή και θετική σχέση μεταξύ TP και Χλωροφύλλης-α (0.85), ενώ υψηλές αρνητικές σχέσεις εντοπίστηκαν μεταξύ βάθους Secchi disk με TP (-0.84) και Χλωροφύλλη-α (-0.83). Οι αναλύσεις παλινδρόμησης μεταξύ των δεδομένων Landsat και βάθους Secchi οδήγησαν σε 3 μοντέλα: για όλες τις λίμνες ($Secchi_{general}$; R = 0.78; RMSE = 2 m), για τις φυσικές ($Secchi_{natural}$; R = 0.95; RMSE = 1.87 m) και για τις τεχνητές ($Secchi_{artificial}$, R = 0.62, RMSE = 1.36 m), με αξιόπιστη ακρίβεια. Τα ευρήματα έδειξαν ότι οι αναλύσεις παλινδρόμησης που σχετίζονται με τον Ολικό Φώσφορο απέτυχαν να αναπτύξουν ένα στατιστικά αποδεκτό μοντέλο για τις τεχνητές λίμνες. Ωστόσο, ανέδειξαν ένα ισχυρό για όλες τις λίμνες $TP_{general}$ (R = 0.71, RMSE = 0.008 mg/L) και για τις φυσικές $TP_{natural}$ (R = 0.93, RMSE = 0.018 mg/L). Στη**

συνέχεια, όσον αφορά στα αποτελέσματα του δείκτη τροφικής κατάστασης TSI, η μεγαλύτερη απόκλιση των προβλεπόμενων τιμών σε σχέση με τις αντίστοιχες *in-situ* ανιχνεύθηκε σε τεχνητές και πιο ρηχές λίμνες (μέσο βάθος < 5 m), επιβεβαιώνοντας την ύπαρξη αξιοσημείωτων διαφοροποιήσεων μεταξύ φυσικών και τεχνητών υδάτινων σωμάτων. Συνοψίζοντας, με βάση τα αποτελέσματα της όλης επιστημονικής προσπάθειας, αποδείχθηκε ότι η συγκεκριμένη μεθοδολογία είναι ικανή να υποστηρίξει ικανοποιητικά τη διαρκή παρακολούθηση της ποιότητας υδάτων και την αξιολόγηση της τροφικής κατάστασης των ελληνικών λιμνών. Κατ' επέκταση, διευκολύνεται η βιώσιμη διαχείρισή τους, ιδιαίτερα σε περιπτώσεις που τα *in-situ* δεδομένα είναι σχετικά περιορισμένα.

Άλλα βασικά ερωτήματα τα οποία απάντησε η παρούσα διδακτορική διατριβή είναι αρχικά εάν οι προαναφερθέντες εμπειρικοί αλγόριθμοι παρακολούθησης της ποιότητας λιμνών που αναπτύχθηκαν με εικόνες Landsat μπορούν να εφαρμοστούν επιτυχώς σε εικόνες Sentinel 2 και στη συνέχεια εάν η συνδυασμένη χρήση εικόνων πολλαπλών αισθητήρων δύναται να βελτιώσει την ακρίβεια ποσοτικοποίησής τους. Επιπλέον, ανεξάρτητα από την ύπαρξη ή όχι κάποιας βελτίωσης, ένας άλλος στόχος αυτής της συνδυασμένης προσέγγισης είναι να καθοριστεί εάν οι εικόνες πολλαπλών αισθητήρων θα μπορούσαν να χρησιμοποιηθούν με τουλάχιστον εξίσου αξιόπιστα αποτελέσματα όπως αυτά που προκύπτουν από τη χρήση εικόνων μόνο ενός αισθητήρα. Μεταξύ των πολλαπλών αισθητήρων ανιχνεύονται πολυάριθμες τροχιακές, χωρικές και φασματικές διαφοροποιήσεις, ωστόσο η δυνατότητα μεταφοράς των αλγόριθμων εκτίμησης ποιότητας υδάτων μεταξύ τους, δεν έχει εξεταστεί επαρκώς στη διεθνή βιβλιογραφία. Προς αυτή την κατεύθυνση, χρησιμοποιήθηκαν εικόνες Sentinel 2 MSI της χρονιάς 2018 με ημερομηνίες ταυτόχρονες με αυτές των μετρήσεων πεδίου με σκοπό την αξιολόγηση της απόδοσης των εμπειρικών μοντέλων και τη σύγκριση των αποτελεσμάτων επικύρωσής τους με τα αντίστοιχα των δορυφόρων Landsat.

Επιπλέον, μια άλλη προσπάθεια βελτίωσης της ακρίβειας ποσοτικοποίησης των μοντέλων πραγματοποιήθηκε μέσω της συνδυασμένης χρήσης εικόνων Landsat (7 ETM+, 8 OLI) και Sentinel 2 MSI, ενώ η επιλογή κάθε εικόνας για κάθε περίπτωση βασίστηκε στην πλησιέστερη ημερομηνία λήψης της σε σχέση με αυτή της αντίστοιχης δειγματοληπτικής. Αναφορικά με τα αποτελέσματα, **τα γενικά μοντέλα όλων των ποιοτικών παραμέτρων (Chl-*a*, βάθος Secchi Disk και TP), με εφαρμογή σε όλες τις λίμνες ήταν πιο αποτελεσματικά και ακριβή όταν εφαρμόστηκαν σε εικόνες Landsat, ενώ δεν παρατηρήθηκε βελτίωση ύστερα από τη χρήση εικόνων πολλαπλών αισθητήρων. Τα**

μοντέλα που αναπτύχθηκαν και εφαρμόστηκαν στις φυσικές λίμνες), παρουσίασαν διαφορετική συμπεριφορά. Τα μοντέλα Chl-*a* και TP παρουσίασαν μικρότερες τιμές σφάλματος όταν εφαρμόστηκαν σε εικόνες Sentinel 2 (RMSE Chl-*a*=16.4 μg/l vs 21.5 μg/l; RMSE TP=0.03 mg/l vs 0.031 mg/l) και μόνο το μοντέλο Secchi$_{natural}$ είχε καλύτερη απόδοση με δεδομένα Landsat (2.8 m vs 2.9 m). Η συνδυαστική χρήση εικόνων Landsat και Sentinel 2 δεν προσέφερε καμία βελτίωση στα αντίστοιχα μοντέλα Chl-*a* και βάθους Secchi Disk, ενώ η χρήση εικόνων πολλαπλών αισθητήρων οδήγησε στην εκτίμηση συγκεντρώσεων TP με εξίσου αξιόπιστα αποτελέσματα όπως αυτά των δεδομένων Sentinel 2. Όσον αφορά στους αλγόριθμους των τεχνητών λιμνών, το μοντέλο της Chl-*a* απέδωσε καλύτερα ύστερα από τη χρήση Landsat εικόνων (RMSE= 3.7 μg/l vs 7.7 μg/l; Sentinel 2), ενώ το μοντέλο εκτίμησης βάθους Secchi Disk πέτυχε ελαφρώς καλύτερη απόδοση χρησιμοποιώντας εικόνες Sentinel 2 (RMSE= 1.5 m vs 1.6 m; Landsat). Τη χειρότερη απόδοση ανάμεσα στις υπό μελέτη ποιοτικές παραμέτρους παρουσίασαν τα μοντέλα της Chl-*a* (οι διάμεσες MAPE τιμές κυμάνθηκαν από 42 % έως 58% ενώ του βάθους Secchi από 24% έως 44% και του Ολικού Φωσφόρου από 22% έως 38%), γεγονός που επιβεβαιώνει για ακόμα μία φορά την πολυπλοκότητα που κρύβει η χαρτογράφηση της Chl-*a* στα ύδατα της Περίπτωσης 2 (παράκτια, εσωτερικά ύδατα). Συνοψίζοντας, προτείνεται τα εμπειρικά μοντέλα παρακολούθησης της ποιότητας των ελληνικών λιμνών να εφαρμόζονται κυρίως σε εικόνες Landsat. Ωστόσο, η χρήση των δεδομένων Sentinel 2 δυνητικά παράγει εξίσου αξιόπιστα αποτελέσματα με ορισμένες (όχι σημαντικές) αποκλίσεις από τα αντίστοιχα αποτελέμστα των Landsat αλλά και από τις επιτόπιες μετρήσεις της ποιότητας υδάτων των λιμνών.

Σήμερα, οι πλατφόρμες υπολογιστικής νέφους ανοιχτού κώδικα έχουν αναδειχθεί σε πολύτιμα εργαλεία γεωχωρικής ανάλυσης δεδομένων εικόνας διαφόρων δορυφόρων. Ιδιαίτερα η χρήση της πλατφόρμας Google Earth Engine (GEE), είναι η πιο διαδεδομένη στο επιστημονικό πεδίο της δορυφορικής τηλεπισκόπησης. Οι πρόσφατες εκτοξεύσεις διαφόρων δορυφόρων σε συνδυασμό με την πλατφόρμα GEE, διευκολύνουν σε μεγάλο βαθμό την παρακολούθηση των λιμνών σε εθνική κλίμακα. Με βασικούς στόχους αφενός την αξιοποίηση των *in-situ* δεδομένων ποιότητας της ΟΠΥ, και αφετέρου το συνδυασμό τους με τη δυναμική της πλατφόρμας GEE, το επόμενο βήμα ήταν η εξακρίβωση της απόδοσης των ποιοτικών μοντέλων όταν χρησιμοποιούν τιμές ανάκλασης οι οποίες έχουν υποβληθεί σε διαφορετικές μεθόδους ατμοσφαιρικής διόρθωσης από αυτές όταν αναπτύχθηκαν.

Πιο συγκεκριμένα, τιμές ανάκλασης από εικόνες Landsat και Sentinel 2 του έτους 2018, αποκτήθηκαν μέσω της πλατφόρμας GEE στους σταθμούς δειγματοληψίας λιμνών της ΟΠΥ ενώ ήταν ατμοσφαιρικά διορθωμένες με

τις μεθόδους LaSRC (Landsat 8 OLI), LEDAPS (Landsat 7 ETM+) και Sen2Cor (Sentinel 2). Οι εν λόγω τιμές ανάκλασης αντιστοιχίστηκαν με τα *in-situ* δεδομένα ποιότητας του 2018 με διαφορά ημερομηνιών ±7 ημερών (από την ημερομηνία της δειγματοληψίας), αποδίδοντας 192 και 210 ζεύγη δεδομένων με εικόνες Landsat και Sentinel 2, αντίστοιχα. Τα ίδια ζεύγη δημιουργήθηκαν με τις τιμές ανάκλασης προερχόμενες από τις αντίστοιχες εικόνες, οι οποίες αποκτήθηκαν και προ-επεξεργάστηκαν χειροκίνητα, με τη μέθοδο της ατμοσφαιρικής διόρθωσης DOS1. Τα εμπειρικά μοντέλα Chl-*a*, βάθους Secchi Disk και TP (για όλες και ξεχωριστά για φυσικές και τεχνητές λίμνες), εφαρμόστηκαν δύο φορές χρησιμοποιώντας και τα δύο είδη ανάκλασης, την DOS1-διορθωμένη ανάκλαση και εκείνη που αποκτήθηκε από εικόνες ενσωματωμένες στη GEE πλατφόρμα. Επιπρόσθετα, η διπλή εφαρμογή των εμπειρικών μοντέλων εκτίμησης της ποιότητας υδάτων πραγματοποιήθηκε ξεχωριστα για τους δορυφόρους Landsat (7 ETM+/OLI) και Sentinel 2. Η διπλή εφαρμογή των ποιοτικών μοντέλων είχε ως αποτέλεσμα τη διπλή εκτίμηση των τιμών της εκάστοτε ποιοτικής παραμέτρου στον εκάστοτε σταθμό δειγματοληψίας και στη συνέχεια αυτές οι διπλές τιμές (προερχόμενες από DOS1 και λοιπές ατμοσφαιρικές μεθόδους) εισήχθησαν σε ανάλυση γραμμικής παλινδρόμησης. **Οι γραμμικές εξισώσεις που προέκυψαν μεμονωμένα για κάθε αισθητήρα (Landsat, Sentinel 2) υπέδειξαν τιμές υψηλής συσχέτισης (R^2 από 0.68 έως 0.98) και αποτέλεσαν τα διορθωμένα πλέον μοντέλα εκτίμησης ποιότητας για χρήση και εφαρμογή τους με δεδομένα που έχουν υποβληθεί σε ίδιες μεθόδους ατμοσφαιρικής διόρθωσης με αυτές που εντοπίζονται στη GEE πλατφόρμα.**

Τα αρχικά αλλά και τα διορθωμένα μοντέλα εκτίμησης της ποιότητας επικυρώθηκαν περαιτέρω με βάση τα διαθέσιμα *in-situ* δεδομένα των ετών 2019 και 2020, περιλαμβάνοντας 239 (Landsat-GEE) και 242 (Sentinel 2-GEE) και 220 (Landsat-GEE) και 286 (Sentinel 2-GEE) αντιστοιχισμένα ζεύγη, αντίστοιχα. Η επικύρωση των αρχικών και των διορθωμένων μοντέλων εκτίμησης της ποιότητας υδάτων βασίστηκε στη χρήση στατιστικών δεικτών και μετρικών σφαλμάτων όπως η ρίζα μέσου τετραγωνικού σφάλματος (RMSE), το κανονικοποιημένο μέσο τετραγωνικό σφάλμα (NRMSE) και η μέση εκατοστιαία απόλυτη απόκλιση (MAPE).

Η διόρθωση των εμπειρικών μοντέλων ανάλογα με τον αισθητήρα αποδείχθηκε απαραίτητη για ορισμένα από αυτά, ενώ οι τιμές RMSE κυμάνθηκαν για τη Chl-*a* από 11.68 μg/l (Landsat) έως 14.88 μg/l (Sentinel 2), για το βάθος Secchi Disk από 2.02 m (Landsat) έως 2.57 m (Sentinel 2) και TP από 0.14 mg/l (Landsat) έως 0.09 mg/l (Sentinel 2), τιμές που επιβεβαίωσαν τη σταθερότητα και τη δυνατότητα

μεταφοράς των εμπειρικών μοντέλων σε εικόνες διαφορετικά ατμοσφαρικά διορθωμένες, στην πλατφόρμα GEE. Με αυτό τον τρόπο, εξοικονομείται πολύτιμος χρόνος που αντιστοιχεί στην προεπεξεργασία των δορυφορικών εικόνων από το χρήστη πριν τη χρήση τους, διευκολύνοντας έτσι την παρακολούθηση της ποιότητας λιμνών σε εθνικό επίπεδο και παρέχοντας παράλληλα στους διαχειριστές υδάτινων πόρων ένα πρόσθετο εργαλείο για τη λήψη μέτρων προστασίας των υδάτων.

Εφόσον διερευνήθηκε η επιτυχής ή μη εφαρμογή των μοντέλων εκτίμησης ποιότητας υδάτων αρχικά σε Sentinel 2 εικόνες και στη συνέχεια σε εικόνες που έχουν υποστεί ατμοσφαιρική διόρθωση μέσω διαφορετικών μεθόδων στη GEE πλατφόρμα, απαντήθηκε ακόμα ένα περίπλοκο ερώτημα. Το πρώτο σκέλος του ερωτήματος που ερευνήθηκε είναι το εάν τα μοντέλα εκτίμησης ποιότητας υδάτων (Chl-*a*, βάθους Secchi Disk και Ολικού Φωσφόρου) δύναται να είναι αποτελεσματικά για την παρακολούθηση της ποιότητας και των ολιγοτροφικών υδάτων και το δεύτερο σκέλος αφορά στην εξακρίβωση της ύπαρξης ή μη ανάγκης να αναπτυχθούν ειδικοί αλγόριθμοι αποκλειστικά προσανατολισμένοι σε αυτή την κατηγορία υδάτων. Η ταξινόμηση των υδάτων στην Περίπτωση 1 (ωκεάνια) και στην Περίπτωση 2 (παράκτιες περιοχές, ποτάμια και λίμνες), είναι ιδιαίτερα σημαντική. Τα νερά της Περίπτωσης 1 προσδιορίζονται με βάση το φυτοπλαγκτόν και λοιπές ουσίες, ενώ τα νερά της Περίπτωσης 2 είναι πιο πολύπλοκα όσον αφορά στη σύσταση και τις οπτικές τους ιδιότητες. Ένας από τους βασικούς παράγοντες που εμποδίζει την εκτίμηση της ποιότητας υδάτων στα νερά της Περίπτωσης 2 με ακρίβεια είναι το γεγονός ότι τα αιωρούμενα υλικά, η συγκέντρωση της έγχρωμης διαλελυμένης οργανικής ύλης (CDOM- Colored Dissolved Organic Matter) και η ανάκλαση του πυθμένα μεταβάλλονται ανεξάρτητα το ένα από το άλλο. Οι ολιγοτροφικές λίμνες αποτελούν μια ξεχωριστή κατηγορία των περίπλοκων οπτικά υδάτων της Περίπτωσης 2 και δε ταξινομούνται στην Περίπτωση 1 δεδομένου ότι συνήθως λαμβάνουν σημαντικές εισροές ιζημάτων και η διαύγεια των υδάτων τους ελέγχεται συνήθως από τις συγκεντρώσεις του διαλυμένου οργανικού άνθρακα (DOC-Dissolved Organic Carbon) και όχι από το φυτοπλαγκτόν. Με βάση τα προαναφερθέντα, τα μοντέλα εκτίμησης ποιότητας υδάτων λιμνών (Chl-*a*, βάθους Secchi Disk και TP), τα οποία αναπτύχθηκαν και «εκπαιδεύτηκαν» με ένα αρκετά μεγάλο εύρος συγκεντρώσεων, εφαρμόστηκαν σε εικόνες Landsat 8 OLI που απεικονίζουν δύο (2) ελληνικά ολιγοτροφικά υδάτινα σώματα, τις λίμνες Τριχωνίδα και Αμβρακία. Τα διαθέσιμα *in-situ* δεδομένα ποιότητας συλλέχθηκαν τα έτη 2013 και 2014, ενώ οι δορυφορικές εικόνες που χρησιμοποιήθηκαν έχουν ταυτόχρονες ημερομηνίες. **Η εφαρμογή των μοντέλων στις ολιγοτροφικές λίμνες Τριχωνίδα και Αμβρακία ήταν αναποτελεσματική˙ το γενικό μοντέλο υπολογισμού της**

Χλωροφύλλης-α απέδωσε τιμές RMSE=1.9 μg/l, NRMSE=1.6 και median MAPE=256.8, το μοντέλο υπολογισμού Χλωροφύλλης-α στις φυσικές λίμνες απέδωσε: RMSE=1.8 μg/l, NRMSE=1.5 και median MAPE=176.6 ενώ τα αποτελέσματα των λοιπών μοντέλων Ολικού Φωσφόρου και βάθους Secchi Disk δεν ήταν στατιστικά σημαντικά σύμφωνα με το δείκτη συνάφειας Spearman r και με επιθυμητό επίπεδο σημαντικότητας 0.01. Τα προαναφερθέντα ευρήματα είναι παρόμοια με εκείνα άλλων μελετών που διερεύνησαν ολιγοτροφικά υδάτινα σώματα. Επιπρόσθετα, οι ιδιαίτερα χαμηλές και ομοιογενείς συγκεντρώσεις Chl-*a*, μέση τιμή 0.6 μg/l και 0.7 μg/l, που μετρήθηκαν στις λίμνες Τριχωνίδα και Αμβρακία, αντίστοιχα, υπέδειξαν ότι η μεγαλύτερη συμβολή οπτικά προέρχεται από σωματίδια μη σχετιζόμενων με το φυτοπλαγκτόν, γεγονός που υποδηλώνει την ανάγκη ανάπτυξης ειδικών αλγορίθμων.

Λαμβάνοντας υπόψιν τα προαναφερθέντα αποτελέσματα, έγινε μια προσπάθεια ανάπτυξης μοντέλων εκτίμησης της ποιότητας υδάτων στη λίμνη Τριχωνίδα. Στο πλαίσιο αυτής της προσπάθειας, διερευνήθηκε η καταλληλότητα εικόνων του δορυφόρου Landsat 8 OLI για την ακριβή εκτίμηση των συγκεντρώσεων Chl-*a*, των θρεπτικών αλάτων και της απορρόφησης CDOM σε συγκεκριμένα μήκη κύματος. Η λίμνη Τριχωνίδα αποτελεί τη μεγαλύτερη φυσική λίμνη στην Ελλάδα, είναι ολιγοτροφική και χαρακτηρίζεται από ανύπαρκτη ποσοτική, χρονική και χωρική μεταβλητότητα ως προς τις συγκεντρώσεις των μελετούμενων ποιοτικών παραμέτρων. Δείγματα νερού συλλέχθηκαν από 22 σταθμούς στα τέλη Αυγούστου 2014 και οι μετρούμενες *in-situ* συγκεντρώσεις τους συνδυάστηκαν μέσω ανάλυσης παλινδρόμησης με μία δορυφορική εικόνα Landsat 8 OLI της ίδιας ημερομηνίας με σκοπό την ανάπτυξη μοντέλων εκτίμησης ποιότητας υδάτων. Στη συνέχεια, τα βέλτιστα μοντέλα εφαρμόστηκαν στη δορυφορική εικόνα του 2013 και η επικύρωση των αποτελεσμάτων πραγματοποιήθηκε χρησιμοποιώντας τα αντίστοιχα *in-situ* δεδομένα του 2013. **Τα αποτελέσματα από τη διαδικασία επικύρωσης έδειξαν μια σχετικά διαφορετική στατιστική σχέση μεταξύ των *in-situ* δεδομένων και των ανακλάσεων (R logchl-*a*: 0.4, R NH_4^+: 0.7, R Chl-*a*: 0.5, R CDOM στα 420 nm: 0.3).** Οι επιτόπιες μετρήσεις των συγκεντρώσεων νιτρικών, νιτρωδών, φωσφορικών αλάτων και ολικού αζώτου του 2014 μετρήθηκαν ως χαμηλότερες από το όριο ανίχνευσης του χρησιμοποιούμενου οργάνου, επομένως δεν πραγματοποιήθηκε στατιστική επεξεργασία. **Επιπλέον, η ανάλυση παλινδρόμησης μεταξύ της ανάκλασης και των συγκεντρώσεων TP οδήγησε σε χαμηλές και στατιστικά μη σημαντικές συσχετίσεις (οι τιμές R^2 κυμάνθηκαν από 0.06 έως 0.07).** Τα ευρήματά μας ήταν επίσης παρόμοια με αυτά άλλων μελετών στη διεθνή βιβλιογραφία, υποδεικνύοντας ότι η παρακολούθηση ποιότητας στις ολιγοτροφικές λίμνες είναι λιγότερο ακριβής σε σχέση με

τις ευτροφικές και μεσοτροφικές, λόγω της απουσίας αιωρούμενων σωματιδίων τα οποία είναι ανιχνεύσιμα από τους δορυφορικούς αισθητήρες. Συμπερασματικά προκύπτει ότι παρά την ύπαρξη πολλών *in-situ* μετρήσεων σε πυκνό δίκτυο σταθμών, όπως συνέβη στην παρούσα διατριβή, μια ποσοτικά ακριβής εκτίμηση ποιοτικών παραμέτρων στα παράκτια/εσωτερικά ύδατα παραμένει μια μεγάλη πρόκληση.

VI. EXTENDED SUMMARY

Given the great importance of lakes in Earth's environment and human life, continuous water quality (WQ) monitoring within the frame of the Water Framework Directive (WFD) is the most crucial aspect for lake management. *In-situ* monitoring of lake WQ in synergy with satellite remote sensing (RS) represents the latest scientific trend in many WQ monitoring programs worldwide. Therefore, the wide methodological framework developed herein has as an ultimate goal the generation of lake WQ quantitative models, supporting sustainable water resources management at a national scale. The practical use of this approach was developed and evaluated in a total of 50 lake water bodies (natural and artificial) from 2013–2018, constituting the National Lake Network Monitoring of Greece in the context of the WFD. Concerning the utilized Earth Observation (EO) data, images from Landsat 7 Enhanced Thematic Mapper Plus (ETM+) and Landsat 8 Operational Land Imager (OLI) sensors have been combined with co-orbital WQ *in-situ* measurements with the main objective of delivering robust WQ assessment models.

In the first instance, assessing Chlorophyll-*a* (Chl-*a*) pigments in complex inland water systems (Case-2 waters) is of key importance as this parameter constitutes a major ecosystem integrity indicator. From the statistical point of view, principal component analysis (PCA) was performed to explore Greek lakes' interrelationships among their Chl-*a* values and certain criteria, e.g. their characteristics (artificial/natural), WFD typology, climatic type (according to the Köppen-Geiger climate classification), season of water samplings and the date difference between sampling and satellite overpass. PCA highlighted the lake characteristics (natural/artificial) and WFD typology as the variables that mostly contribute to the variance of Chl-*a* concentration; thus, numerous stepwise multiple regression analyses (MLRs) among different groups of cases, formed by the PCA criteria, were implemented with basic aim the generation of different remote sensing-derived Chl-*a* algorithms for different types of lakes. MLRs analysis was also implemented employing datasets without considering certain criteria for cases where no information is available about their characteristics.

Moreover, correlation analysis among *in-situ* co-orbital WQ data including Chl-*a*, Secchi depths and Total phosphorus (TP) concentrations, was conducted to explore and detect their inter-relationships. Subsequently, based on correlation analysis's results, further stepwise MLRs employing available *in-situ* TP and Secchi depth datasets were further implemented to establish optimal quantitative models (general, oriented to natural-only and artificial-only lakes). Eventually, trophic status classification was conducted herein, calculating Carlson's Trophic State Index (TSI) of each lake, initially throughout all lakes and then oriented toward natural-only and artificial-only

lakes. Those three types of TSI (general, natural, artificial) were calculated based on hereby specially designed WQ models (Chl-*a*, Secchi depth, TP).
All in all, the results, concerning Chl-*a*, evidenced the suitability of Landsat data to estimate log-transformed Chl-*a*. The proposed scheme resulted in the development of models separately for natural (R = 0.78) and artificial (R = 0.76) lakes, while the model developed without criteria proved weaker (R = 0.65) in comparison to the other ones examined. Correlation matrix results among *in-situ* Chl-*a*, Secchi and TP data showed a high and positive relationship between TP and Chl-*a* (0.85), whereas high negative relationships were found between Secchi depth with TP (-0.84) and Chl-*a* (-0.83). MLRs among Landsat data and Secchi depths resulted in 3 optimal models concerning the assessment of Secchi depth of all lakes ($Secchi_{general}$; R = 0.78; RMSE = 2 m), natural ($Secchi_{natural}$; R = 0.95; RMSE = 1.87 m) and artificial ($Secchi_{artificial}$; R = 0.62; RMSE = 1.36 m), with reliable accuracy. Study findings showed that TP-related MLR analyses failed to deliver a statistically acceptable model for the reservoirs; nevertheless, they delivered a robust $TP_{general}$ model for all lakes (R = 0.71; RMSE = 0.008 mg/L) and a $TP_{natural}$ model for natural lakes (R = 0.93; RMSE = 0.018 mg/L). Subsequently, regarding the TSI results, the higher deviation of satellite-derived TSI values in relation to *in-situ* ones was detected in reservoirs and shallower lakes (mean depth < 5 m), indicating noticeable divergences among natural and artificial waterbodies. Summarizing, this particular part of the whole scientific effort was proven capable of providing important support towards the perpetual WQ monitoring and trophic status assessment of Greek lakes and, by extension, their sustainable management, particularly in cases when ground truth data is limited.

Some other key questions that this thesis answered are initially whether aforementioned Landsat-based empirical WQ algorithms can be efficiently applied to Sentinel 2 images and then whether the combined use of multi-sensor data improves those algorithms' prediction accuracy. Additionally, independently from whether there is some improvement or not, another goal of this combined approach is to decide whether multi-sensor images could be used with at least equally reliable results as those accrued from only-one sensor's utilization. Among sensors numerous orbital, spatial and spectral differences are detected, however transferability of WQ algorithms across them remains poorly examined. Towards this direction, Sentinel 2 images of 2018 with concurrent dates with those of field measurements were utilized to facilitate a WQ models' efficiency evaluation and comparison with the respective Landsat's validation results. Additionally, another effort has been made to improve WQ models' quantification capability through the combined use of Landsat (7 ETM+, 8 OLI) and Sentinel 2 images, while the selection of each image for each case was based on the nearest acquisition date to the sampling one. As far as the

results are concerned, in particular for general models of all WQ elements (Chl-*a*, Secchi depth and TP), all models were more efficient and accurate when were accompanied by Landsat images while no improvement was observed by using multi sensor images. Models developed and applied to natural lakes, though, demonstrated a different behavior. **Chl-*a* and TP models (natural lakes) presented lower values of error metrics when employing Sentinel 2 images (RMSE Chl-*a*=16.4 μg/l vs 21.5 μg/l; RMSE TP=0.03 mg/l vs 0.031 mg/l) and only $Secchi_{natural}$ model performed better with Landsat data (2.8 m vs 2.9 m).** Combined utilization of Landsat and Sentinel 2 images did not provide any improvement to corresponding Chl-*a* and Secchi models whereas the multi sensor images resulted in TP concentrations with equally reliable outcomes as those employing Sentinel 2. **Regarding the algorithms developed and applied in artificial lakes, performance of Chl-*a* model was better by exploiting Landsat data (RMSE= 3.7 μg/l vs 7.7 μg/l of Sentinel 2) while Secchi model achieved slightly better efficiency with Sentinel 2 images (RMSE= 1.5 m vs 1.6 m of Landsat).** The largely worse performance of Chl-*a* models compared to rest of WQ elements (median MAPE values ranged from 42 % to 58%, Secchi depth from 24% to 44% and TP from 22% to 38%), emphasized once again the complexity that mapping of Chl-*a* in Case 2 waters (coastal and/or inland waters) hides. Summing up, it is proven that hereby WQ models are proposed to employ principally Landsat images; however, the employment of Sentinel 2 data potentially produces reliable results with some (not significant) deviations in assessment of lake WQ.

Today, open source Cloud Computing platforms have emerged as a valuable tool for geospatial analysis of image data from various satellites. In particular, the Google Earth Engine (GEE) platform is the most widespread in the scientific field of satellite RS. Newest launches of various satellites in combination with the GEE platform, facilitate in a great extent national-scale lake monitoring. In order to take advantage of *in-situ* lake WQ data derived from the ongoing WFD implementation in Greece and the high potential of GEE platform, next step was to test the transferability and performance of hereby-developed empirical WQ algorithms when employing Landsat (7 +ETM/8 OLI) and Sentinel 2 surface reflectance (SR) values embedded in GEE and subjected to different atmospheric correction (AC) methods from those used as they were developed. More particularly, GEE-Landsat and -Sentinel 2 SR of year 2018 was retrieved from the WFD lake sampling stations and were atmospherically corrected by the methods LaSRC (Landsat 8 OLI), LEDAPS (Landsat 7 ETM+) and Sen2Cor (Sentinel 2). Those SR values (GEE) were matched with WQ *in-situ* data of 2018 within ±7 days (from sampling date) of satellite overpasses, yielding 192 (Landsat) and 210 (Sentinel 2) matched pairs. Same pairs were created with SR derived from manually downloaded and

pre-processed, with AC DOS1 method, respective images. Empirical WQ models of Chl-*a*, Secchi depth and TP (for all and separately for natural and artificial lakes), were applied twice employing both types of SR (DOS1- and rest in GEE- corrected). Furthermore, double application of WQ models was conducted separately for Landsat (7 ETM+/OLI) and Sentinel 2 data. Double application of WQ models resulted in double quantifications of each studied WQ element in each sampling station while those double WQ values were inserted in a linear regression analysis. **Yielded linear equations (corrected WQ models), for each sensor, were accompanied by strong associations (R^2 ranging from 0.68 to 0.98).** Initial and corrected sensor-specific WQ models were combined with available *in-situ* WQ datasets, yielding 239 (Landsat) and 242 (Sentinel 2) matched pairs of 2019 and 220 (Landsat) and 286 (Sentinel 2) matched pairs of 2020, retrieved from the GEE platform. **Sensor-specific correction of WQ models was proven essential for some of them while RMSE values ranged for Chl-*a* from 11.68 μg/l (Landsat) to 14.88 μg/l (Sentinel 2), for Secchi depth from 2.02 m (Landsat) to 2.57 m (Sentinel 2) and TP from 0.14 mg/l (Landsat) to 0.09 mg/l (Sentinel 2), values that confirmed the stability and transferability of empirically developed models even when apply differently-from-DOS1 method corrected SR embedded in GEE platform.** In this way, valuable time concerning the images' pre-processing is saved while in parallel national lake WQ monitoring is facilitated by providing water resources managers an additional tool for taking water protection measures.

Since the transferability of hereby developed WQ models initially across different sensors and then when employing SR corrected with different AC methods embedded in GEE have been explored, one more ambiguous question has been examined; whether those universal models are efficient for WQ monitoring of oligotrophic Case-2 waters and then reach final conclusions whether there is a need for the development of special algorithms exclusively oriented to oligotrophic waterbodies. The classification of waters in Case 1 (oceanic) and Case 2 (coastal regions, rivers, and lakes), is characterized by great importance; Case 1 waters are determined by phytoplankton and co-varying substances, while Case 2 waters are more complex concerning their composition and optical properties. Oligotrophic lakes are classified as Case-2 rather than Case-1 waters since they typically receive significant levels of terrigenous input and their water clarity is primarily controlled by the concentration of Dissolved Organic Carbon (DOC). One of the main factors hindering accurate WQ monitoring in Case 2 waters is the fact that suspended material, yellow substances, and bottom reflectance vary independently of each other. In purview of the above, lake WQ quantitative models (Chl-*a*, Secchi depth and TP), developed and trained based on wide concentration ranges derived from WFD implementation, were applied to Landsat 8 OLI images illustrating two (2) Greek oligotrophic

waterbodies, Trichonis and Amvrakia lakes. The respective available *in-situ* WQ datasets was collected by HCMR staff from both of lakes and concern years 2013 and 2014 while satellite dates were concurrent with sampling ones. **Conclusively, application of hereby developed WQ models in oligotrophic Trichonis and Amvrakia lakes was ineffective: Chl-$a_{general}$ model yielded values of RMSE=1.9 μg/l, NRMSE=1.6 and median MAPE=256.8 %, Chl-$a_{natural}$ model yielded values of RMSE=1.8 μg/l, NRMSE=1.5 and median MAPE=176.6 % while results of all models of Secchi Disk and Total Phosphorus were statistically insignificant according to Spearman's rank correlation coefficient values at significance level 0.01.** Aforementioned results agree with those of other studies investigating oligotrophic waterbodies. Moreover, particularly low and homogeneous *in-situ* measured Chl-*a* concentrations, mean values equal to 0.6 μg/l (Trichonis) and 0.7 μg/l (Amvrakia), indicated that, in those lakes, the greatest optical contribution originates from non-algae particles implying the need for the development of special designed WQ algorithms.

Based on the previous approach, an effort has been made to develop special designed WQ algorithms in Trichonis lake. In the framework of this effort, the suitability of Landsat 8 OLI in accurately estimating Chl-*a*, nutrient concentrations and CDOM (Colored Dissolved Organic Matter) absorption at specific wavelengths was investigated. As a case study, the largest freshwater and oligotrophic body of Greece e.g. Trichonis Lake, is characterized by inexistent quantitative, temporal and spatial variability. Water samples were collected at 22 different stations on late August of 2014 and the satellite image of the same date was used to statistically correlate the *in-situ* measurements with various combinations of L8 bands in order to develop algorithms that best describe those relationships and calculate accurately the aforementioned WQ components. Subsequently, the most statistically promising quantitative models – accrued from statististical elaboration of 2014 data- were applied to the satellite image of 2013 and validation was conducted using *in-situ* data of 2013 as reference. **Results from the validation process showed a relatively variable statistical relationship between the *in-situ* data and reflectances (R logchl-*a*: 0.4, R NH_4^+: 0.7, R Chl-*a*: 0.5, R CDOM at 420 nm: 0.3).** *In-situ* nitrate, nitrite, phosphate and total nitrogen concentrations of 2014 were measured as lower than the detection limit of the instrument used, hence no statistical elaboration was conducted. **On the other hand, MLR analysis among reflectance measures and TP concentrations resulted in low and statistical insignificant correlations (R^2 values ranged from 0.06 to 0.07).** Our findings were concurrent with other studies in international literature, indicating that estimations for oligotrophic are less accurate than eutrophic and mesotrophic lakes, owing to the lack of suspended particles that are

detectable by satellite sensors. Yet, even with the presence of a lot of ground information as was the case in our study, a quantitatively accurate estimation of WQ constituents in coastal/inland waters remains a great challenge. Nevertheless, although those regression models, developed and applied to Trichonis oligotrophic lake are less accurate, may still be useful indicators of its WQ deterioration.

1. GENERAL INTRODUCTION

Lakes' significance

Surface freshwater is one of the most essential resources for the terrestrial ecosystem and the predominant source of drinking water on Earth (Whyte et al., 2018). Part of this resource is stored in lakes and reservoirs, while 14 million lakes (>10 ha) have been recorded in the world (Meyer et al., 2020). Lake water is used to satisfy environmental and human requirements while itplays a key role in the European and the global economy since it is exploited for civil (e.g., irrigation), industrial (e.g., processing and cooling, energy production, fishery) and recreational purposes. These activities, though, critically depend on a sufficient amount of freshwater. In particular, lakes in land-locked countries are valuable since they are among the most significant water sources (Knoll et al., 2019).

Given the great significance of lakes for human well-being, the rationale of universal access to water in quality and quantity is fundamental. Therefore, it is an objective of several global environmental agendas, like the United Nations Agenda for Sustainable Development (United Nations, 2015), which has 17 Sustainable Development Goals (SDG). Since lakes are a principal source of food and water supply, they are considered as essential ecosystems contributing to SDG 2, (Target 2.1. "*By 2030, end hunger and ensure access by all people, in particular the poor and people in vulnerable situations, including infants, to safe, nutritious and sufficient food all year round*") and SDG 6, sustainable management of water and sanitation (e.g., Target 6.1. "*By 2030, achieve universal and equitable access to safe and affordable drinking water for all*"). Furthermore, the employment of lakes in producing hydropower and supplying biomass participate as well to SDG 7, affordable and clean energy (e.g., Target 7.2. "*By 2030, increase substantially the share of renewable energy in the global energy mix*") and reduce reliance on fossil fuel (Inácio et al., 2022)

However, scarcity of freshwater resources is already perceptible, constraining development and societal well-being in many countries (Coppin et al., 2004), while the expected growth of global population over the coming decades, together with growing economic prosperity, is expected to increase water demand, aggravating those problems (Vörösmarty et al., 2000; Arnell, 2004; Alcamo et al., 2007; Schewe et al., 2013). Over the past few decades, the effect of climate change (global warming) and the anthropogenic pressure on natural resources have deteriorated their water quality (WQ; Michalak, 2016). The impacts of climate change on lake ecosystems are a well-studied topic, highlighting that alterations in temperature and precipitation patterns, result in regulation of other components such as water balance, limnology, and biogeochemical characteristics (Paulsson and Widerlund, 2022).

Intergovernmental Panel on Climate Change (IPCC) predicts that the increase in temperature will be higher for lakes of high latitudes than the global average (IPCC, 2021), due to their high exposure to atmospheric conditions (Schirpke and Ebner, 2022). Those lakes are further predicted to be subjected to a loss of perennial ice cover, and to stronger stratification in the water column due to inflowing snow melt, increasing duration of open water conditions, and shifts in the water balance. Climate warming affects as well lake ecological dynamics, in particulate of alpine (Schirpke and Ebner, 2022) and shallow lakes located in arid regions, where climate and hydrological regime exercise a strong control on water constituents' concentrations (Weyhenmeyer et al., 2019; Teubner et al., 2020).

Various direct and indirect human-induced pressures have severe impacts on ecosystem conditions and processes of lakes (Mammides, 2020), varying according to certain characteristics. Low elevation lakes are highly affected by overfishing and environmental pollution, whereas mountain lakes are less exploited due to their far proximity (Lyche Solheim et al., 2019). The lakes' shorelines are often transformed for touristic infrastructure, or by constructed dams, causing a severe degradation of littoral habitats, resulting in a decrease of biodiversity and ecological integrity (Porst et al., 2019). Accelerated touristic use, including aquatic recreation and hiking activities, alters species composition and ecosystem functions (Senetra et al., 2020; Tiberti et al., 2019). Agricultural activities in proximity to lakes as well as the use of fertilizers for intensive farming, lead to the World's lakes' greatest threat; increasing eutrophication of lakes and WQ degradation through increased nutrient inflows (Van Colen et al., 2018, Pedreros-Guarda et al., 2021).

Limnological research concerning the last two decades supports that existing global warming tends to intensify the responses of lakes to cultural eutrophication including accelerating hypolimnetic anoxia and nutrient release from lake sediments, intensified nutrient recycling, and increased algal production (Salmaso and Tolotti, 2021). Moreover, European surface water bodies were studied by European Environment Agency (EEA, 2018) concerning their impact from nutrient loads. Based on this report, 60% of the surface water bodies fail to achieve the objectives of good water quality defined by the international directives on water quality, such as the European Water Framework Directive (2000), with diffuse emission from agriculture being the second most important pressure affecting surface waters (Nikolaidis et al., 2022).

The need for sustainable management of water bodies highlights the fact that water resources are not inexhaustible and have limited resistance under anthropogenic pressures (ongoing drainage, conversion, and pollution). Hence, one of the most significant aspects for the sustainable management of water bodies is the constant monitoring of their quality, as well as of their

watersheds (Gholizadeh et al., 2016). WQ monitoring is the most crucial aspect for lake management while the term "lake management" refers to management designed to maintain an ongoing viability of lake ecosystems that provide the basis for aquatic and non-aquatic life (Bonansea et al., 2015).

Water Quality (WQ) monitoring

WQ is the most significant indicator of a water body's ecological status, while its assessment assumes the continuous monitoring of mainly physico-chemical and biological elements (Fatima, 2018; Nikolaidis et al., 2022). The continuous monitoring of large water bodies is a complex task, since it demands frequent and detailed data collection and interpretation efforts. Only exhaustive sampling field works can fully attain the spatial and temporal variance of common key WQ indicators. This results to an essential compromise between the number of sampling stations and the need of maintaining costs within reasonable limits (Strobel et al., 2000).

WQ parameters are traditionally measured based on *in-situ* measurements, collection of water samples and laboratory analysis (Li et al., 2016). Although *in-situ* WQ monitoring provides high accuracy (at specific location and time), it is a time-consuming procedure, and it cannot ensure a simultaneous WQ dataset on a regional or greater scale (Duan et al., 2013; Gholizadeh et al., 2016; Topp et al., 2020). Furthermore, traditional point sampling methods are not capable of detecting the spatial or temporal variations in WQ, as required in extensive assessment and management of water bodies. Additionally, patchy distribution of elements such as nutrients, algal blooms, and TSM (Total Suspended Matter) classify those methods as unsuitable for monitoring a large number of water bodies at a regional or national scale (Japitana and Burce, 2019).

Nowadays *in-situ* monitoring of lake WQ in synergy with satellite RS represents the latest scientific trend in many WQ monitoring programs worldwide (Japitana and Burce, 2019; Neil et al., 2019; Topp et al, 2020). Although the ability of RS to assess WQ is undeniable, this technique alone is not adequately precise and should be combined with field water sampling (Gholizadeh et al., 2016). Therefore, point-specific WQ datasets, which lack spatiotemporal trends, in conjuction with simultaneous RS datasets which provide a synoptic spatiotemporal view of ongoing earth surface processes, facilitate the monitoring, assessment and identification of WQ management strategies.

Contribution of RS in WQ monitoring

Recent developments in geoinformation technologies and in particular of RS and Geographical Information Systems (GIS), concerning pollution loads and lake WQ, offer a number of advantages that practically address the

limitations of traditional water sampling (Brivio et al., 2001; Pozdnyakov et al., 2005; Tyler et al., 2006). Among the key advantages of RS is the ability to cover large areas (Chatziantoniou et al., 2017) and to collect spectral information at variable spatial scales- including lakes that are otherwise inaccessible (MacKay et al., 2009; Whyte et al., 2018), at multi-scale temporal analysis and at dramatically lower cost compared to field measurements (Haddad and Harris, 1985).

Satellite RS is an efficient, beneficial tool for the assessment of spatial and temporal differentiations in WQ (Bonansea et al., 2015; Japitana and Burce, 2019). RS technologies enable researchers to acquire a unique, holistic perspective of the ecosystems. From the vantage point of space, satellite data becomes an invaluable tool in support of lake management while this is of especial importance in the context of the increasingly strict environmental regulations approved by governments worldwide such as Water Framework Directive (WFD; 2000/60/EC) and the European Marine Strategy Framework Directive (MSFD; 2008/56/EC) (Nikolaidis et al., 2022)

Apart from the law-required WQ components, the major factors which can influence the quality of inland water bodies are the suspended sediments (turbidity; Avdan et al. 2019), phytoplankton and cyanobacteria (i.e., chlorophylls, carotenoids), dissolved organic matter (DOM; Olmanson et al., 2020), organic and inorganic nutrients, pesticides, metals, thermal releases, macrophytic algae, pathogens and oils (Topp et al., 2020). The above-mentioned factors affect the optical properties of waters (except for nutrients); thus, directly change the signal acquired by optical sensors over water bodies (Gholizadeh et al., 2016). The parameters which can be directly quantified using RS techniques are the suspended particulate matter (SPM), which is placed in suspension by wind-wave stirring of shallow waters and can be a tracer for inflowing pollutants (Eleveld, 2012), the phytoplankton mainly as chlorophyll-*a* (chl-*a*) or phycocyanin (PC), that can be used to indicate the trophic level, to evaluate the presence of potentially toxic algal blooms and as a proxy of phytoplankton biomass (Pahlevan et al., 2020) and the coloured DOM (CDOM), commonly called yellow substances, which might indicates the presence of either fulvic or humic acids; CDOM is also investigated because of its role in protecting aquatic biota from ultraviolet solar radiation and its influence on specifically heterotrophic bacterial productivity in the water column, indicative of the shift from net autotrophy to net heterotrophy (Gholizadeh et al., 2016; Topp et al., 2020; Pizani and Maillard, 2022).

WFD and WQ monitoring

Several WQ monitoring programs, such as the US Clean Water Act (CWA) and Safe Drinking Water Act (SDWA), the Australian Reef Water Quality Protection Plan (Reef Plan) and Water Framework Directive (WFD)

have been implemented worldwide requiring large datasets of several WQ parameters to be monitored on a regular basis.

WFD, in particular, has been applied in a broad framework of catchment management while it provides a scheme for the conservation and improvement of inland, ground, and coastal waters' ecological status and aims to harmonize European legislation on water. Thus, pan-European hydromorphological, physicochemical, and biological datasets are used to determine ecological status of surface waters (Article 8) in order to assure and further improve future WQ and quantity (Mavromati et al., 2017; Nikolaidis et al., 2022). For each one category of datasets, a descriptive definition of high, good, moderate, poor, and bad status is given. Each National authority should set standards for those elements most relevant to the pressures faced by the water body under its responsibility and classify waters accordingly (Nikolaou et al., 2008).

Since the European Commission WFD (EC, 2000) was declared, Member States have started to establish lake ecological status assessment schemes, and integrating the setting of TP (Total Phosphorus) and Chl-*a* as reference conditions for European lakes of different types and ecoregions (Cardoso et al., 2007; Carvalho et al., 2008; Poikane et al., 2010; Huo 2013; Nikolaidis et al., 2022)

At the national level, the Greek Water Monitoring Network according to the Joint Ministerial Decision 140384/2011, operates for WFD and is implemented by the Goulandris Natural History Museum, Greek Biotope/Wetland Centre (EKBY). The monitoring network consists of 50 lake water bodies with an area of 0.5 km^2, including 26 artificial and 24 natural ones. At the majority of the lakes only one sampling station is detected, except for transboundary lakes (Megali Prespa, Mikri Prespa and Doirani), where two sampling stations are located. From the total of 53 sampling sites, the 27 are the surveillance and the 26 the operational ones. Surveillance stations operate in water bodies of good status for a certain period of time (one year in every monitoring cycle) unless during the previous monitoring period a specific lake system was determined to have reached the good condition and no changes are detected. The minimum monitoring frequency for the physicochemical parameters is 3 months unless longer time intervals are justified based on expert judgement. Furthermore, in the framework of surveillance monitoring, biological and hydromorphological parameters are monitored at least once during the whole period. On the other hand, operational stations run continuously on water bodies which fail to achieve good status while the suitable monitoring frequency for each element is determined based on the acquisition of sufficient enough data to provide a reliable assessment of the ecological status. In general, monitoring should be

performed at intervals not exceeding the limits listed in the table below (Table 1-1; WFD; 2000/60/EC).

Table 1-1. Specifications of Directive 2000/60/EC regarding monitoring frequency.

WQ element	Lakes
Biological	
Phytoplankton	6 months
Aquatic flora	3 years
Macroinvertebrates	3 years
Fish	3 years
Hydromorphological	
Hydrology	1 month
Morphology	6 years
Physicochemical	
Temperature	3 months
Oxygen	3 months
Salinity	3 months
Nutrients	3 months
pH	3 months
Pollutants	3 months
Priority substances	1 month

EO data and Landsat's contribution to lake WQ monitoring

Even though inland WQ measurements based on RS approaches dates back nearly 50 years (Topp et al., 2020) and during the last 20 years new instruments (platforms and sensors) have been developed for this purpose (Pizani and Maillard, 2022), a slow evolution has been observed compared to terrestrial and oceanic RS techniques. The effectiveness of each RS application for WQ monitoring depends on the selection of appropriate platforms and instruments (Pizani and Maillard, 2022) while there are several categories of the most commonly sensors used in WQ assessments, including airborne - along with unmanned aerial vehicles (UAV)- and satellite sensors (with visible and infrared wavelengths), passive microwave radiometers (MWR) and synthetic aperture radar (SAR). According to Gholizadeh et al. (2016), Sagan et al. (2020), Topp et al. (2020) and Pizani and Maillard (2022), the most utilized sensors for inland WQ assessment are the passive optical and thermal

ones, since water is highly absorptive within the near and shortwave infrared spectrum and the majority of water-leaving radiance occurs within the visible spectrum. In optically complex waters though, sediment reflectance exceeds the absorptive properties of water in the near/shortwave infrared wavelengths where high absorption within the visible spectrum results in low range of reflectance values. This low range demands high sensitivity to detect small changes in reflectance and therefore significant improvement concerning spectral, spatial and radiometric resolutions, revisit time, number of satellite bands and free access to data have been made (Topp et al., 2020). EO data from several optical ocean color sensors such as Moderate Resolution Imaging Spectroradiometer (MODIS), Earth Observing 1-Hyperion, and Medium Resolution Imaging Spectrometer (MERIS) and Sentinel-3 Ocean and Land Cover Instrument (OLCI) or optical land surface including Landsat series, Sentinel 2 A/B Multispectral Instrument (MSI) and Satellite Pour l' Observation de la Terre (SPOT), have been widely used for the study of surface WQ (Odermatt et al., 2018; Cao et al., 2019; Chelotti et al., 2019). In general, satellite data from the aforementioned sensors have been utilized in the development of models leveraging the relationship between a waterbody's optical qualities and its concentration of optically active water quality constituents (Topp et al., 2020).

According to Kutser (2009) and Matthews (2011), Advanced Land Imager (ALI) (30 m), Advanced Land Observation Satellite (ALOS) (10 m), SPOT-5 (10 m) and Landsat sensors, compared to sensors of higher spatial resolution, are characterised by a better radiometric performance which contributes to a more accurate assessment of the concentrations of quality parameters over water.

Landsat 7 (launched in 1999) introduced the Enhanced Thematic Mapper Plus (ETM+), whose analysis was similar to Thematic Mapper (TM) except for two bands, a 60 m thermal and a new 15 m panchromatic band, respectively (Loveland and Dwyer 2012). Since 2003, Landsat 7 had a sensor deficiency where the Scan Line Corrector (SCL) was off and even though those images are characterized by black line gaps (Tebbs et al. 2013), its radiometric and geometric analyses remain undisturbed (Bonansea et al. 2015). On May 30, 2013, data from the Landsat-8 satellite (launched on 11 February, 2013) became available allowing the continuance of studies on WQ of lakes (Giardino et al., 2014); this satellite bears two sensors, the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIR) while it includes a narrower near-infrared band, and a 12-bit radiometric resolution compared to the 8 bits of previous Landsat satellites (Olmanson et al. 2016; Bonansea et al., 2018).

Moreover, Landsat 9 (OLI-2) was successfully launched on Monday, Sept. 27, 2021 continuing the Landsat program's critical role in monitoring,

understanding and managing the water resources needed to sustain human life. Furthermore, since June 2015, Sentinel 2 mission provides simultaneous image data with those of Landsat 8 OLI offering great opportunities for long term high-frequency WQ monitoring (Mandanici and Bitelli, 2016) through building time-series.

The Sentinel-2 mission carries two satellites, Sentinel-2A and Sentinel-2B. They are both equipped with identical Multispectral Instruments (MSI) capable of acquiring data at 13 bands at different spatial resolutions (between 10 m and 60 m) while the revisit frequency of each satellite is 10 days. Landsat (30 m spatial resolution) and Sentinel-2 (10–60 m spatial resolution) missions provide fine-scale spatial data and have been reported to be suitable for the quantification of multiple WQ indices in freshwater lakes and reservoirs (Allan et al., 2011; Giardino et al., 2014; Kim et al., 2014; Markogianni et al., 2014; Bresciani et al., 2018; Markogianni et al., 2018; Bramich et al., 2021).

Inland waters, and especially lakes, are small water bodies that are not detected by current ocean color satellites, and even though this lack prevents the monitoring and estimation of their WQ components, it has been replenished by the use of Landsat, Sentinel-2, and ASTER (Advanced Spaceborne Thermal Emission and Reflection Radiometer) multispectral images. Their fine spatial resolutions (10~60 m) enable them to resolve small freshwater lakes and rivers more than a few hundred meters wide. Therefore, the application of those images has been preferred for freshwater lake mapping projects (Wang et al., 2020). Furthermore, recent reviews of state-of-the-art RS-based approaches by Gholizadeh et al. (2016) and Pizani and Maillard (2022) underpin the use of particularlyLandsat sensors, TM (Thematic Mapper), MSS (Multi-Spectral Scanner), ETM (Enhanced Thematic Mapper), OLI (Operational Land Imager) and OLI-2 (Landsat 9) as fairly successful choices to assess the important WQ parameters, including Chl-*a*, Secchi Disk Depth (SDD), TP, and trophic status.

Although Landsat sensors were not designed for aquatic applications (Kutser, 2012; McCullough et al., 2012a), there are numerous examples of Landsat images' employment for estimating and/or monitoring lake WQ. Several studies have proposed reliable algorithms between Landsat data and WQ parameters, including chlorophyll; phytoplankton and phycocyanin concentrations (Yacobi et al., 1995; Vincent et al., 2004; Brezonik et al., 2005; Tyler et al., 2006; Torbick et al., 2008; Karakaya et al., 2011; Tebbs et al., 2013), water clarity (Stadelmann et al., 2001; Hadjimitsis et al., 2006; Olmanson et al., 2008; Guan et al., 2011; Zhao et al., 2011; McCullough et al., 2012a), CDOM (Brezonik et al., 2005; Zhu et al., 2014; Brezonik et al., 2015), blooms of cyanobacteria (Vincent et al., 2004), macrophyte (Albright and Ode, 2011) and TSM (Guang et al., 2006; Zhou et al., 2006; Onderka and Pekárová, 2008; Kulkarni, 2011; Bonansea and Fernandez, 2013;).

Traditionally used approaches for lake WQ monitoring via satellite RS

The most commonly used approach to monitor WQ of inland waters via RS involves fitting a standard linear regression between spectral band/band ratio values and temporally coincident *in-situ* WQ measurements (e.g. Topp et al., 2020).

In general, according to Topp et al. (2020), there are three (3) well-documented methodologies to estimate the concentration of WQ elements in inland waters: empirical, semi-empirical, and physical or analytical methodology (Table 1-2). Empirical methods attempt to establish relationships between *in-situ* WQ measurements and water leaving radiance measured by the sensor without the precondition of prior understanding of the complex water and light interactions. Those relationships imply effective data improvement but limited transferability (Austin and Petzold, 1981). Moreover, empirical methods incorporate machine learning techniques, which are differentiated by their robust ability to handle complicated non-linear relationships, typical of WQ remote sensing data (Sagan et al., 2020; Topp et al., 2020). Machine learning algorithms include artificial neural networks (ANN), genetic algorithms (GA), support vector machines (SVM), random forest regression trees, and empirical orthogonal functions (Topp et al., 2020). On the other hand, through semi-empirical techniques, spectral and physical knowledge of studied WQ constituents' properties are combined and then correlated to the *in-situ* concentrations. Regarding physical or analytical approaches, the acquisition of certain biogeochemical parameter values (e.g., Chl-*a*, CDOM) is required, as well as inherent (IOP) and apparent optical properties (AOP), and those models are based on radiative transfer and calibrated using field observations.

Although analytical methods, including fuzzy logic and Principal Component Analysis (PCA), have already been extensively used, empirical and semi-empirical predicting models are still widely utilized (Gholizadeh et al., 2016). Analytical methods' complexity in terms of their theory and calculation difficulties (Gholizadeh et al., 2016) and the non-availability of required detailed spectral information of the optically active water constituents (optical properties, radiometric quantities) have contributed to the maintenance and development of empirical models. This trend is further observed especially in cases where machine learning models are utilized, as most of them reduce overall error and maximize model fit (Topp et al., 2020). However, it should be noted that empirical algorithms are more specific to certain water types, regional or optical (Odermatt et al., 2012). It should also be noted that semi-analytical methods are superior to empirical ones mainly concerning the reliability of results and the fact that no *in-situ* data are required afterwards for recalibrating the retrieval algorithm. On the other hand, those approaches require the utilization of a spectroradiometer and the

collection of *in situ*-measured Rrs spectra including the radiance of skylight, radiance from a standard gray board, and the total upwelling radiance from the water (Jiang et al., 2019).

Table 1-2. Summary of the common approaches for remote sensing of lake WQ monitoring.

Modelling Approach	including	Advantages	Limitations
Empirical		Easily interpretable, no a priori assumptions required	Limited transferability (dependant on range of training data), uncapable of handling non-linear relationships
	Machine learning techniques	Capable of handling complicated non-linear Relationships, no a priori assumptions required	Computationally expensive, risk of overfitting
Semi-Empirical		Easily interpretable	Uncapable of handling non-linear relationships
Physical/ Analytical		Theoretically generalizable	Computationally expensive, knowledge of optical properties and collection of *in-situ* Rrs spectra are required

WQ elements measured by RS

There are numerous parameters measured for WQ monitoring; the optically-active ones (e.g. Chlorophyll-*a*, transparency, turbidity, total suspended matters, coloured dissolved organic matter, true colour, temperature) can be monitored remotely and others that are non-optically active (total phosphorus, total nitrogen, pH, dissolved oxygen) and can be assessed indirectly through their relationship with the optically active ones (as proxies; Gholizadeh et al., 2016; Pizani and Maillard, 2022).

Chlorophyll-*a* concentration, is indicative of phytoplankton abundance in waters, and can be directly quantified using EO techniques implying the trophic level, the existence of toxic algal blooms and the phytoplankton biomass (Randolph et al., 2008; Ruiz-Verdu et al., 2008). Chl-*a* is the major indicator of trophic state and considered as one of the top water pollution indices related to public health, eutrophication, and deterioration of ecosystem habitat.

Findings from numerous published studies have indicated that biological and chemical water quality parameters such as Chl-*a* have distinctive spectral characteristics and can be measured using spectral indices. A variety of spectral indices derived from remote sensing data based on empirical or semi-empirical relationships have been developed for

transforming spectral data into WQ parameters. These indices may involve three (Song et al., 2013; Sun et al., 2014; Huang et al., 2014;) and four spectral bands (Le et al., 2009). The majority of spectral indices are based on reflectance ratios of two spectral bands (near infrared and red) for operational purpose. A band ratio between the near infrared (NIR, ~0.7 μm) and Red (~0.6 μm) has frequently been used to estimate Chl-*a* in waters due to a positive reflectivity of Chl-*a* in the NIR and an inverse behavior in the red (Rundquist et al., 1996; Pepe et al., 2001) while NIR and red bands are involved in most indices (Yang et al., 2017). Spectral band ratios are generalizable and easily applicable across wide geographic ranges. These indices, however, appear to be less reliable in diverse water bodies including lakes, ponds, rivers and streams in coastal regions (Yang et al., 2017; Sagan et al., 2020), as they assume constant water and atmospheric conditions. This assumption may result in significant estimation errors, especially when applied across time series. Therefore, spectral indices are proposed to identify spatial distribution of WQ rather than make exact predictions (Sagan et al., 2020).

On the other hand, few studies have aimed to assess and model nutrient concentrations, due to their weak optical characteristics and low signal noise ratio as nutrients constitute non-optically active WQ constituents (Gholizadeh et al., 2016). Nutrient models have not yielded statistically strong results or at least similar as of those constituents that have optical properties (Dewidar and Khedr, 2001; Wu et al., 2010; Chen and Quan, 2012; Isenstein and Park, 2014). Indirect methods, however, can be utilized to estimate nitrogen (N) and phosphorus (P) concentrations. RS has been widely demonstrated as an effective solution for detecting the relationship between algae concentration and corresponding nutrients (Sagan et al., 2020). Nitrogen (N) and phosphorus (P) are vital micronutrients for algae, while P (existing either in a particulate or dissolved phase) is the key limiting nutrient responsible for eutrophication in most lakes (Correll, 1999). In general, special attention should be paid depending on which nutrient is growth limiting, as in one water body the correlation with Chl-*a* might be with N, while in a different water body the correlation might be with P (Topp et al., 2020). Total phosphorus (TP) estimation via RS has been explored due to its high correlation with optically active constituents (Kutser et al., 1995; Wang et al., 2004; Wu et al., 2010) since it cannot be measured directly using optical RS instruments. The Chl-*a* and TP relationship has been investigated in individual lakes (Smith, 1982; Malve and Qian, 2006), and it is well documented to be accompanied by a strong and positive correlation among lakes (Healey and Hendzel, 1979; Busse et al., 2006). As TP is highly correlated to Chl-*a* concentration, and TSM usually reflects TP loading, TP is also closely related to Secchi depth (SD) with an exponential equation according to Carlson's findings (Carlson, 1977).

As the algae and suspended inorganic matter increase in a lake, the depth to which light can penetrate is reduced (Fuller et al., 2004). Therefore, SD is often used as a trophic state indicator (Carlson, 1977). Generally, there are two methodologies followed to retrieve Secchi Disk Transparency (SDT) based on RS data. Empirical approach estimating SD through regression analysis and semi-analytical approach retrieving SD based on an underwater visibility theory (Jiang et al., 2019). Regarding empirical models, reflectance at the red spectrum has been almost globally used to retrieve water clarity (Baban, 1993; Nelson et al., 2003; Wu et al., 2008; McCullough et al., 2012; Hicks et al., 2013) since increased brightness is accompanied by decreased water clarity (Matthews, 2011). Moreover, further studies have also documented the usefulness of spectral response of the blue, green, and near-infrared spectral bands in combination with *in-situ* measurements of SD and Chl-*a* concentrations in predicting water clarity for inland lakes (Avdan et al., 2019). Since water clarity has long been proven to interact with nutrient availability and Chl-*a* concentrations within lakes (Song et al., 2022), RS studies frequently use it to assess overall lake trophic status (oligotrophic, mesotrophic, or eutrophic).

Utilization of multi sensor image data for lake WQ monitoring

Effective and accurate remote sensing of lake WQ requires frequent *in-situ* time series WQ data accompanied by simultaneous satellite images. Performance of high-frequency time series and multi-temporal analyses becomes more possible when multi sensor image data is available (Mandanici and Bitelli, 2016). Furthermore, the use of various multi-spectral sensors, with different radiometric characteristics- makes possible to measure many of the WQ parameters required by law (Mantas et al., 2013; Mandanici and Bitelli, 2016).

In the framework of this research, the emphasis has been mainly given on the combination initially of Landsat sensors (7 ETM+;8 OLI) and then of Landsat and Sentinel 2 image data. This selection was based on the fact that the majority of available *in-situ* WQ data were recorded during 2013–2015, hence images of sensors Landsat 7 ETM+ and Landsat 8 OLI were the exclusive choice for the implementation of current research. Moreover, images from both platforms have been proven particularly valuable for inland lakes while both offer free open access data-archive (Deutsch et al., 2018).

Incorporation of Landsat sensors was attempted to increase the temporal range of observations; temporal resolution is sixteen (16) days while Landsat 7 ETM+ and 8 OLI together, provide four (4) satellite images for every 32 days (Pedreros-Guarda et al., 2021). Based on relevant literature review, Landsat 7 ETM+ and Landsat 8 OLI images have similar spatial resolution (30 m), are statistically comparable and homogeneous over WQ sample sites (Wang et al., 2020) while both have similar spectral band placements for the Blue (ETM+

band 1, 0.45–0.52 μm; 8 OLI band 2, 0.45– 0.51 μm) and Green bands (ETM+ band 2, 0.52–0.60 μm; 8 OLI band 3: 0.53–0.59 μm). Differences are particularly observed in the NIR (ETM+ Band 4, 0.76–0.90 μm; 8 OLI Band 5, 0.85–0.88 μm) and to a lesser extent in Red bands (ETM+ Band 3, 0.63–0.69 μm; 8 OLI Band 4, 0.64–0.67 μm) (Olmanson et al., 2016; Deutsch et al., 2018). Moreover, Landsat 8 OLI is characterized by 12-bit radiometric resolution, higher signal to noise ratios, increased spectral bands and narrower near-infrared bands compared to ETM+, features that have contributed to a more accurate monitoring of freshwater quality (Li et al., 2021).

Sentinel 2 MSI data were then selected initially based on their significant match with the corresponding spectral bands of Landsat 8 OLI data and then based on their high spatial resolution (10m at visible and near infrared bands) and short revisit interval (5 days; Li et al., 2021). Joint use of Landsat and Sentinel 2 images achieve globally a 2-3-day revisit time (Li et al., 2021) while since both platforms are characterized by 12-bit quantization, provide an improved radiometric quality resulting in an also improved inland water monitoring (Mandanici and Bitelli, 2016). Conclusively, combined use of Landsat and Sentinel 2 data grants access to a greater amount of satellite images while facilitates high frequency time series analyses.

On the other hand, when multi sensor image data are combined, a number of conceptual and technical challenges may accrue originating from their orbital, spatial and spectral differences (Deutsch et al., 2018). Moreover, even though Mandanici and Bitelli (2016) highlighted a significant match between Landsat 8 OLI and Sentinel 2 MSI spectral bands, differences in the recorded radiometric values were also observed. What is important though, concerning those differences, is the application and the approach adopted to implement multi-sensor time series analyses. On one hand, many empirical approaches based on multispectral indices be more affected by the problem (Werff and Meer, 2016) but when methods and processing are applied separately on discrete images and the training is also independent, results are less affected (Mandanici and Bitelli, 2015; 2016). Furthermore, independent elaboration of only-one sensor images does not require implementation of a resampling procedure, which is mostly essential in change detection analyses, given the different spatial resolution of the two sensors (e.g. Landsat 30 vs. Sentinel 10 m).

Big Earth Data Cloud Processing Platforms-the GEE platform

Despite the advantages that RS offers compared to field works, computing WQ properties from RS images may become time-consuming and complicated because of the processing data chain that a large-scale WQ assessment and long-time-series analyses demand (Kumar and Mutanga, 2018). Cloud computing has emerged as a significant tool to process Big Data

with main advantages the convenient access and processing of big geospatial data and substantial computational capabilities (Zhao et al., 2022). Some of the most currently popular Big Earth cloud processing platforms include Google Earth Engine (GEE), Amazon Web Services (AWS), Microsoft Azure, NASA Earth Exchange (NEX), Sentinel Hub (SH), Processing and Analysis for Land Monitoring (SEPAL), open EO and Open Data Cube (ODC). Gomes et al. (2020) compared seven platforms for big EO data regarding the following criteria: data abstraction, processing abstraction, physical infrastructure abstraction, open governance, reproducibility of science, infrastructure replicability, processing scalability, storage scalability, data access interoperability and extensibility. Based on this evaluation but as well to Zhao et al. (2022) and Pizani and Maillard (2022), GEE is the most significant cloud processing platform for the remote sensing community due to its ease of use and maturity.

Google Earth Engine (GEE) platform has emerged as a valuable tool for geospatial analysis of image data from various satellites based on open source Cloud Computing (Bioresita et al., 2021). In addition to this, several involved operators such as the United State Geological Survey (USGS), National Aeronautics and Space Administration (NASA), and European Space Agency (ESA) -among others- are collaborating with Google Inc. and have made satellite data available online through the Google Earth Engine (GEE) cloud platform (Wang et al., 2020).

GEE provides a Javascript API (Application Programming Interface) and a Python API for data management and analysis while offers a data catalog that stores a large repository including among others geospatial data, environmental variables, climate forecasts, land cover and topographic datasets (Gomes et al., 2020). Concerning optical imagery of satellites, the data repository of GEE includes among others the entire datasets collected by Landsat 4/5/7/8, Sentinel 1/2, and ASTER while it is updated on a daily basis with around 6000 new image scenes. The GEE offers a parallel computation capability and utilizes many processors to conduct individual tasks, hence accelerating the time-consuming computing, required for large-scale applications. Moreover, satellite images are pre-processed to various processing levels and products, such as surface reflectance, top of atmospheric reflectance (TOA), and vegetation indices (Wang et al., 2020).

A plethora of recent studies have been published employing GEE for inland waters' WQ monitoring (Jia et al., 2019; Zong et al., 2019; Maeda et al., 2019; Wang et al., 2020; Weber et al., 2020; Somasundaram et al., 2021; Bioresita et al., 2021; Lobo et al., 2021; Vaičiūtė et al., 2021; Kislik et al., 2022; Wen et al., 2022). Wang et al. (2020) used GEE to automatically search matching cloud- and haze-free image pixels across multiple sensors using online scripts for Chl-*a* samples while a SVM was trained and eventually predicted Chl-*a* concentrations with reasonable accuracy. Wen et al. (2022)

established empirical models between satellite reflectance- derived from archived Landsat images embedded in GEE- and *in-situ* TSM observations over 426 Chinese lakes during a 10-year time span (2011-2020) and managed to confirm their temporal stability and suitability for examination of long-term TSM trend in lakes. Bioresita et al. (2021), assessed Chl-*a* and TSS concentrations through certain formulas along Kali Porong estuary (Indonesia) and validated their results by utilizing Sentinel-2 reflectance values retrieved from GEE platform. Validation of results was conducted by using available *in-situ* data and was accompanied by high correlation values (Chl-*a*: 0.654; TSS: 0.652).

Case 1-Case 2 waters and oligotrophic waterbodies

The classification of waters in Case 1 (oceanic) and Case 2 (coastal regions, rivers, and lakes), refined by Morel and Gordon (1983), is characterized by great importance when RS techniques are utilized to monitor their WQ and/or trophic status. The distinction between the two cases has some significant effects on the interpretation and modelling of optical data. In particular, according to this classification scheme, the optical properties of Case 1 waters are determined by phytoplankton and co-varying substances, while Chl-*a* is considered a proxy of phytoplankton concentration. This assumption has facilitated the implementation of large-scale optical models and the development of Chl-*a* predicting algorithms for Case 1 waters (Markogianni et al., 2022).

On the other hand, single variable models should be abandoned when Case 2 waters are the case. It is, on the whole, acknowledged that Case 2 waters are more complex than Case 1 concerning their composition and optical properties. Monitoring the WQ of Case 2 waters is a more sophisticated task since phytoplankton, suspended material, yellow substances, and perhaps bottom reflectance vary independently of each other. The main difficulty lies in the fact that the alterations in optical signal and the concentrations of the dissolved constituents are often so small that they hinder the ability to extract reliable information or the optical signal may be affected in a similar way by more than one substance, which results in an inability to discriminate the different materials (Gholizadeh et al., 2016). Runoff and discharges from rivers/streams are also one of the main factors adding to the complexity of the water constituent retrieval process in Case 2 waters while those inflows from streams introduce different organic/inorganic particles, known as total suspended solids (TSS).

Hence, given the difficulty that WQ monitoring of Case 2 waters constitutes a multi-variable, non-linear problem, it is more realistic to establish a series of algorithms rather than a single all-purpose one. In this way, more than one algorithm will attempt to capture and solve the problem

for all variables and over several and different ranges of concentrations (IOCCG, 2000).

In parallel, the Case 1/Case 2 classification can substantially improve RS products when associated with individual optical water types (OWTs). In particular, coastal regions and inland waters are characterized by such optical diversity that any further information about their variability in IOPs and biogeochemical significance would be particularly valuable. Some OWTs can be hypereutrophic waters, turbid waters with high organic content, sediment-laden waters, CDOM-rich waters, or even very clear blue waters. Several hierarchical, partitional, and hybrid clustering techniques have been utilized to further discriminate distinct OWTs within and between Case 1 and Case 2 waters (Spyrakos et al., 2018). After all, a reliable OWT classification optimizes the selection of the finest constituent algorithms when simpler approaches cannot yield reliable results.

The different ranges of concentrations within Case 2 waters correspond to classes of trophic status. Carlson (1977) developed a method of trophic status classification for inland waters considering Chl-*a* and phosphorus concentrations and Secchi disk depths (*ZSD,* m). Ranges of those WQ elements are associated with three (3) main trophic classes: oligotrophic, mesotrophic and eutrophic (McCullough, 2012) including also transitional categories (e.g. ultra-oligotrophic, hypertrophic; etc; Watanabe et al., 2020).

Based on this rationale, very clear lakes are classified as oligotrophic Case-2 rather than Case-1 (Gons et al., 2008) since they typically receive significant levels of terrigenous input (Gons and Auer, 2004) and their water clarity is primarily controlled by the concentration of Dissolved Organic Carbon (DOC) (Lisi and Hein, 2019; Song et al., 2022). Water clarity, in turn, affects a plethora of chemical, physical and biological processes, including thermal structure, light transmission for photosynthesis, attenuation of damaging levels of ultraviolet light, vertical distribution of plants and animals, as well as the form and availability of toxic metals (Schindler et al., 1997; Williamson et al., 1999a; Pérez-Fuentetaja et al., 1999; Gunn et al., 2001).

In purview of the above and based on the relevant literature, it has been reported that there is a need for further algorithm development, especially for oligotrophic water bodies, while, of principle value is the selection of the appropriate wavelengths.

Gons and Auer (2004) measured spectra in the Keweenaw Bay (Lake Superior) which were typical of oligotrophic lacustrine waters. However, strong absorption by water in the red region hindered the accurate detection of Chl-*a* highlighting the need of algorithm development for oligotrophic waterbodies.

Furthermore, Gons et al. (2008) managed to adequate assess Chl-*a* concentrations of the Great Lakes (North America) through an empirically developed algorithm employing blue-to-green bands. Additionally, there is a

plethora of studies highlighting the utilization of blue-to-green band ratios as the most optimal choice for the monitoring of WQ elements in oligotrophic lakes (Binding et al. 2019; O'Reilly and Werdell, 2019; Warren et al., 2019)
Even though for clear waters the results warrant the use of the blue-green ratio, blue-green algorithms are not suitable in turbid regions. Therefore, AC processors need to further improve so that the NIR-red band ratio algorithms can be used in more turbid waterbodies (Case-2 waters; Warren et al., 2019)

1.1 Scope and objectives of the current thesis

All in all, the present PhD thesis constitutes, to the author's knowledge the first attempt to achieve the continuous monitoring and assessment of WQ and trophic state of Greek lakes. Taking advantage of the ongoing implementation of WFD in Greece, collection of large *in-situ* WQ datasets in synergy with satellite RS, will further provide the essential means for the monitoring of lake eutrophication and its spatio-temporal changes. Hereby delivered "tools" will be proven fundamentally valuable in the framework of national environmental policy in general, and in particular of lake management at a national scale.
Overall, the main objectives of this research are to:

1. Establish a methodological framework that aims to model WQ and trophic status of optically diverse Greek lakes (Case 2 waters) by assessing key WQ elements with fine spatial resolution (10-30 m) RS image data. Ultimate goal of this proposed methodology is the accurate spatial assessment of WQ and trophic status over various types of lakes, thus acquiring the valuable information about their variability.
   The unique contribution of this objective lies in the fact that spatially distributed WQ of Greek lakes can be monitored continuously, for the first time, reflecting their trophic status and detecting the possible pollutant threats in near real time and in fine spatial resolution.
2. Explore the spatio-temporal transferability of Landsat-developed WQ models across sensors; initially across Sentinel 2 and then across multi-sensor image data (Landsat 7 ETM+, 8 OLI and Sentinel 2 MSI). The transferability is tested along the National Lake Network Monitoring of Greece (WFD) and concerns the sampling campaigns of 2018. In particular, the smoothly and operating transferability of WQ models across different sensors will facilitate the acquisition of high-frequency time series and multi-temporal WQ analyses, further contributing to continuous lake WQ monitoring at a national scale (Greece).
3. An examination of the influence of different atmospheric correction methods to WQ models' performance after employing differently-atmospherically corrected SR values. Statistically-modified WQ models

harmonize the differences accrued from the application of DOS1 (manually applied to images of all sensors) and LaSRC, LEDAPS and Sen2Cor correction methods applied to Landsat 8 OLI, Landsat 7 ETM+ and Sentinel 2 images, respectively through the GEE platform. This analysis is performed across the National Lake Network Monitoring of Greece (WFD) and the comparison of results is based on the *in-situ* WQ data of years 2018, 2019 and 2020. High performance of WQ models employing SR from GEE environment further contributes to the continuous lake WQ monitoring across Greece in an even faster manner, whilst liberate researchers from the time-demanding and complicated atmospheric correction of raw image products.

4. Assess WQ models' performance in a distinct category of optically complex Case-2 waters, oligotrophic Trichonis and Amvrakia lakes. The unique contribution of this objective lies in the final decision on whether national WQ models adequately support perpetual WQ monitoring of Greek oligotrophic lakes or special oligotrophic algorithms should be developed and under which circumstances. Furthermore, it includes the background information required for the designation of the WQ monitoring methodology of oligotrophic waterbodies.
5. Model WQ of oligotrophic Trichonis lake by assessing WQ key elements (Chl-*a*, nutrient concentrations and CDOM absorption at 420 nm) through satellite RS. Trichonis is the largest freshwater lake of Greece while the available *in-situ* and satellite datasets concern years 2013 and 2014.

1.2 Thesis 's Significance

The most significant aspect concerning the contribution of the present PhD thesis lies in the fact that the methodology has been developed, applied and validated in 50 different Greek lake systems of varied chemistry, limnological conditions and trophic level, while covering a broad geographic area and a wide range of WQ elements' concentrations collected over different seasons. The hereby developed WQ models were proven to efficiently accommodate initially the spectral composition differences among Landsat (7 ETM+, 8 OLI) and Sentinel 2 sensors and then the differences regarding the pre-processing procedures among SR values that are subjected to different atmospheric correction methods (DOS1, LaSRC, LEDAPS and Sen2Cor).

Wide WFD *in-situ* lake WQ datasets in conjuction with satellite images managed to generate uniform models for the systematic assessment of Chl-*a* and TP concentrations and Secchi Disk depths at a greater scale (country

level), compared to the majority of the respective literature focusing on regional scales and discrete inland water bodies.

Furthermore, WQ models exhibited spatial and temporal stability to variations of the optical properties of lakes while their good performance when employing SR retrieved from GEE platform facilitates and significantly improves and accelerates the perpetual lake WQ and trophic status monitoring especially when *in-situ* data are limited. Addittionally, the detailed spatial variability of WQ and trophic status over lakes is delivered spatially finer compared to similar, large-scale, purely based on EO applications, offered worldwide (e.g. SDG6 Hydrology TEP Reporting portal; 90m spatial resolution). By extension, WQ empirical models were also proved priceless means for the monitoring of lake eutrophication and the drivers of its dynamics, particularly nowadays that this phenomenon has been evolved into a growing public concern and lakes are undergone the dual impact of human activities and climate change.

Ultimate goal of this thesis and the delivered WQ models is to constitute a valuable tool, part of a wider national early warning system, in the hands of scientists and competent public authorities for the timely identification of pollution events and by extension the promptly performance of sustainably efficient solutions. Moreover, what is the most desired is the uninterrupted continuation of WFD implementation in Greece as the on-going combination of RS and WFD *in-situ* data will further improve the temporal resolution of lake WQ monitoring while offering a multi-platform observation by acquiring more comprehensive information.

1.3 References

Albright, T.P.; Ode, D.J. Monitoring the dynamics of an invasive emergent macrophyte community using operational remote sensing data. Hydrobiologia **2011**, 661:469-474.

Alcamo, J.; Flörke, M.; Märker, M. Future long-term changes in global water resources driven by socio-economic and climatic changes. Hydrol. Sci. J. **2007**, 52, 247–275.

Allan, M.G.; Hamilton, D.P.; Hicks, B.J.; Brabyn, L. Landsat remote sensing of chlorophyll a concentrations in central north island lakes of New Zealand. Int. J. Remote Sens. **2011**, 32, 2037–2055.

Arnell, N.W. Climate change and global water resources: SRES emissions and socio-economic scenarios. Glob. Environ. Chang. **2004**, 14, 31–52.

Austin, R.W.; Petzold, T.J. Water Colour Measurements. In Oceanography from Space; Gower, J., Ed.; Plenum: New York, NY, USA, **1981**; pp. 239–256.

Avdan ZY, Kaplan G, Goncu S and Avdan U. Monitoring the water quality of small water bodies using high-resolution remote sensing data. ISPRS International Journal of Geo-Information **2019**, 8(12): 553.

Baban, S.M.J. Detecting water quality parameters in the norfolk broads, U.K., using landsat imagery. Int. J. Remote Sens. **1993**, 14, 1247–1267.

Binding, C.E.; Zastepa, A.; Zeng, C. The impact of phytoplankton community composition on optical properties and satellite observations of the 2017 western Lake Erie algal bloom. J. Great Lakes Res. **2019**, 45, 573–586. https://doi.org/10.1016/j.jglr.2018.11.015.

Bioresita, Filsa *et al IOP Conf. Ser.: Earth Environ. Sci.* **2021**, 936 012011

Bonansea, M.; Bazán, R.; Ledesma, C.; Rodriguez, C.; Pinotti, L. Monitoring of regional lake water clarity using Landsat imagery. Hydrology Research **2015**, 46 (5), 661–670.

Bonansea, M.; Fernandez, R.L. Remote sensing of suspended solid concentration in a reservoir with frequents wildland fires on its watershed. Water Science and Technology **2013**, 67(1), 217–223.

Bonansea, M.; Rodriguez, M.C.; Pinotti, L.; Ferrero, S. Using multi-temporal Landsat imagery and linear mixed models for assessing water quality parameters in Río Tercero reservoir (Argentina). Remote Sens. Environ. **2015**, 15, 28–41.

Bonansea, Matias; Rodriguez, Claudia; Pinotti, Lucio. Assessing the potential of integrating Landsat sensors for estimating chlorophyll-a concentration in a reservoir. Hydrology Research **2018**, 49 (5): 1608–1617.

Bramich, J.; Bolch, C.J.S.; Fischer, A. Improved red-edge chlorophyll-a detection for Sentinel 2. Ecol. Indic. **2021**, 120, 106876.

Bresciani, M.; Cazzaniga, I.; Austoni, M.; Sforzi, T.; Buzzi, F.; Morabito, G.; Giardino, C. Mapping phytoplankton blooms in deep subalpine lakes from Sentinel-2A and Landsat-8. Hydrobiologia **2018**, 824, 197–214.

Brezonik, P. L; Olmanson, G. L.; Finlay, J.C.; Bauer, M. E. Factors affecting the measurement of CDOM by remote sensing of optically complex inland waters. Remote Sensing of Environment, **2015**, vol. 157, pp. 199-215, ISSN 0034-4257, http://dx.doi.org/10.1016/j.rse.2014.04.033.

Brezonik, P.; Menken, D.; Bauer, M. Landsat-based remote sensing of lake water quality characteristics, including chlorophyll and colored dissolved organic matter (CDOM). Lake and Reservoir Management **2005**, 21(4), 373–382.

Brivio, P.A.; Giardino, C.; Zilioli, E. Determination of chlorophyll concentration changes in Lake Garda using an image-based radiative transfer code for Landsat TM images. Int. J. Remote Sens. **2001**, 22, 487–502.

Busse, L.B.; Simpson, J.C.; Cooper, S.D. Relationships among nutrients, algae, and land use in urbanized southern California streams. Can. J. Fish. Aquat. Sci. **2006**, 63, 2621–2638.

Cao Z, Ma R, Duan H, Xue K and Shen M. Effect of satellite temporal resolution on long-term suspended particulate matter in inland lakes. Remote Sensing **2019**, 11(23): 2785.

Cardoso, A. C.; Solimini, A.; Premazzi, G.; Carvalho, L.; Lyche A, Rekolainen S. Phosphorus reference concentrations in European lakes. *Hydrobiologia* **2007**, 584(1): 3–12.

Carlson, R.E. A trophic state index for lakes. Limnol. Oceanogr. **1977**, 22, 361–369.

Carvalho, L.; Solimini, A.; Phillips, G.; Berg, M.; Pietilainen, O. P.; Solheim, A. L et al. Chlorophyll reference conditions for European lake types used for intercalibration of ecological status. *Aquatic Ecology*, **2008**, 42(2): 203–211.

Chatziantoniou, A.; Psomiadis, E.; Petropoulos, G. Co-orbital sentinel 1 and 2 for LULC mapping with emphasis on wetlands in a Mediterranean setting based on machine learning. Remote Sens. (Basel) **2017**, 9 (12), 1259. https://doi.org/10.3390/rs9121259.

Chelotti GB, Martinez JM, Roig HL and Olivietti D (2019). Space-temporal analysis of suspended sediment in low concentration reservoir by remote sensing. RBRH 24.

Chen, J., Quan, W.T. Using Landsat/TM imagery to estimate nitrogen and phosphorus concentration in Taihu Lake, China. IEEE J-STARS **2012**, 5 (1), 273–280. http://dx.doi.org/10.1109/JSTARS.2011.2174339.

Colen Van, W., Mosquera, P.V., Hampel, H., Muylaert, K. Link between cattle and the trophic status of tropical high mountain lakes in p´aramo grasslands in Ecuador. Lakes Reserv.: Sci. Policy Manag. Sustain. Use **2018**, 23, 303–311. https://doi.org/ 10.1111/lre.12237.

Coppin, P.; Jonckheere, I.; Nackaerts, K.; Muys, B.; Lambin, E. Digital change detection methods in ecosystem monitoring: A review. Int. J. Remote Sens. **2004**, 25, 1565–1596.

Correll, D.L. Phosphorus: A rate limiting nutrient in surface waters. Poult. Sci. **1999**, 78, 674–682.

Deutsch, E.S.; Alameddine, I.; El-Fadel, M. Monitoring water quality in a hypereutrophic reservoir using Landsat ETM+ and OLI sensors: how transferable are the water quality algorithms? Environ Monit Assess **2018**, 190, 141. https://doi.org/10.1007/s10661-018-6506-9

Dewidar, K.; Khedr, A. Water quality assessment with simultaneous Landsat-5 TM at Manzala Lagoon, Egypt. Hydrobiologia **2001**, 457 (1–3), 49–58.

Duan, W.; He, B.; Takara, K.; Luo, P.; Nover, D.; Sahu, N.; Yamashiki, Y. Spatiotemporal evaluation of water quality incidents in Japan between 1996 and 2007. Chemosphere **2013**, 93, 946–953.

Duan, W.; Takara, K.; He, B.; Luo, P.; Nover, D.; Yamashiki, Y. Spatial and temporal trends in estimates of nutrient and suspended sediment loads in the Ishikari river, Japan, 1985 to 2010. Sci. Total Environ. **2013**, 461, 499–508.

EEA, **2018**. EuropeanWaters—Assessment of Status and Pressures. https://doi.org/10.2800/ 303664 EEA Report No 7/2018.

Eleveld MA. Wind-induced resuspension in a shallow lake from Medium Resolution Imaging Spectrometer (MERIS) full-resolution reflectances. Water Resour. Res. **2012**, 48: W04508.

Fatima R. Studies on physical, chemical and bacteriological characteristics on quality of spring water in Hajigak iron ore mine, Bamyan province, central Afghanistan, *Open J. Geol.*, **2018**, vol. 8, pp. 313–332.

Fuller, L.M.; Aichele, S.S.; Minnerick, R.J. Predicting Water Quality by Relating Secchi-Disk Transparency and Chlorophyll a Measurements to Satellite Imagery for Michigan Inland Lakes, August 2002; US Geological Survey Scientific Investigations Report; US Geological Survey: Denver, CO, USA, 2004.

Gholizadeh, M.; Melesse, A.; Reddi, L. A comprehensive review on water quality parameters estimation using remote sensing techniques. Sensors **2016**, 16, 1298.

Giardino, C.; Bresciani, M.; Cazzaniga, I.; Schenk, K.; Rieger, P.; Braga, F.; Matta, E.; Brando, V.E. Evaluation of multi-resolution satellite sensors for assessing water quality and bottom depth of lake garda. Sensors **2014**, 14, 24116–24131.

Giardino, Claudia; Bresciani, Mariano; Stroppiana, Daniela; Oggioni, Alessandro; Morabito, Giuseppe. Optical remote sensing of lakes: an overview on Lake Maggiore. *J. Limnol.*, **2014**, *73(s1): 201-214, DOI: 10.4081/jlimnol.2014.817*

Gomes, V.C.F.; Queiroz, G.R.; Ferreira, K.R. An Overview of Platforms for Big Earth Observation Data Management and Analysis. *Remote Sens.* **2020**, *12*, 1253. https://doi.org/10.3390/rs12081253

Gons, J. Herman; Auer, T. Martin. Some Notes on Water Color in Keweenaw Bay (Lake Superior), Journal of Great Lakes Research **2004**, Volume 30, Supplement 1, Pages 481-489, ISSN 0380-1330, https://doi.org/10.1016/S0380-1330(04)70408-6.

Gons, J. Herman; Auer, T. Martin; Effler, W. Steven. MERIS satellite chlorophyll mapping of oligotrophic and eutrophic waters in the Laurentian Great Lakes, Remote Sensing of Environment **2008**, Volume 112, Issue 11, Pages 4098-4106, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2007.06.029.

Guan, X.; Li, J.; Booty, W.G. Monitoring lake Simcoe water clarity using Landsat-5 TMimages. Water Resources Management **2011**, 25(8), 2015–2033.

Guang, J.; Wey, Y.; Huang, J. A model for the retrieval of suspended sediment concentrations in Taihu Lake from TM images. Journal of Geographical Sciences **2006**, 16(4), 458–464.

Gunn, J. M.; Snucins, E.; Yan, N. D.; Arts, M. T. Use of water clarity to monitor the effects of climate change and other stressors on oligotrophic lakes. Environmental Monitoring and Assessment, **2001**, 67, 69–88.

Haddad, K.D.; Harris, B.A. Use of remote sensing to assess estuarine habitats. In Proceedings of the 4th Symposium on Coastal and Ocean Management; Magoom, O.T., Converse, H., Miner, D., Clark, D., Tobin, L.T., Eds.; American

Society of Civil Engineers: New York, NY, USA, **1985**; pp. 662–675.

Hadjimitsis, D.G.; Hadjimitsis, M.G.; Clayton, C.; Clarke, B.A. Determination of turbidity in Kourris dam in Cyprus utilizing Landsat TM remotely sensed data. Water Resources Management **2006**, 20(3), 449–465.

Healey, F.P.; Hendzel, L.L. Indicators of phosphorus and nitrogen deficiency in five algae in culture. Can. J. Fish Aquat. Res. **1979**, 36, 1364–1369.

Hicks, B.J.; Stichbury, G.A.; Brabyn, L.K.; Allan, M.G.; Ashraf, S. Hindcasting water clarity from Landsat satellite images of unmonitored shallow lakes in the Waikato region, New Zealand. Environ. Monit. Assess. **2013**, 185, 7245–7261. https://doi.org/10.1016/j.rse.2019.03.018.

Huang, C.; Zou, J.; Li, Y.; Yang, H.; Shi, K.; Li, J.; Chena, X.; Zheng, F. Assessment of NIR-red algorithms for observation of chlorophyll-a in highly turbid inland waters in China. ISPRS J. Photogramm. Remote Sens. **2014**, 93, 29–39.

Huo, Shouliang; Beidou, Xi; Jing, Su; Fengyu, Zan; Qi, Chen, Danfeng, Ji; Chunzi, Ma. Determining reference conditions for TN, TP, SD and Chl-*a* in eastern plain ecoregion lakes, China. *Journal of Environmental Sciences* **2013**, 25(5) 1001–1006.

Inácio, M.; Barceló, D.; Zhao, W.; Pereira, P. Mapping lake ecosystem services: A systematic review. Sci. Total Environ. **2022**, 847, 157561

IOCCG. Remote Sensing of Ocean Colour in Coastal, and Other Optically-Complex, Waters; Reports of the International Ocean-Colour Coordinating Group, No. 3; Sathyendranath, S., Ed.; IOCCG: Dartmouth, NH, Canada, 2000.

IPCC, 2021: Summary for Policymakers. In: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Masson-Delmotte, V., P. Zhai, A. Pirani, S.L. Connors, C. Péan, S. Berger, N. Caud, Y. Chen, L. Goldfarb, M.I. Gomis, M. Huang, K. Leitzell, E. Lonnoy, J.B.R. Matthews, T.K. Maycock, T. Waterfield, O. Yelekçi, R. Yu, and B. Zhou (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 3–32, doi:10.1017/9781009157896.001.

Isenstein, Elizabeth, M.; Park, Mi-Hyun. Assessment of nutrient distributions in Lake Champlain using satellite remote sensing. Journal of Environmental Sciences **2014**, 26, 1831–1836, http://dx.doi.org/10.1016/j.jes.2014.06.019

J. Li *et al.* MODIS observations of water color of the largest 10 lakes in China between 2000 and 2012, *Int. J. Digit. Earth* **2016**, vol. 9, no. 8, pp. 788–805.

Japitana, M.; Burce, M. A Satellite-based Remote Sensing Technique for Surface Water Quality Estimation. Eng. Technol. Appl. Sci. Res. **2019**, 9, 3965–3970.

Jia, T.; Zhang, X.; Dong, R. Long-Term Spatial and Temporal Monitoring of Cyanobacteria Blooms Using MODIS on Google Earth Engine: A Case Study in Taihu Lake. Remote Sensing **2019**, 11, 2269.

Jiang, D.; Matsushita, B.; Setiawan, F.; Vundo, A. An improved algorithm for estimating the Secchi disk depth from remote sensing data based on the new underwater visibility theory. ISPRS J. Photogramm. Remote Sens. **2019**, 152, 13–23.

Karakaya, N.; Evrendilek, F.; Aslan, G.R.; Gungor, K.; Karakas, D. Monitoring of lake water quality along with trophic

gradient using landsat data. International Journal of Environmental Science and Technology **2011**, 8(4), 817–822.

Kim, S.I.; Kim, H.C.; Hyun, C.U. High resolution ocean color products estimation in Fjord of Svalbard, Arctic sea using Landsat-8 OLI. Korean J. Remote Sens. **2014**, 30, 809–816.

Kislik, Chippie; Dronova, Iryna; Grantham, E. Theodore; Kelly, Maggi. Mapping algal bloom dynamics in small reservoirs using Sentinel-2 imagery in Google Earth Engine, Ecological Indicators **2022**, Volume 140, 109041, ISSN 1470-160X, https://doi.org/10.1016/j.ecolind.2022.109041.

Knoll, L. B., Sharma, S., Denfeld, B. A., Flaim, G., Hori, Y., Magnuson, J. J., et al. Consequences of lake and river ice loss on cultural ecosystem services. Limnology and Oceanography Letters **2019**, 4(5), 119–131. https://doi.org/10.1002/lol2.10116

Kulkarni, A. Water quality retrieval from Landsat TM imagery. Procedia Computer Science **2011**, 6, 475–480.

Kumar, L.; Mutanga, O. Google Earth Engine applications since inception: Usage, trends, and potential. Remote Sens. **2018**, 10, 1509.

Kutser, T. Passive optical remote sensing of cyanobacteria and other intense phytoplankton blooms in coastal and inland waters. Int. J. Remote Sens. **2009**, 30:4401-4425.

Kutser, T. The possibility of using the Landsat image archive for monitoring long time trends in coloured dissolved organic matter concentration in lake waters. Remote Sensing of Environment **2012**, 123, 334–338.

Kutser, T.; Arst, H.; Miller, T.; Käärmann, L.; Milius, A. Telespectrometrical estimation of water transparency, chlorophyll-a and total phosphorus concentration of Lake Peipsi. Int. J. Remote Sens. **1995**, 16, 3069–3085.

Le, C.; Li, Y.; Zha, Y.; Sun, D.; Huang, C.; Lu, H. A four-band semi-analytical model for estimating chlorophyll-a in highly turbid lakes: The case of Taihu Lake, China. Remote Sens. Environ. **2009**, 113, 1175–1182.

Li Peng; Chen Shenliang; Ji Hongyu; Ke Yinghai; Fu Yutao. Combining Landsat-8 and Sentinel-2 to investigate seasonal changes of suspended particulate matter off the abandoned distributary mouths of Yellow River Delta, Marine Geology, **2021**, 441, 106622. https://doi.org/10.1016/j.margeo.2021.106622.

Lisi, P.J., Hein, C.L. Eutrophication drives divergent water clarity responses to decadal variation in lake level. Limnol. Oceanogr. **2019**, 64 (S1), S49–S59.

Lobo, F. de L.; Nagel, G.W.; Maciel, D.A.; Carvalho, L.A.S. de; Martins, V.S.; Barbosa, C.C. F.; Novo, E.M.L. de M. Algae MAp: Algae Bloom Monitoring Application for Inland Waters in Latin America. Remote Sensing **2021**, 13, 2874.

Loveland, T.; Dwyer, J. Landsat: building a strong future. Remote Sensing of Environment **2012**,122, 22–29.

MacKay, H.; Finlayson, C.M.; Fernández-Prieto, D.; Davidson, N.; Pritchard, D.; Rebelo, L.M. The role of earth observation (EO) technologies in supporting implementation of the Ramsar convention on wetlands. J. Environ. Manag. **2009**, 90, 2234–2242.

Maeda, E.E.; Lisboa, F.; Kaikkonen, L.; Kallio, K.; Koponen, S.; Brotas, V.; Kuikka, S. Temporal patterns of phytoplankton phenology across high latitude lakes unveiled by long-term time series of

satellite data. Remote Sens. Environ. **2019**, 221, 609–620.

Malve, O.; Qian, S.S. Estimating nutrients and chlorophyll a relationships in Finnish lakes. Environ. Sci. Technol. **2006**, 40, 7848–7853.

Mammides, C.A global assessment of the human pressure on the world's lakes. Global Environ. Change **2020**, 63, 102084. https://doi.org/10.1016/j.gloenvcha.2020.102084.

Mandanici, E.; Bitelli, G. Multi-image and multi-sensor change detection for long-term monitoring of arid environments with Landsat series. Remote Sens. **2015**, 7, 14019–14038.

Mandanici, E.; Bitelli, G. Preliminary Comparison of Sentinel-2 and Landsat 8 Imagery for a Combined Use. *Remote Sens.* **2016**, *8*, 1014. https://doi.org/10.3390/rs8121014

Mantas, V.M.; Pereira, A.J.S.C.; Neto, J.; Patrício, J.; Marques, J.C. Monitoring estuarine water quality using satellite imagery. The Mondego river estuary (Portugal) as a case study. Ocean & Coastal Management **2013**, 72: 13-21

Markogianni, V.; Dimitriou, E.; Karaouzas, I. Water quality monitoring and assessment of an urban Mediterranean lake facilitated by remote sensing applications. Environ. Monit. Assess. **2014**, 186, 5009–5026.

Markogianni, V.; Kalivas, D.; Petropoulos, G.; Dimitriou, E. An appraisal of the potential of Landsat 8 in estimating chlorophyll-a, ammonium concentrations and other water quality indicators. Remote Sens. **2018**, 10, 1018.

Markogianni, V.; Kalivas, D.; Petropoulos, G.P.; Dimitriou, E. Estimating Chlorophyll-a of Inland Water Bodies in Greece Based on Landsat Data. *Remote Sens.* **2020**, *12*, 2087. https://doi.org/10.3390/rs12132087.

Markogianni, V.; Kalivas, D.; Petropoulos, G.P.; Dimitriou, E. Modelling of Greek Lakes Water Quality Using Earth Observation in the Framework of the Water Framework Directive (WFD). Remote Sens. **2022**, 14, 739. https://doi.org/10.3390/rs14030739.

Matthews, M.W. A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters. International Journal of Remote Sensing, **2011**, 32(21), 6855–6899.

Mavromati, E.; Kagalou, I.; Kemitzoglou, D.; Apostolakis, A.; Tsiaoussi, D. Linkages between physicochemical status and hydromorphology in greek lakes under WFD policy. European Water Resources Association. Special Issue: 10th Word Congress on Water Resources and Environment (EWRA2017—Issue II). Eur. Water **2017**, 58, 273–279.

McCullough, I.M.; Loftin, C.S.; Sader, S.A. Combining lake and watershed characteristicswith Landsat TMdata for remote estimation of regional lake clarity. Remote Sensing of Environment **2012a**, 123, 109–115.

McCullough, Ian M. Remote Estimation of Regional Lake Clarity with Landsat TM and MODIS Satellite Imagery. *Electronic Theses and Dissertations*. **2012**, 1744.

Meyer, M.F., Labou, S.G., Cramer, A.N. et al. The global lake area, climate, and population dataset. Sci Data **2020**, 7, 174. https://doi.org/10.1038/s41597-020-0517-4

Michalak, A. Study role of climate change in extreme threats to water quality. Nature **2016**, 535, 349–350.

Morel, A.; Gordon, H.R. Report of the working group on water colour. Bound. -Layer Meteorol. **1980**, 18, 343–355.

Nelson, S.A.C.; Soranno, P.A.; Cheruvelil, K.S.; Batzli, S.A.; Skole, D.L. Regional Assessment of lake water clarity using satellite remote sensing. J. Limnol. **2003**, 62, 27–32.

Nikolaidis, N. P., Phillips, G., Poikane, S., Várbíro, G., Bouraoui, F., Malago, A., & Lilli, M.A. River and lake nutrient targets thatsupport ecological status: European scale gap analysis and strategiesfor the implementation of the Water Framework Directive.Science ofthe Total Environment **2022**, 813(March), 151898, https://doi.org/10.1016/j.scitotenv.2021.151898

Nikolaou, A.D.; Meric, S.; Lekkas, D.F.; Naddeo, V.; Belgiorno, V.; Groudev, S.; Tanik, A. Multi-parametric water quality monitoring approach according to the WFD application in Evros trans-boundary river basin: Priority pollutants. Desalination **2008**, 226, 3–320.

Odermatt D, Danne O, Philipson P and Brockmann C. Diversity II water quality parameters from ENVISAT (2002-2012): a new global information source for lakes. Earth System Science Data **2018**, 10(3).

Odermatt, D.; Gitelson, A.; Vittorio, E.V.; Schaepman, M. Review of constituent retrieval in optically deep and complex waters from satellite imagery. Remote Sens. Environ. **2012**, 118, 116–126.

Olmanson, L. G.; Brezonik, P. L.; Finlay, J. C.; Bauer, M. E. Comparison of Landsat 8 and Landsat 7 for regional measurements of CDOM and water clarity in lakes. Remote Sensing of Environment **2016**, 185, 119–128. https://doi.org/10.1016/j.rse.2016.01.007

Olmanson, L.G., Page, B.P., Finlay, J.C., Brezonik, P.L., Bauer, M.E., Griffin, C.G., and Hozalski, R.M. Regional measurements and spatial/temporal analysis of CDOM in 10,000+ optically variable Minnesota lakes using Landsat 8 imagery. Sci. Total Environ. **2020**, 724: 138141. doi:10.1016/j.scitotenv.2020.138141.
Olmanson, L.G.; Bauer, M.E.; Brezonik, P.L. A 20-year Landsat water clarity census of Minnesota's 10,000 lakes. Remote Sens. Environ. **2008**, 112 4086-4097.

Onderka, M.; Pekárová, P. Retrieval of suspended particulate matter concentrations in the Danube River from Landsat ETM data. Science of the Total Environment **2008**, 397(1), 238–243.

O'Reilly, J.E.; Werdell, P.J. Chlorophyll algorithms for ocean color sensors-OC4, OC5 & OC6. Remote Sens. Environ. **2019**, 229, 32–47.

Pahlevan, N., Smith, B., Schalles, J., Binding, C., Cao, Z., Ma, R., Alikas, K., Kangro, K., Gurlin, D., Hà, N., Matsushita, B., Moses, W., Greb, S., Lehmann, M.K., Ondrusek, M., Oppelt, N., Stumpf, R. Seamless retrievals of chlorophyll-a from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in inland and coastal waters: a machine-learning approach. Remote Sens. Environ. **2020**, 24, 111604. https://doi.org/10.1016/j.rse.2019.111604.

Paulsson, O., Widerlund, A. Modelled impact of climate change scenarios on hydrodynamics and water quality of the R ¨avlidmyran pit lake, northern Sweden. Appl. Geochem. **2022**, 139, 105235 https://doi.org/10.1016/j.apgeochem.2022.105235.

Pedreros-Guarda, M.; Abarca-del-Río, R.; Escalona, K.; García, I.; Parra, Ó. A Google Earth Engine Application to Retrieve Long-Term Surface Temperature for Small Lakes. Case: San Pedro Lagoons, Chile. Remote Sens. **2021**, 13, 4544. https://doi.org/10.3390/ rs13224544

Pepe, M.; Giardino, C.; Borsani, G.; Cardoso, A.C.; Chiaudani, G.; Premazzi, G.; Rodari, E.; Zilioli, E. Relationship between apparent optical properties and

photosynthetic pigments in the sub-alpine Lake Iseo. Sci. Total Environ. **2001**, 268, 31–45.

Pérez-Fuentetaja, A.; Dillon, P.J.; Yan, N.D. et al. Significance of dissolved organic carbon in the prediction of thermocline depth in small Canadian shield lakes. Aquatic Ecology **1999**, 33, 127–133. https://doi.org/10.1023/A:1009998118504

Pizani, F.; Maillard, P. The determination of water quality parameters by remote sensing technologies: 2000 -2020. Universidade Federal de Minas Gerais, [S. l.] **2022** p. 1–30. DOI: 10.13140/RG.2.2.32203.87849

Poikane, S.; Alves, M. H.; Argillier, C.; Berg, M.; Buzzi, F.; Hoehn, E. et al. Defining chlorophyll-*a* reference conditions in European lakes. *Environmental Management* **2010**, 45(6): 1286– 1298.

Porst, G., Brauns, M., Irvine, K., Solimini, A., Sandin, L., Pusch, M., Miler, O. Effects of shoreline alteration and habitat heterogeneity on macroinvertebrate community composition across European lakes. Ecol. Indicat. **2019**, 98, 285–296. https://doi.org/10.1016/j.ecolind.2018.10.062.

Pozdnyakov, D.; Shuchman, R.; Korosov, A.; Hatt, C. Operational algorithm for the retrieval of water quality in the Great Lakes. Remote Sens. Environ. **2005**, 97, 352–370.

Randolph, K.; Wilson, J.; Tedesco, L.; Li, L.; Pascual, D.L.; Soyeux, E. Hyperspectral remote sensing of cyanobacteria in turbid productive water using optically active pigments, chlorophyll-a and phycocyanin. Remote Sens. Environ. **2008**, 112, 4009–4019.

Ruiz-Verdu, A; Simis, SGH; de Hoyos, C.; Gons, H.J.; Peɑa- Martvnez, R. An evaluation of algorithms for the remote sensing of cyanobacterial biomass. Remote Sens. Environ. **2008**, 112:3996-4008.

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

Sagan, V.; Peterson, T.K.; Maimaitijiang, M.; Sidike, P.; Sloan, J.; Greeling, A.B.; Samar, M.; Adams, C. Monitoring inland water quality using remote sensing: Potential and limitations of spectral indices, bio-optical simulations, machine learning, and cloud computing. Earth-Sci. Rev. **2020**, 205, 103187.

Salmaso, N., Tolotti, M. Phytoplankton and anthropogenic changes in pelagic environments. Hydrobiologia **2021**, 848, 251–284. https://doi.org/10.1007/s10750-020-04323-w

Schewe, J.; Heinke, J.; Gerten, D.; Haddeland, I.; Arnell, N.; Clark, D.; Dankers, D.; Eisner, S.; Fekete, B.M.; Colón-González, F.J.; et al. Multimodel assessment of water scarcity under climate change. Proc. Natl. Acad. Sci. USA **2013**, 111, 3245–3250.

Schindler, D.W.; Curtis, P.J.; Bayley, S.E. et al. Climate-induced changes in the dissolved organic carbon budgets of boreal lakes. Biogeochemistry, **1997,** 36, 9–28. https://doi.org/10.1023/A:1005792014547

Schirpke Uta, Ebner Manuel. Exposure to global change pressures and potential impacts on ecosystem services of mountain lakes in the European Alps. Journal of Environmental Management **2022**, 318, 115606. https://doi.org/10.1016/j.jenvman.2022.115606.

Senetra, A., Dynowski, P., Cie´slak, I., Zrobek-Sokolnik, A. An evaluation of the impact of hiking tourism on the ecological status of alpine lakes-a case study of the valley of dolina pieciu stawow polskich in the tatra mountains. Sustainability **2020**, 12, 2963. https://doi.org/10.3390/su12072963.

Smith, V.H. The nitrogen and phosphorus dependence of algal biomass in lakes: An empirical and theoretical analysis. Limnol. Oceanogr. **1982**, 27, 1101–1112.

Solheim Lyche, A., Globevnik, L., Austnes, K., Kristensen, P., Moe, S.J., Persson, J., Phillips, G., Poikane, S., van de Bund, W., Birk, S. A new broad typology for rivers and lakes in Europe: development and application for large-scale environmental assessments. Sci. Total Environ. **2019**, 697, 134043 https://doi.org/ 10.1016/j.scitotenv.2019.134043.

Somasundaram, D.; Zhang, F.; Ediriweera, S.; Wang, S.; Yin, Z.; Li, J.; Zhang, B. Patterns, Trends and Drivers of Water Transparency in Sri Lanka Using Landsat 8 Observations and Google Earth Engine. Remote Sens. **2021**, 13, 2193. https://doi.org/10.3390/rs13112193

Song Kaishan; Wang Qiang; Liu Ge; Jacinthe Pierre-Andre; Li Sijia; Tao Hui; Du Yunxia; Wen Zhidan; Wang Xiang; Guo Wenwen; Wang Zongming; Shi Kun; Du Jia; Shang Yingxin; Lyu Lili; Hou Junbin; Zhang Baohua; Cheng Shuai; Lyu Yunfeng; Fei Long. A unified model for high resolution mapping of global lake (>1 ha) clarity using Landsat imagery data, Science of The Total Environment, **2022**, 810, 151188. https://doi.org/10.1016/j.scitotenv.2021.151188.

Song, K.; Li, L.; Li, S.; Tedesco, L.; Hall, B.; Li, L. Hyperspectral Remote Sensing of Total Phosphorus (TP) in Three Central Indiana Water Supply Reservoirs. Water Air Soil Pollut. **2012**, 223, 1481–1502.

Song, K.S.; Li, L.L.; Tedesco, S.; Li, H.; Duan, T.; Liu, D.W.; Hall, B.; Du, J.; Li, Z.C.; Shi, K.; et al. Remote estimation of chlorophyll-a in turbid inland waters: Three-band model versus GA-PLS model. Remote Sens. Environ. **2013**, 136, 342–357.

Spyrakos, E.; O'donnell, R.; Hunter, P.D.; Miller, C.; Scott, M.; Simis, S.G.; Neil, C.; Barbosa, C.C.; Binding, C.E.; Bradt, S.; et al. Optical types of inland and coastal waters. Limnol. Oceanogr. **2018**, 63, 846–870.

Stadelmann, T.H.; Brezonik, P.L.; Kloiber, S.M. Seasonal patterns of chlorophyll-a and Secchi disk transparency in lakes of east-central Minnesota: implications for design of ground and satellite-based monitoring programs. Lake Reserv. Manage. **2001**,17:299-314.

Strobel, C.J.; Paul, J.; Hughes, M.; Buffum, H.W.; Brown, B.S.; Summers, K. Using information on spatial variability of small estuaries in designing large-scale estuarine monitoring programs. Environmental Monitoring and Assessment **2000**, 63, 223-236.

Sun, D.; Hu, C.; Qiu, Z.; Cannizzaro, J.P.; Barnes, B.B. Influence of a red band-based water classification approach on chlorophyll algorithms for optically complex estuaries. Remote Sens. Environ. **2014**, 155, 289–302.

Tebbs, E.J.; Remedios, J.J.; Harper, D.M. Remote sensing of chlorophyll-a as a measure of cyanobacterial biomass in Lake Bogoria, a hypertrophic, saline–alkaline, Flamingo Lake, using Landsat ETM+. Remote Sensing of Environment **2013**, 135, 92–106.

Teubner, K.; Teubner, I.; Pall, K.; Kabas, W.; Tolotti, M.; Ofenbock, T.; Dokulil, M. T. New Emphasis on Water Transparency as Socio-Ecological Indicator for Urban Water: Bridging Ecosystem Service Supply and Sustainable Ecosystem Health.

Frontiers in Environmental Science **2020**, 8, 8:573724.

Tiberti, R., Buscaglia, F., Armodi, M., Callieri, C., Ribelli, F., Rogora, M., Tartari, G., Bocca, M. Mountain lakes of Mont Avic Natural Park: ecological features and conservation issues: mountain lakes of a natural park. J. Limnol. **2019**, 79 https://doi.org/ 10.4081/jlimnol.2019.1923.

Topp, S.; Pavelsky, T.; Jensen, D.; Simard, M.; Ross, M. Research trends in the use of remote sensing for inland water quality science: Moving towards multidisciplinary applications. Water **2020**, 12, 169.

Torbick, N.; Feng, H.; Zhang, J.; Qi, J.; Zhang, H.; Becker, B. Mapping chlorophyll-a concentrations in West Lake, China using Landsat 7 ETM+. Journal of Great Lakes Research **2008**, 34(3), 559–565.

Tyler, A.N.; Svab, E.; Preston, T.; Presing, M.; Kovacs, W.A. Remote sensing of the water quality of shallow lakes: A mixture modeling approach to quantifying phytoplankton in water characterized by high suspended sediment. Int. J. Remote Sens. **2006**, 27, 1521–1537.

United Nations, **2015**. The 2030 Agenda for Sustainable Development - Sustainable Development Goals.

Vaičiūtė, D.; Bučas, M.; Bresciani, M.; Dabulevičienė, T.; Gintauskas, J.; Mėžinė, J.; Tiškus, E.; Umgiesser, G.; Morkūnas, J.; De Santi, F.; Bartoli, M. Hot moments and hotspots of cyanobacteria hyperblooms in the Curonian Lagoon (SE Baltic Sea) revealed via remote sensing-based retrospective analysis. Sci. Total Environ. **2021**, 769, 145053.

Van der Werff, H.; van der Meer, F. Sentinel-2A MSI and Landsat 8 OLI provide data continuity for geological remote sensing. Remote Sens. **2016**, 8, doi:10.3390/rs8110883.

Vincent, R.K.; Qin, X.; McKay, R.M.; Miner, J.; Czajkowski, K.; Savino, J. et al. Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie. Remote Sensing of Environment **2004**, 89(3), 381–392.

Vörösmarty, C.J.; Green, P.; Salisbury, J.; Lammers, R.B. Global water resources: Vulnerability from climate change and population growth. Science **2000**, 289, 284–288.

Wang, L.; Xu, M.; Liu, Y.; Liu, H.; Beck, R.; Reif, M.; Emery, E.; Young, J.; Wu, Q. Mapping Freshwater Chlorophyll-a Concentrations at a Regional Scale Integrating Multi-Sensor Satellite Observations with Google Earth Engine. *Remote Sens.* **2020**, *12*, 3278. https://doi.org/10.3390/rs12203278

Wang, Y.P.; Xia, H.; Fu, J.; Sheng, G.Y. Water quality change in reservoirs of Shenzhen, China: Detection using LANDSAT/TM data. Sci. Total Environ. **2004**, 328, 195–206.

Warren, M.A.; Simis, S.G.H.; Martinez-Vicente, V.; Poser, K.; Bresciani, M.; Alikas, K.; Spyrakos, E.; Giardino, C.; Ansper, A. Assessment of atmospheric correction algorithms for the Sentinel-2A MultiSpectral Imager over coastal and inland waters, Remote Sensing of Environment, **2019**, 225, 267-289.

Watanabe, Fernanda S.Y.; Miyoshi, Gabriela T.; Rodrigues, Thanan W.P.; Bernardo, Nariane M.R.; Rotta, Luiz H.S.; Alcântara, Enner; Imai, Nilton N. Inland water's trophic status classification based on machine learning and remote sensing data, Remote Sensing Applications: Society and Environment **2020**, Volume 19, 100326, ISSN 2352-9385, https://doi.org/10.1016/j.rsase.2020.100326.

Weber, S.J.; Mishra, D.R.; Wilde, S.B.; Kramer, E. Risks for cyanobacterial harmful algal blooms due to land

management and climate interactions. Sci. Total Environ. **2020**, 703, 134608.

Wen Zhidan; Qiang, Wang; Ge, Liu; Pierre-Andre, Jacinthe; Xiang, Wang; Lili, Lyu; Hui, Tao; Yue, Ma; Hongtao, Duan; Yingxin, Shang; Baohua, Zhang; Yunxia, Du; Jia, Du; Sijia, Li; Shuai, Cheng; Kaishan, Song. Remote sensing of total suspended matter concentration in lakes across China using Landsat images and Google Earth Engine, ISPRS Journal of Photogrammetry and Remote Sensing **2022**, Volume 187, Pages 61-78, ISSN 0924-2716, https://doi.org/10.1016/j.isprsjprs.2022.02.018.

Weyhenmeyer, G.A., Hartmann, J., Hessen, D.O., Kopáčˇek, J., Hejzlar, J., Jacquet, S., et al. Widespread diminishing anthropogenic effects on calcium infreshwaters. Sci. Rep. **2019**, 9, 1–10.

Whyte, A.; Feredinos, K.; Petropoulos, G.P. A new synergistic approach for monitoring wetlands using Sentinels -1 and 2 data with object-based Machine Learning algorithms. Environ. Model Softw. **2018**, 104, 40–57.

Williamson, C. E.; Morris, D. P.; Pace, M. L.; Olson, O. G. Dissolved organic carbon and nutrients as regulators of lake ecosystems: resurrection of a more integrated paradigm. Limnology and Oceanography, **1999**, 44: 795–803.

Wu, C.F.; Wu, J.P.; Qi, J.G.; Zhang, L.S.; Huang, H.Q.; Lou, L.P.; et al. Empirical estimation of total phosphorus concentration in the mainstream of the Qiantang River in China using Landsat TM data. Int. J. Remote Sens. **2010**, 31 (9), 2309–2324.

Wu, G.; de Leeuw, J.; Skidmore, A.K.; Prins, H.H.T.; Liu, Y. Comparison of MODIS and Landsat TM5 images for mapping tempo-spatial dynamics of Secchi disk depths in Poyang Lake National Nature Reserve, China. Int. J. Remote Sens. **2008**, 29, 2183–2198.

Yacobi, Y.Z.; Gitelson, A.; Mayo, M. Remote sensing of chlorophyll in Lake Kinneret using highspectral-resolution radiometer and Landsat TM: spectral features of reflectance and algorithm development. Journal of Plankton Research **1995**, 17(11), 2155–2173.

Yang, Z.; Reiter, M.; Munyei, N. Estimation of chlorophyll-a concentrations in diverse water bodies using ratio-based NIR/Red indices. Remote Sens. Appl. Soc. Environ. **2017**, 6, 52–58.

Zhao, D.; Cai, Y.; Jiang, H.; Xu, D.; Zhang, W.; An, S. Estimation of water clarity in Taihu Lake and surrounding rivers using Landsat imagery. Advances in Water Resources **2011**, 34(2), 165–173.

Zhao, Q.; Yu, L.; Du, Z.; Peng, D.; Hao, P.; Zhang, Y.; Gong, P. An Overview of the Applications of Earth Observation Satellite Data: Impacts and Future Trends. Remote Sens. **2022**, 14, 1863. https://doi.org/10.3390/rs14081863

Zhou, W.; Wang, S.; Zhou, Y.; Troy, A. Mapping the concentrations of total suspended matter in lake Taihu, China, using Landsat-5 TM data. Int. J. Remote Sens. **2006**, 27:1177-1191.

Zhu, W.; Yu, Q.; Tian, Y. Q.; Becker, B.L.; Zheng, T.; Carrick, H. J. An assessment of remote sensing algorithms for colored dissolved organic matter in complex freshwater environments. Remote Sensing of Environment **2014**, 140, 766–778.

Zong, J.-M.; Wang, X.-X.; Zhong, Q.-Y.; Xiao, X.-M.; Ma, J.; Zhao, B. Increasing Outbreak of Cyanobacterial Blooms in Large Lakes and Reservoirs under Pressures from Climate Change and Anthropogenic Interferences in the Middle-Lower Yangtze River Basin. Remote Sensing **2019**, 11, 1754.

2. Towards the Modelling of Greek Lakes Water Quality Using Satellite Remote Sensing technology

2.1 Estimating Chlorophyll-*a* of Inland Water Bodies in Greece Based on Landsat Data

Published as: **Markogianni, V.**; Kalivas, D.; Petropoulos, G.P.; Dimitriou, E. Estimating Chlorophyll-*a* of Inland Water Bodies in Greece Based on Landsat Data. *Remote Sens.* **2020**, *12*, 2087. https://doi.org/10.3390/rs12132087

Preamble

Assessing chlorophyll-*a* (Chl-*a*) pigments in complex inland water systems is of key importance as this parameter constitutes a major ecosystem integrity indicator. In this study, a methodological framework is proposed for quantifying Chl-*a* pigments using Earth observation (EO) data from Landsat 7 Enhanced Thematic Mapper Plus (ETM+) and 8 Operational Land Imager (OLI) sensors. This effort aimed at exploring different remote sensing-derived Chl-*a* algorithms for various types of lakes. The practical use of the proposed approach was evaluated in a total of 50 lake water bodies (natural and artificial) from 2013–2018, constituting the National Lake Network Monitoring of Greece in the context of the Water Framework Directive (WFD). All in all, the results evidenced the suitability of Landsat data when used with the proposed technique to estimate log-transformed Chl-*a*. The methodological framework proposed herein can be used as a useful resource toward a continuous monitoring and assessment of lake water quality, supporting sustainable water resources management.

2.1.1 Introduction

Accumulating passive exploitation of natural resources, improper land-use practices and irregular development activities in lake basins undermine various significant functions of water resources (Alymkulova et al., 2016). Surface water provides exceptional financial benefits, regarding water supply (quantity and quality), fisheries, agriculture, wildlife resources and recreation

and tourism opportunities (Ramsar Information Paper no. 1 2007; Ramsar Convention Bureau). The need for sustainable management of water bodies highlights the fact that water resources are not inexhaustible and have limited resistance under anthropogenic pressures such as ongoing drainage, conversion and pollution.

Hence, one of the most significant aspects for the sustainable management of water bodies - lakes in particular - is the constant monitoring of their quality as well as of their watersheds (Gholizadeh et al., 2016). Water quality parameters comprising physical, chemical, and biological properties are conventionally measured by collecting samples from the field and then analysing those samples in the laboratory. Although *in-situ* monitoring provides high accuracy, it is a time-consuming procedure, and cannot ensure a simultaneous water quality dataset on a regional or greater scale [Gholizadeh et al., 2016; Duan et al., 2013; Duan et al., 2013b). Furthermore, traditional point sampling methods are not capable of detecting the spatial or temporal variations in water quality, as required in extensive assessment and management of water bodies. On the other hand, geoinformation technologies provide a promising direction in that respect. In particular, the combined use of Earth Observation (EO) and Geographical Information Systems (GIS) allows monitoring in an efficient and robust way lake parameters over variable spatial scales, including lakes that are otherwise inaccessible (MacKay et al., 2009; Whyte et al., 2018).

Various EO instruments mounted on either airborne or satellite platforms, acquire spectral information and measure the energy from the water's surface at different wavelengths (Gholizadeh et al., 2016). The most commonly used approach to inland water remote sensing involves fitting a standard linear regression between spectral band/band ratio values and temporally coincident *in-situ* water quality measurements (Topp et al., 2020). Visible, near and short infrared bands of the solar spectrum have been usually used by many researchers to acquire powerful correlations -through empirical approaches- among water column reflectance values and constituents, in different water bodies (Ritchie et al., 2003; Gitelson et al., 2008; Olmanson et al., 2008; Gholizadeh et al., 2016; El-Din et al., 2013; Giardino et al., 2014; Markogianni et al., 2018). EO data from several satellite and airborne sensors such as SPOT, MODIS, Earth Observing 1-Hyperion and MERIS have been used for Chl-*a* estimation (Nas et al., 2007; Kim et al., 2014; Zhang et al., 2015; Lim et al., 2015; Bonansea et al., 2018). Nonetheless, it revealed that the Landsat seems to be more appropriate and widely used for Chl-*a* assessment due to its temporal coverage, spatial resolution and easy accessibility (Gholizadeh et al., 2016).

Chl-*a* is the major indicator of trophic state and considered as one of the top water pollution indices related to public health, eutrophication and deterioration of ecosystem habitat. In Case 2 waters (i.e. inland and coastal

waters), the optical properties are measured based on a compound of dissolved organic matter, dead organic and inorganic particulate matter and phytoplankton (Chl-*a*). Therefore, determination of Chl-*a* concentration is much more complex and less accurate (Ritchie et al., 1990; Dekker and Peters, 1993; George, 1997; Gholizadeh et al., 2016). Oligotrophic to mesotrophic waterbodies with low biomass, present a Chl-*a* spectrum characterized by a sun-induced fluorescence peak centered at 680 nm (Gitelson et al., 1994; Topp et al., 2020), while eutrophic waterbodies (high biomass) present a florescence signal which is masked by absorption features and backscatter peaks around 665 nm and 710 nm, respectively (Matthews et al., 2012; Topp et al., 2020). The ratio between these two wavelengths has been widely used to quantify Chl-*a* concentrations with high accuracy (Le et al., 2011; Topp et al., 2020).

Many studies have focused on monitoring eutrophication or trophic state through Chl-*a* concentration retrieval in Greek lakes. For example, Peppa et al. (2020) applied Chl-*a* detection algorithms in Lake Pamvotis using Sentinel-2 Data, (Markogianni et al., 2014) developed empirical Chl-*a* quantitative models based on Landsat 5 images in the brackish urban shallow Koumoundourou lake while Markogianni et al. (2018) investigated the suitability of the OLI instrument on-board the Landsat 8 satellite platform in accurately estimating Chl-*a* in the largest freshwater body of Greece (Trichonis Lake). In another study, Kontopoulou et al. (2017), tried to exploit the Water Framework Directive (WFD) dataset and measure Chl-*a* concentrations by using Landsat 8 data in 11 (6 natural; 5 artificial) of the 50 lakes comprising the national sampling network.

In this study, authors discuss the utility of remotely sensed techniques in the qualitative assessment of 50 lake water bodies and particularly of Chl-*a* concentrations, derived from the WFD (2000/60/EC) monitoring network for lakes in Greece. WFD provides a scheme for the conservation and improvement of inland, ground and coastal waters' ecological status and aims to harmonize European legislation on water. Thus, pan-European hydromorphological, physicochemical and biological datasets are used to determine ecological status of surface waters (Article 8) in order to assure and further improve future water quality and quantity (Mavromati et al., 2017).

In purview of the above, this study proposes a methodological framework that aims to provide Chl-*a* in Case 2 complex inland waters of Greece by generating accurate quantitative models with EO data from the Landsat 7 ETM+ and 8 OLI satellite series. The methodology applied initially includes the implementation of stepwise MLR analysis of the whole available Chl-*a* dataset with the basic aim of exploring its potential to establish robust Chl-*a* quantitative algorithms, regardless of lake characteristics. Then, PCA is performed to highlight which are the most significant parameters (artificial/natural, WFD typology, water sampling's season and climatic type)

affecting Chl-*a* values. This procedure was considered to be proven valuable for the next step, involving the execution of multiple stepwise MLR analyses -based on PCA results- among different groups of cases. This effort aimed at exploring different remote sensing derived Chl-*a* algorithms for various types of lakes according to the most significant lake characteristics. The practical use of the proposed approach is evaluated in a total of 50 lake water bodies (natural and artificial) during 2013-2018, consisting the National Lake Network Monitoring of Greece in the context of Water Framework Directive (WFD).

2.1.2 Materials and Methods

2.1.2.1 Study area

The National Monitoring Network of Waters in Greek lakes, according to the Joint Ministerial Decision 140384/2011, is implemented by the Goulandris Natural History Museum, Greek Biotope/Wetland Centre (EKBY). The network consists of 50 lake water bodies, natural and artificial.

At the majority of the lakes only one sampling station is detected, except for trans-boundary lakes (Megali Prespa, Mikri Prespa and Doirani), where two sampling stations are located (Table 2.1.2-1; Figure 2.1.2-1). From the total of 53 sampling sites, the 27 are surveillance and the 26 operational ones.

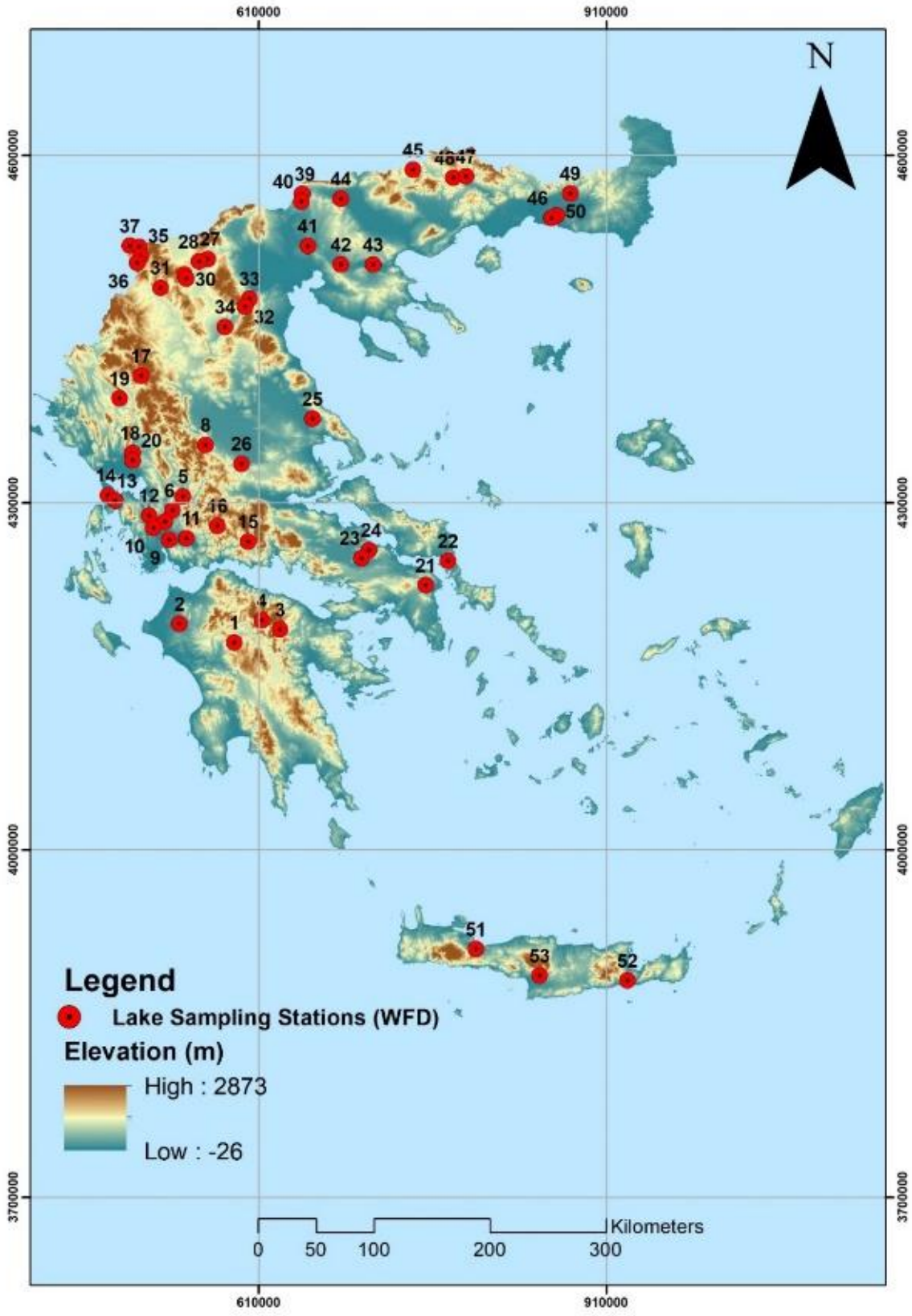


Figure 2.1.2-1 Monitoring network for lakes in Greece (GGRS_1987 coordinate system; Transverse Mercator projection; numbers of sampling stations coincide with the numbers presented in Table 2.1.2-1).

Table 2.1.2-1. General characteristics of the lakes comprising the National Lake Network Monitoring in Greece (WFD; Mavromati et al., 2018).

No	National Name Station	(N)atural/ (A)rtificial	Typology	Koppen climate classification	Mean depth (m)	No	National Name Station	(N)atural/ (A)rtificial	Typology	Koppen climate classification	Mean depth (m)
1	Lake Ladona	A	L-M8	Csa	-	28	Lake Petron	N	GR-VSNL	Cfa	3.1
2	Lake Pineiou	A	L-M8	Csa	15.1	29	Lake Zazari	N	GR-SNL	Cfa	3.95
3	Lake Stymfalia	N	GR-VSNL	Csa	1.3[1]	30	Lake Cheimaditida	N	GR-VSNL	Cfa	1.01
4	Lake Feneou	A	L-M5/7W	Csa	10.5	31	Lake Kastorias	N	GR-SNL	Cfa	3.7
5	Lake Kremaston	A	L-M8	Csa	47.2	32	Lake Sfikias	A	L-M5/7W	BSk	23.2
6	Lake Kastrakiou	A	L-M5/7W	Csa	33.2	33	Lake Asomaton	A	L-M5/7W	BSk	20.8
7	Lake Stratou	A	GR-SR	Csa	9.6	34	Lake Polyfytou	A	L-M5/7W	Cfa	22.4
8	Lake Tavropou	A	L-M5/7W	Csa	15.0	35	Lake Mikri Prespa A	N	GR-SNL	Csa	3.95
9	Lake Lysimacheia	N	GR-SNL	Csa	3.5	36	Lake Mikri Prespa B	N	GR-DNL	Csa	-
10	Lake Ozeros	N	GR-SNL	Csa	3.8	37	Lake Megali Prespa A	N	GR-DNL	Csa	17
11	Lake Trichonida	N	GR-DNL	Csa	29.6	38	Lake Megali Prespa B	N	GR-DNL	Csa	-
12	Lake Amvrakia	N	GR-DNL	Csa	23.4	39	Lake Doirani 1	N	GR-SNL	Dfc	4.6
13	Lake Voulkaria	N	GR-VSNL	Csa	0.96	40	Lake Doirani 2	N	GR-SNL	Dfc	-
14	Lake	N	GR-	Csa	-	41	Lake	N	GR-	Cfb	1.2

	Saltini		SP1 (lagoon)				PikroLake		SP2 (special category)		
15	Lake Mornou	A	L-M5/7W	Csa	38.5	42	Lake Koroneia	N	GR-VSNL	Csa	3.8
16	Lake Evinou	A	L-M5/7W	Csa	31.5	43	Lake Volvi	N	GR-DNL	Csa	12.3
17	Lake Pigon Aoou	A	L-M5/7W	Csa	20.8	44	Lake Kerkini	A	GR-SR	Dfc	2.19
18	Lake Pournariou	A	L-M5/7W	Csa	29.8	45	Lake Leukogeion	A	GR-SR	Dfc	4.05
19	Lake Pamvotida	N	GR-SNL	Csa	5.3	46	Lake Ismarida	N	GR-VSNL	Csb	0.9
20	Lake Pournariou II	A	GR-SR	Csa	11.7	47	Lake Platanovrysis	A	L-M5/7W	Dfc	26.4
21	Lake Marathona	A	L-M8	Csa	15.8	48	Lake Thisavrou	A	L-M5/7W	Dfc	38.4
22	Lake Dystos	N	GR-VSNL	Csa	-	49	Lake Gratinis	A	L-M5/7W	Csb	14.2
23	Lake Yliki	N	GR-DNL	Csa	20.1	50	Lake N. Adrianis	A	GR-SR	Csb	-
24	Lake ParaLake	N	GR-SNL	Csa	2.99	51	Lake Kourna	N	GR-DNL	Csa	15
25	Lake Karlas	A	GR-SR	BSk	0.9	52	Lake Bramianou	A	L-M8	Csa	10.1
26	Lake Smokovou	A	L-M8	Csa	-	53	Lake Faneromenis	A	L-M8	Csa	9.98
27	Lake Vegoritida	N	GR-DNL	Dfa	26.52	[1](Zacharias et al., 2002)					

2.1.2.2 Data acquisition

2.1.2.2.1 In-situ data

In this study Chl-*a* concentrations, measured from 2013 up to 2018 (summer, autumn and spring) throughout the studied lake stations were acquired while the Chl-*a* concentrations (in μg/l) were determined spectrophotometrically (Method 10200 H; APHA, 1989). Those data were retrieved from the EKBY's site (Goulandris Natural History Museum, Greek Biotope/Wetland Centre; http://biodiversity-info.gr/index.php/el/lakes-data#!IMGP4731), where more details about the sampling periods, stations and the variables measured can be found. Apart from the Chl-*a* measurements, some basic characteristics of the 50 studied lakes have been also considered, such as whether they are natural or artificial, their typology according to the Water Framework Directive (Table 2.1.2-2) and the climatic type according to the Köppen-Geiger climate classification (Table 2.1.2-3; Peel et al., 2007). The determination of the Mediterranean lake types is based on the 2013/480/EU decision (Table 2.1.2-2) while the WFD typology of each lake has been retrieved from the respective reports, acquired from the Environment and Energy Ministry's website (http://wfdver.ypeka.gr/).

Table 2.1.2-2. WFD national lake types.

Type	Characteristics	Elevation (m)	Precipitation (mm) and Temperature (°C) (mean annual values)	Surface (km²)	Mean depth (m)	Catchment (km²)	Thermal Stratification
Artificial							
L-M5/7W	Deep, large reservoirs, silicate, wet areas	<1000	>800 or/and <15	>0.5	>15	<20000	
L-M8	Deep, large reservoirs, limestones	<1000	-	>0.5	>15	<20000	
GR-SR	Shallow reservoirs	<1000	-	>0.5	<15	-	
Natural							
GR-DNL	Deep lakes	0-1000		>0.5	>9		Thermal monomictic
GR-SNL	Shallow lakes	0-1000		>0.5	3-9		Polymictic
GR-VSNL	Very shallow lakes	0-1000		>0.5	<3		Polymictic

Table 2.1.2 -3. Köppen-Geiger's Classification of Climatic Regions of Greece.

Climate type	Description
Csa	Mediterranean hot summer climates
Csb	Mediterranean warm/cool summer climates
Cfa	Humid subtropical climates
Cfb	Oceanic climate
Dfa	Hot summer continental climates
Dfc	Subarctic or boreal climates
BSk	Cold semi-arid climate

2.1.2.2.2 EO Data

Landsat 8 satellite was launched in 2013 and bears two sensors, the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIR). It also includes a narrower near-infrared band, and a 12-bit radiometric resolution compared to the 8 bits of previous Landsat satellites (Bonansea et al., 2018; Olmanson et al., 2016). The main aim was to match as many records from the available Chl-*a* dataset as possible with Landsat 8 images. In cases where no Landsat 8 images were available or were cloud-covered, Landsat 7 ETM+ images were used. Landsat 7, launched in 1999, introduced the Enhanced Thematic Mapper Plus (ETM+), while its analysis was similar to TM except for a 60 m thermal and a new 15 m panchromatic band (Loveland and Dwyer, 2012). Since 2003, Landsat 7 had a sensor deficiency where the Scan Line Corrector (SCL) was off and even though those images are characterized by black line gaps (Tebbs et al., 2013), its radiometric and geometric analyses remain undisturbed (Bonansea et al., 2015).

In order to cover the 50 lakes throughout Greece and the different sampling dates, a 2013-2016 and 2018 time series of 296 Landsat imageries (102 L7 ETM+, 194 L8 OLI) were downloaded from the USGS (United States Geological Survey) Data Centre (https://earthexplorer.usgs.gov/). Images from both sensors reside in Landsat Collection 1 Level-1 category data products. Furthermore, the mean time window between the satellite overpass and the *in-situ* measurements is about ±15 days. More information about the bandwidths and spatial resolution of the aforementioned sensors can be found at NASA's official website (https://landsat.gsfc.nasa.gov/).

2.1.2.3 Satellite Data Pre-Processing

Pre-processing was carried out to the Landsat images using the Semi-automatic Classification plugin (SCP) of the free and open-source cross-platform desktop Quantum Geographic Information System (Q-GIS), v. 3.6.3-

Noosa. Images of both Landsat 7 ETM+ and Landsat 8 sensors were subject to the following pre-processing steps:

1. Conversion of images from Digital Numbers (DN) to the physical measure of Top of Atmosphere reflectance (TOA)
2. Atmospheric correction using the DOS1 method (Dark Object Subtraction 1; image-based technique), which was applied to all bands except for thermal ones
3. Creation of band stack set for each image. The band stack set of L7 ETM+ includes the bands B1 (blue), B2 (green), B3 (red), B4 (NIR), B5 (SWIR1) and B7 (SWIR2) while L8 set incorporates bands B2 (blue), B3 (green), B4 (red), B5 (NIR), B6 (SWIR1) and B7 (SWIR2).

More information about each correction including the theoretical background can be found at the SCP Documentation Release 6.2.0.1 (Congedo 2019). To ensure the use of only cloud-free pixels over the sampled lakes, the Cloud Masking QGIS plugin (https://smbyc.github.io/CloudMasking) was used. By using this tool, clouds, cloud shadow, cirrus, aerosols and ice/snow were masked for all Landsat images using the combination of the Fmask and Blue Band processes.

Pre-processing procedure of L7 images also included the retrieval of data that coincided with the aforementioned black diagonal stripes. By visual checking, sampling sites covered by those stripes were recognized and by employing GIS operations and Focal Statistics, the mean value within a circle of 7 cells around it for each input cell location was calculated. This radius was the most adequate among several trials. Then, by applying the Con and IsNull functions were replaced only the cells that had no values. In cases where part of sampled lakes was cloud covered, the SetNull and IsNull functions were combined with the Cloud mask calculated in earlier stages to remove the cloud biased pixels.

2.1.2.4 Statistical methods of analysis

2.1.2.4.1 Basic statistics and PCA

Basic statistical analysis among the Chl-*a* datasets of 2013, 2014, 2015, 2016 and 2018 was carried out including the calculation of mean, median, standard deviation and min-max. Based on the available lake characteristics (Table 2.1.2.-1) a Factor Analysis was used with the Varimax-Rotated Principal Component Analysis (PCA) for factor extraction method to interpret the major patterns of Chl-*a* variation within the whole dataset.

Furthermore, PCA 's basic aim was to explore and indicate the presence of inter-correlations among Chl-*a* concentrations, lakes' characteristics, climatic type, WFD typology and season's sampling and further indicate which of them are the most significant parameters affecting Chl-*a*

concentrations in studied lakes. The seasons are defined as: summer (June, July, and August), autumn (September, October and November) and spring (March, April, and May). Chl-*a* concentrations' distribution was grouped and categorized by the most significant variables/criteria that PCA indicated, by using box-plots diagrams. PCA was performed in SPSS Statistical Package (v. 24.0).

2.1.2.4.2 Developing Relationships between Landsat and Chlorophyll-a Data

There are several band combinations and transformations proposed in the relevant literature for establishing relationships between the Landsat reflectance data (independent) and Chl-*a*, log (Chl-*a*) and ln (Chl-*a*) data (Markogianni et al., 2018). Matthews (2011) reported that the wavelength areas where water reflects and scatters the majority of the entering solar radiation are those that are mostly used for the monitoring of water quality constituents. These wavelengths incorporate the water-leaving reflectance in visible and NIR wavelengths of the electromagnetic spectrum. Thus, to compare Chl-*a* concentrations and EO data, visible (blue, green and red), NIR and SWIR bands were used while aerosol, cirrus, panchromatic and TIRS spectral bands were excluded. In addition, Landsat band ratios, additions, subtractions, log and ln-transformations were added to the analysis to establish accurate and reliable estimation algorithms. Based on the respective literature, more than 75 available band transformations/combinations were developed, also including spectral indices such as Enhanced Vegetation Index (EVI; Liu and Huete, 1995), Normalized Vegetation Index (NRVI; Baret and Guyot, 1991), Normalized Difference Water Index (NDWI; McFeeters 1996), Modified Normalized Difference Water Index (MNDWI; Xu 2006), Green Normalized Difference Vegetation Index (GNDVI; Gitelson et al., 1996), Normalized Difference Vegetation Index (NDVI; Rouse et al., 1974) and SABI (Surface Algal Bloom Index; Alawadi 2010).

As a first step, the main objective was to distinguish the highest important predictors among the aforementioned band transformations/combinations by conducting a correlation analysis among them and the Chl-*a*, log(Chl-*a*) and ln (Chl-*a*) concentrations. Taking into account only the significant at the 0.01 level correlations, a threshold value of Spearman r was set to ±0.4 (moderate relationship; Dancey and Reidy, 2007). Variables that presented an r value equal or higher than the aforementioned, were selected. Then, those variables/predictors were further inserted - combined in various ways-in numerous stepwise MLRs while examining for statistical performance and residuals. Further criteria consisting multi-collinearity, tolerance factor and variance inflation factor (VIF) were applied

and checked to a subset of optimal models in order to further compare them and select the simplest models with higher accuracy (higher R^2). To develop a robust Chl-*a* algorithm, stepwise MLR analyses have been evaluated in the context of 2 scenarios, as also illustrated in Figure 2.1.2-2:

1st scenario: Basic aim of this scenario was the exploration of the potential of stepwise MLR analysis to establish robust Chl-*a* quantitative algorithms. The training dataset was randomly established and included the 80% of the whole available dataset (481 out of 565 Chl-*a* measurements, regardless of lake characteristics; Figure 2.1.2-2). This methodology is widely used in remote sensing of inland water quality and attempted to develop an algorithm that can be applied in multiple types of lakes.

2nd scenario: This scenario aimed at exploring different remote sensing derived Chl-*a* algorithms for various types of lakes according to the most significant lake characteristics as resulted by PCA of Chl-*a* values, lakes' climatic type, WFD typology and season's sampling. PCA-indicated parameters were combined in all possible ways, forming the various modelling cases examined to highlight the most optimal model. In this scenario, the various training datasets were subsets of the whole dataset of *in-situ* Chl-*a* values, defined by the respective combination of the PCA derived criteria. Each training dataset was then used in a separate MLR analysis (Figure 2.1.2-2). As an additional criterion, a confined time window of ±5 days was used, defined as the difference between field measurements and satellite overpass. Kloiber et al. (2002) observed that *in-situ* measurements within one day off the satellite image date resulted in the best calibrations, but larger number of ground observations with the longer time difference balances some of the loss of accuracy. Hence, since the available dataset of *in-situ* Chl-*a* measurements is quite large, the ±5 days window was determined presuming that during this period the water quality usually does not exhibit large and rapid water quality fluctuations.

During the analyses which included both scenarios mentioned above, it was ensured that the separate training and validation datasets contained representative samples of the data, 80% and 20%, respectively of the initial dataset. Training datasets of both scenarios comprised the *in-situ* Chl-*a* measurements of 2013-2016.

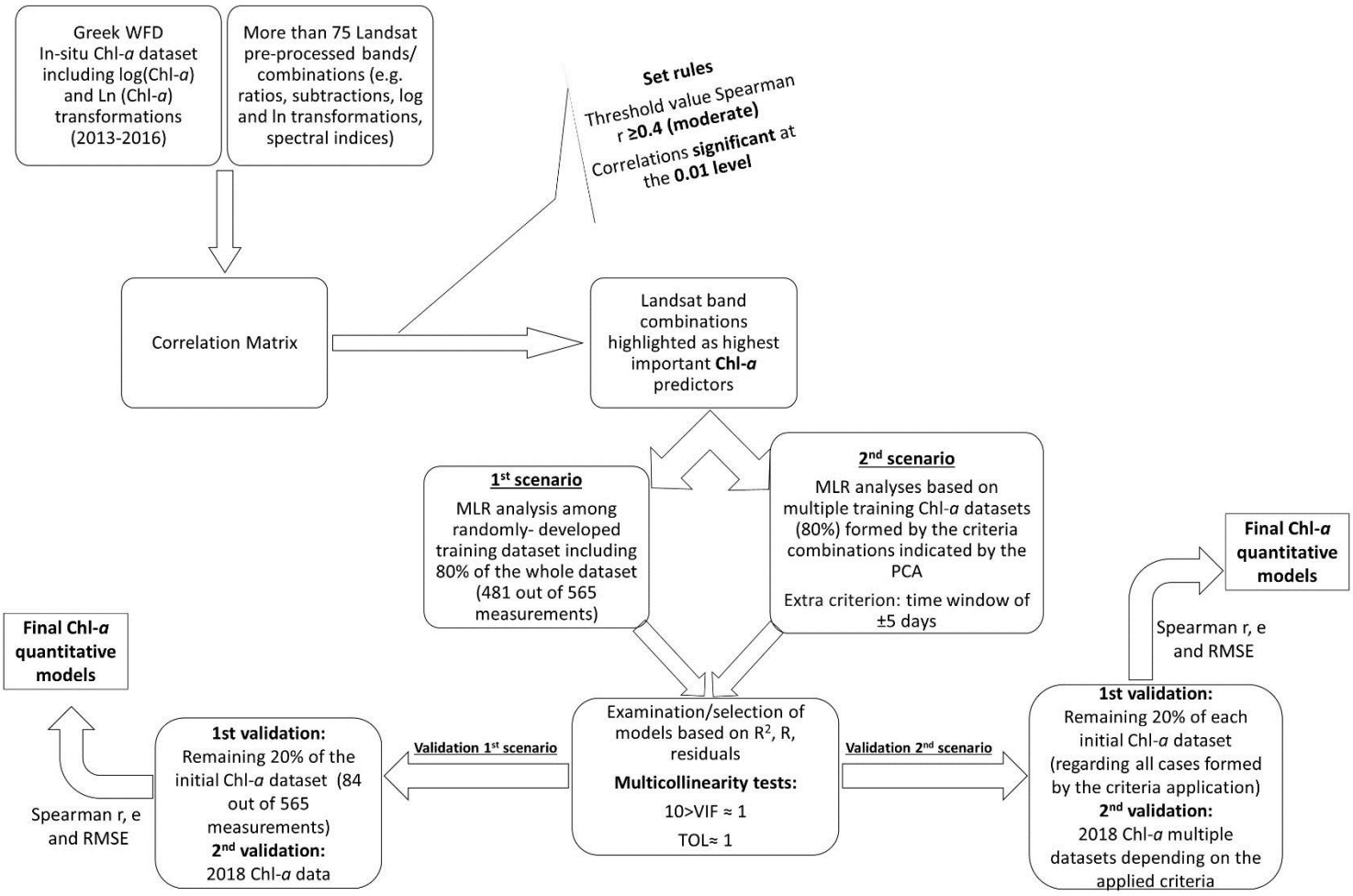


Figure 2.1.2-2. Flowchart summarizing the methodology followed in this study.

2.1.2.5 Validation approach

The qualitative capability of the derived algorithms was evaluated by regression analysis in two ways. Algorithms derived from the MLR based on the 1st scenario (random dataset), were applied in Landsat images associated with the respective validation dataset (20%) and in images of the year 2018 (Figure 2.1.2-2). The algorithms resulted from the 2nd scenario's training datasets (criteria), were applied in available Landsat imageries concerning the Chl-*a* of each respective validation dataset (20%) and the *in-situ* Chl-*a* values of 2018. After applying several models on the available images, the optimal ones were verified based on Spearman's r correlation coefficient, mean residual value (e) and RMSE (Root-Mean-Square Error) statistical indices (Figure 2.1.2-2).

2.1.3 Results

2.1.3.1 Statistical Analyses

Total measurements of *in-situ* Chl-*a* concentrations of years 2013-2016 and 2018 are 702, including all sampling campaigns in all lakes, while the most measurements presented in 2015 (Table 2.1.3-1). Sampling campaigns of 2014 indicated the highest range in Chl-*a* values (1026.7 μg/l) while in the rest of years the annual ranges are significantly lower. Minimum Chl-*a* value is 0.22 μg/l for years 2013 and 2014 whereas this value is increasing over the years. Despite this fact though, mean Chl-*a* value is remaining similar between different years, indicating no special deterioration trend in Greek lakes' water quality. Furthermore, datasets of all years are skewed right with low values while the kurtosis of all years is described as leptokurtic (fat tails).

Table 2.1.3-1. Descriptive statistics-Summary table of *in-situ* Chlorophyll-*a* of years 2013-2016 and 2018.

	Chl-*a* (μg/l)-2013	**Chl-*a* (μg/l)-2014**	**Chl-*a* (μg/l)-2015**	**Chl-*a* (μg/l)-2016**	**Chl-*a* (μg/l)-2018**	**Chl-*a* (μg/l)-all**
N	157	155	172	82	136	702
Minimum	.22	.22	.36	.45	.58	.22
Maximum	263.9	1026.9	286.96	292.7	361.7	1026.9
Mean	24.44	29.13	23.06	27.99	27.03	26.06
Std. Deviation	46.94	101.82	46.27	45.34	55.92	64.31
Skewness	3.05	7.7	3.24	3.23	3.83	7.87
Kurtosis	9.7	67.8	11.7	14.1	16.9	97.5

Performance of Factor Analysis (using as extraction method the Principal Component Analysis - PCA) included five (5) variables/criteria, e.g. sampling's season, *in-situ* Chl-*a* concentrations, lakes' nature (natural/artificial), WFD typology and climatic type. Factor Analysis was implemented to obtain an indication of underlying common factors (components) that explain the interrelationships among those aforementioned variables. The analysis initially extracts 5 components (Table 2.1.3-**2**). Finally, only the first three components with eigenvalues higher than 1 are retained (as those which represents a real underlying factor) in the extraction sums of squared loadings. The percentage of the total variance explained by the three components (calculated after the implementation of the varimax rotation

method) is 35.4 %, 21.9% and 20.8%, respectively and cumulatively those first three explain the 78% of the total variance.

Based on rotated component matrix's results (Table 2.1.3**-3**), communalities of studied variables are further discussed. Thus, the 53% of Chl-*a*'s variance is explained by the second component which also explains the 57% of the sampling's season variance. The first component explains the 10% of Chl-*a* variance which also explains the 81.5% of the variance of lakes' characteristics (natural/artificial) and the 79% of the variance of the lakes' WFD typology. The third component explains only 5% of the Chl-*a* variance and 90% of the climatic types' variance. Considering those results, the variables that mostly contribute and affect the variance of Chl-*a* concentrations are the lakes' characteristics (natural/artificial) and WFD typology followed by the samplings' season.

Table 2.1.3**-2.** Total variance explained, initial eigenvalues and extracted components.

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	1.78	35.5	35.6	1.78	35.6	35.6	1.8	35.4	35.4
2	1.12	22.4	57.9	1.12	22.4	57.9	1.1	21.9	57.4
3	1.01	20.2	78.1	1.01	20.2	78.1	1.04	20.8	78.12
4	.79	15.9	93.98						
5	.30	6.02	100						

Table 2.1.3-3. Rotated component matrix (Rotation Method: Varimax with Kaiser Normalization, Rotation converged in 5 iterations); (*Percentages of the variables total communalities explained by each component).

ITEM	Component								
	PCA Score 1	PCA Score 2	PCA Score 3	Communalities (1)	1*	Communalities (2)	2*	Communalities (3)	3*
Sampling's season	.25	**.75**	.26	0.06	6	0.56	**56**	0.07	7
Chl-*a* (μg/l)	-.31	.73	-.23	0.096	9.6	0.53	53	0.053	5.3
Natural/Artificial	**.9**	-.03	.06	0.81	**81**	0.0009	0.09	0.004	0.4
Köppen/ Climate type	-.08	.02	**.95**	0.006	0.6	0.0004	0.04	0.9	**90**
WFD Typology	**.89**	.01	-.15	0.79	**79**	0.0001	0.01	0.023	2.3

Concentrations of Chl-*a* of natural lakes were found notably higher in comparison to artificial while the highest measured values are detected in artificial lakes during the summer and autumn months (outlier values) while in natural during autumn (Figure 2.1.3-1a). In general, seasonality of Chl-*a*

concentrations is more evident in natural rather than in artificial lakes. The highest discrepancies are detected at natural lakes (Zazari and Voulkaria lakes, 07/2014) followed by autumn sampling campaigns (eg Zazari lake, 09/2014).

Measured Chl-*a* concentrations in both artificial and natural lakes present a greater range during summer while most outliers of artificial lakes are illustrated also during summer (Kerkini-08/2014; 08/2016 and Karla reservoirs, 07/2014) and autumn seasons (Kerkini reservoir, 10/2015). Spring also presents significant differences between natural and artificial lakes where median value of the latter is evidently decreased.

The same Chl-*a* pattern is illustrated in a different way by the second boxplot (Figure 2.1.3-1b). As far as the Chl-*a* distribution based on the lakes' typology is concerned, the highest values are detected at natural GR-VSNL and GR-SNL shallow (Zazari lake) and very shallow lakes (Voulkaria lake). These cases are followed by the values measured in deep natural lakes GR-DNL (Yliki lake), while the lowest Chl-*a* values were detected in artificial shallow reservoirs (Pournari II and Stratos reservoirs).

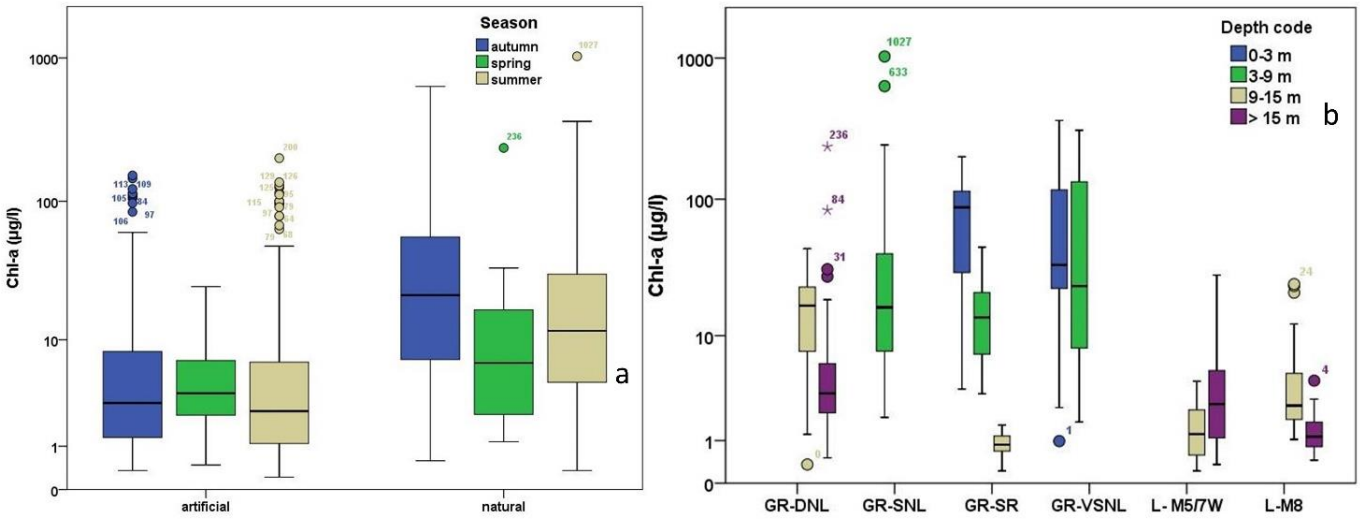


Figure 2.1.3-1. Boxplots presenting basic descriptive statistics (median, percentiles, min-max, outliers and extremes) of chlorophyll-*a* concentrations of years 2013-2016 and 2018 (a) grouped by the lakes' characteristics and categorized by the sampling's season and (b) grouped by the lakes' WFD typology and categorized by lakes' mean depth. Mean depth classification is based on WFD typology characteristics as explained in Table 2.1.2-1.

2.1.3.2 Relationships between Landsat and Chlorophyll-a Data

Blue, Green, Red, NIR and SWIR Landsat 8 and Landsat 7 ETM+ bands and more than 75 different band transformations and indices were used for investigating the most suitable relationship to estimate Chl-*a* concentrations throughout 50 Greek lakes.

Correlation analysis among all available variables and Chl-*a*, log(Chl-*a*) and ln (Chl-*a*) concentrations resulted in Spearman r values that ranged from -0.6 to +0.6. Based on the correlation matrix, the highest important predictors that met the criterion of the set threshold value of Spearman (r) ±0.40 in relation to log(chl-*a*) are the following: Blue/Green, Green/Blue, Blue/Red, Red/Blue, Red/ Green, Red/SWIR1, Log(Blue/Green), Log(Blue/Red), LogBlue/LogGreen, LogBlue/LogRed, (Blue-Red)/Green, LnBlue/LnGreen, LnBlue/LnRed, LnGreen /LnBlue, LnRed/LnBlue, LnRed /LnSWIR1 and LnRed /LnSWIR2. Those variables/predictors were further inserted in several combinations in numerous stepwise linear regressions. As the number of the generated models is quite large, hereby are included only the most significant models from a statistical point of view.

Based on the 1st scenario, MLR analysis among the indicated variables generated several models with good performance. Based on tests on statistical significance of the bi coefficient of the independent factors (t-test with p values less than 0.05) and on tests for multicollinearity (Variance Inflation Factor-VIF with values higher than 1 and less than 10 and Tolerance higher than 0.1) the following model was finally selected as satisfactory (Equation 2.1.3-1; Table 2.1.3-4).

$$\log Chla = 3.599 - 0.63 * \left(\frac{blue}{red}\right) - 2.183 * \left(\frac{\ln red}{\ln swir2}\right) \quad (2.1.3\text{-}1)$$

Table 2.1.3-4. Regression analysis statistics and models' summary among reflectance values and log-chlorophyll-*a* concentrations (dependent variable).

Scenario /Model	R	R^2	Adjusted R^2	Std. Error of the Estimate	Durbin-Watson
1A	.654	.427	.425	.525	1.95
Predictors: (Constant), BLUE/RED, LNRED/LNSWIR2					

Descriptive statistics results obtained for the training and validation datasets (regarding logchl-*a*) suggested that there is no discrepancy in the mean (training: 0.8; validation: 0.79) and standard deviation (training: 0.69; validation: 0.68) values, between the two groups. To ensure statistically significant results, an independent t-test for the mean values was also implemented. Based on these results (t-value 0.276, p= 0.783), it is assumed that there is no difference in the group means.

Random selection of Chl-*a* measurements, and MLR analysis (1st scenario, 1A model) yielded a model accompanied by a value of coefficient of determination equal to 0.43, including the band ratios lnred/lnswir2 and blue/red (Table 2.1.3-4). The standard error of the estimate is equal to 0.53 and the collinearity statistics were considered acceptable.

Concerning the 2nd scenario, MLR analysis conducted among the selected predictors and the various combinations of the criteria (Table 2.1.3-5) indicated by the PCA, generated several models. Examination though of aforementioned statistical indices highlighted two (2) models (Table 2.1.3.2-6; Equations 2.1.3-2 and 2.1.3-3). It should also be noted that the various criteria combinations always included in the examined dataset the values with Chl-*a* concentrations lower than 500 μg/l (based on the normal Q-Q Plot of Chl-*a* concentrations and outliers analysis) and mean depth higher than 5 m to surely avoid the bottom reflectance noise (McKinna and Werdell, 2018). The criterion of the time window of ±5 days, was applied in cases with statistically significant results to examine and explore any further improvement. The 13th criterion case (Table 2.1.3.2-7) aimed to establish accurate algorithms considering only the mean depth differentiation of the lakes by including the shallow reservoirs (<15 m) as well as shallow (3-9 m) and some deep (>9 m) natural lakes.

Table 2.1.3-5. Description of cases including all possible combinations of criteria applied. The criteria were selected according to PCA results.

No of case regarding the criteria application	**Combination of criteria**
1	Mean depth > 5 m, natural lakes, all seasons
2	Mean depth > 5 m, artificial lakes, all seasons
3	Mean depth > 5 m, natural lakes, season spring
4	Mean depth > 5 m, natural lakes, season autumn
5	Mean depth > 5 m, natural lakes, season summer
6	Mean depth > 5 m, artificial lakes, season spring
7	Mean depth > 5 m, artificial lakes, season autumn
8	Mean depth > 5 m, artificial lakes, season summer
9	Mean depth > 5 m, season spring
10	Mean depth > 5 m, season autumn
11	Mean depth > 5 m, season summer
12	Mean depth > 5 m, date difference (sampling/satellite): ±5 days
13	> 3 m mean depth <15 m

Table 2.1.3-6. Regression analysis statistics and models' summary among reflectance values and log-chlorophyll-*a* concentrations (dependent variable).

Scenario /Model	R	R^2	Adjusted R^2	Std. Error of the Estimate	Durbin-Watson
2A	.776	.602	.587	.432	1.666
2B	.757	.574	.563	.291	2.034

The 2A regression model is the following:

$$\log Chla = 4.443 - 1.421 * \left(\frac{blue}{green}\right) - 3.454 * \left(\frac{\ln red}{\ln swir2}\right) + 1.304 * \left(\frac{red}{green}\right) \quad (2.1.3\text{-}2)$$

with 85 number of selected cases; based on the criteria: Chl-*a* concentration <500 μg/l, mean depth > 5 m, natural lakes, all seasons.

The equation of the 2B regression model is:

$$\log Chla = 2.919 - 2.011 * \left(\frac{\ln red}{\ln swir1}\right) + 1.449 * \left(\frac{red}{green}\right) - 1.441 * \left(\frac{\ln red}{\ln blue}\right) \quad (2.1.3\text{-}3)$$

with 125 number of selected cases and developed based on the following criteria: Chl-*a* concentration <500 μg/l, mean depth > 5 m, artificial lakes, all seasons and date difference between sampling and satellite overpass ±5 days.

The optimal Chl-*a* models developed based on the 2nd scenario incorporated the band ratios lnred/lnswir2, blue/green, lnred/lnblue, red/green and lnred/lnswir1 while the coefficient of determination was equal to 0.57 (2B) and 0.6 (2A) and the standard error of the estimate 0.29 (2B) and 0.4 (2A). Collinearity statistics (Tolerance and VIF) of the coefficients were also considered acceptable. In general, the highest beta coefficient values accompany the band ratios red/green, blue/red and blue/green while Durbin-Watson's statistic test indicates an absence of autocorrelation especially in the residuals of models developed in the 2nd scenario.

2.1.3.3 Regression models' validation

Developed regression models were validated for both scenarios while validation datasets were different for each model. Concerning the 1st scenario's analysis (1A), the model was applied in Landsat 8 and 7 ETM+ images connected to the remaining 20% of the respective validation dataset (47 measurements; first validation) and then for the second validation by

using images and all available Chl-*a* measurements of the year 2018 (N= 71; Table 2.1.3-7). Models 2A and 2B were validated based on the 20% of remaining datasets characterized by the set criteria (2A, N=20 measurements; 2B, N=29; Table 2.1.3-7) and subsequently the *in-situ* Chl-*a* values of 2018 were used for the 2nd validation process (2A, N=23; 2B, N=40). More specifically, 2A model was validated by using the measurements concerning the natural lakes while for the validation of 2B model, measurements of artificial lakes with sampling/satellite date difference of ±5 days, regardless the sampling's season were used. It should also be noted that validation results have been enhanced by extracting outlier values, concerning the residuals between *in-situ* and satellite values.

Table 2.1.3-7. Statistical indices used to validate the selected algorithms (**. Correlation significant at the 0.01 level (2-tailed) and *. at the 0.05 level (2-tailed), respectively).

	1st validation (20%)					**2nd validation (2018 data)**				
Scenario/ Models	**Spearman r**	**Average *in-situ* Chl-*a* (*μg/l*)**	**Average satellite Chl-*a* (*μg/l*)**	**Average residuals (*μg/l*)**	**RMSE (*μg/l*)**	**Spearman r**	**Average *in-situ* Chl-*a* (*μg/l*)**	**Average satellite Chl-*a* (*μg/l*)**	**Average residuals (*μg/l*)**	**RMSE (*μg/l*)**
1A Training dataset N=481	.688** N=47	2.9	4.3	-1.3	3.96	.758** N=71	4.8	5.01	-0.19	4.6
2A Training dataset N=85	.782** N=20	10.9	9.7	1.3	5.6	.697** N=23	5.6	6.9	-1.3	4.2
2B Training dataset N=125	.622** N=29	2.3	2	0.3	1.53	.593** N=40	3	2.5	0.5	2.3

The 1st scenario yielded a model (1A) that even though is characterized by a quite low coefficient of determination (0.43), is accompanied by two validation processes with high Spearman (r) values (0.69 and 0.76, respectively, significant at the 0.01 level) and quite large validation datasets (N=47 and 71). Additionally, the differences among the mean *in-situ* and satellite derived values in both validations are not high (Figure 2.1.3-2a), while mean residual (Figure 2.1-5) and RMSE values are quite satisfactory for both validation processes (particularly of 2018 dataset; RMSE 1st Validation:3.96 μg/l; RMSE 2nd Validation: 4.6 μg/l; Table 2.1.3-7). Hence, this specific model could be included as one of the proposed Chl-*a* quantitative

algorithms; however it is inferior compared to special natural and artificial models.

Hence, concerning the models resulted from the 2nd scenario and their validations, 2A and 2B models are those that are mostly proposed. Model 2A is derived from a regression analysis with R^2 equal to 0.6 and Spearman values equal to 0.78 (first validation) and 0.7 (second validation), respectively. Mean *in-situ* and satellite Chl-*a* values are similar (Figure 2.1.3-2b) while RMSE values are quite low, 5.6 μg/l (first validation) and 4.2 μg/l for the second validation, respectively. Likewise, 2B Model, demonstrated high enough Spearman values (0.62 and 0.59), low average residuals (Figure 2.1.3-3), low RMSE values (1.53 and 2.3 μg/l) and is also accompanied by one of the largest validation datasets (2018 dataset; N=40).

Furthermore, Chl-*a* maps of selected lakes were created after the application of the resulted algorithms (Figure 2.1.3-4). Landsat 8 OLI satellite image of 11/08/2013 was used in order to produce the satellite derived spatial distribution of Chl-*a* values of this day while the respective *in-situ* values of those lakes have been sampled with -2 and +5-days difference from the aforementioned date. Application of 1A model resulted in Chl-*a* concentrations that range from 0.18 to 58.9 μg/l and from 0.71 to 61.7 μg/l for artificial and natural lakes, respectively (Figure 2.1.3-4a). Application of 2A model in natural lakes yielded Chl-*a* values ranging between 2.75 and 70.8 μg/l (Figure 2.1.3-4b) while 2B model resulted in Chl-*a* values varying from to 0.019 to 12.6 μg/l (Figure 2.1.3-4c), as far as the artificial lakes are concerned. 1A model yielded quite higher Chl-*a* values in relation to 2B model, concerning the artificial lakes, while in natural lakes the satellite derived Chl-*a* values from 1A model are slightly lower than those produced by the 2A model. Moreover, it should be noted that *in situ* values concern point samples whereas satellite derived values relate to the spatially distributed Chl-*a* concentrations along the lakes' surfaces.

Summarizing the information derived from the validation process, the most optimal Chl-*a* assessment model throughout the WFD Greek lakes when no information is available about their characteristics is 1A. When the studied lakes are natural, the Chl-*a* model becomes more complex (2A) incorporating three band ratios and when artificial lakes are the case, then model 2B is proposed with the condition that the date difference between field measurements and satellite overpass ranges from -5 to +5 days.

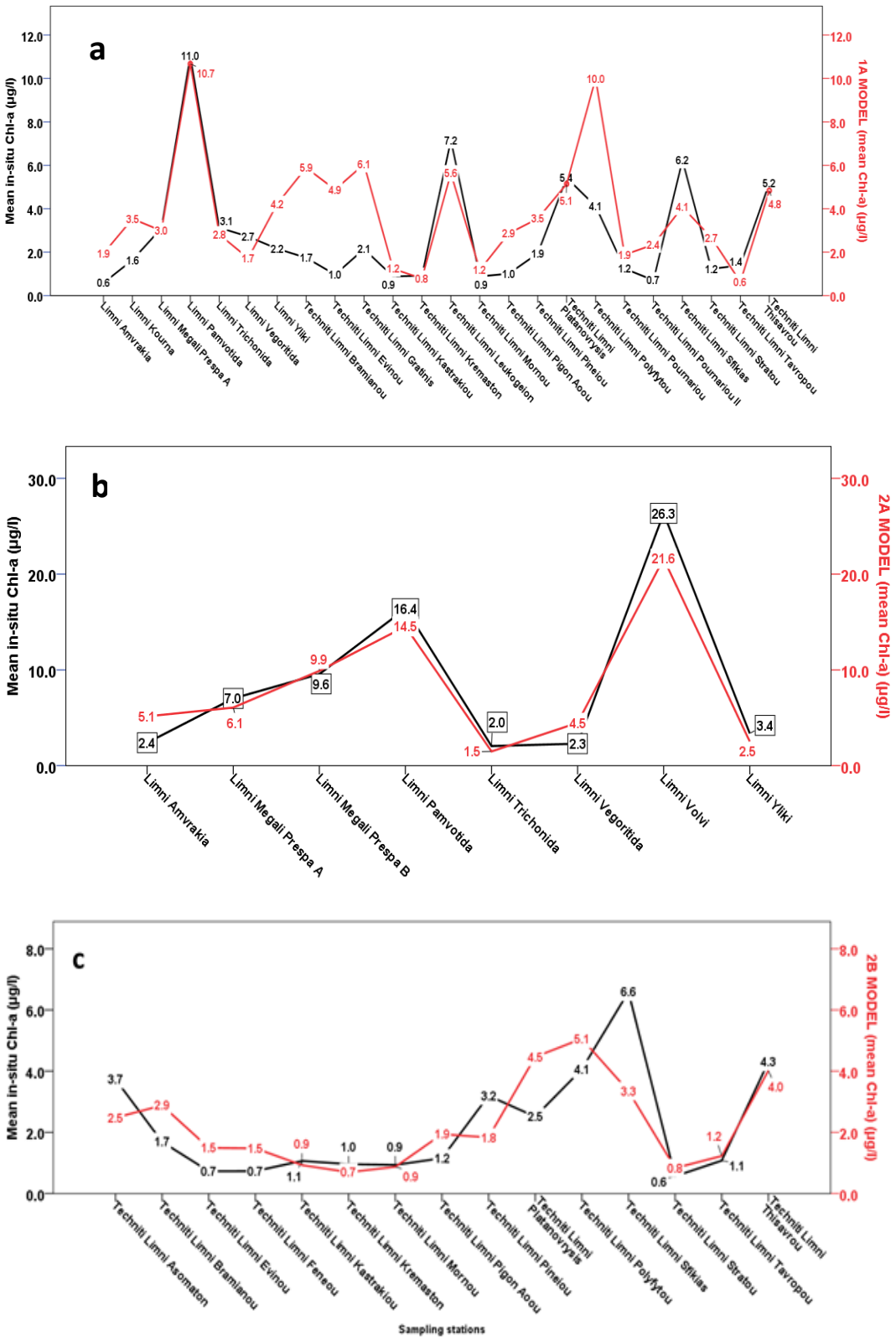


Figure 2.1.3-2. Spatial distribution per lake of mean *in-situ* Chl-*a* concentrations and mean satellite derived Chl-*a* concentrations based on the application of the a)1A Model, b)2A Model and c)2B Model, on the validation datasets including the 20% of the initial dataset).

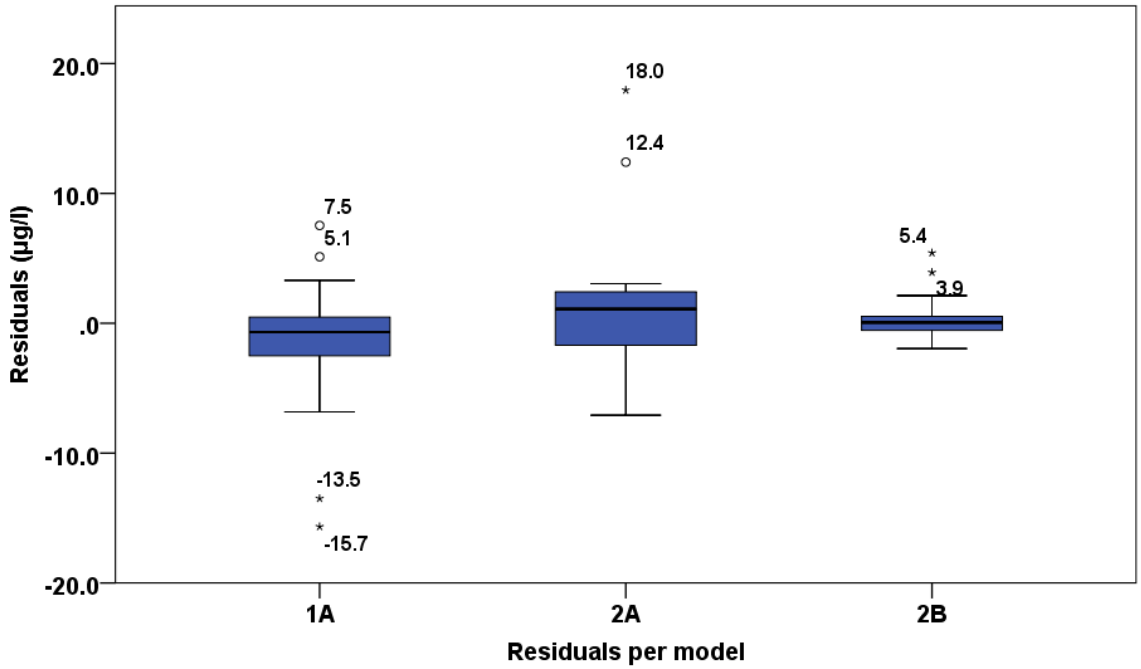


Figure 2.1.3-3. Boxplots presenting basic descriptive statistics of residuals per model.

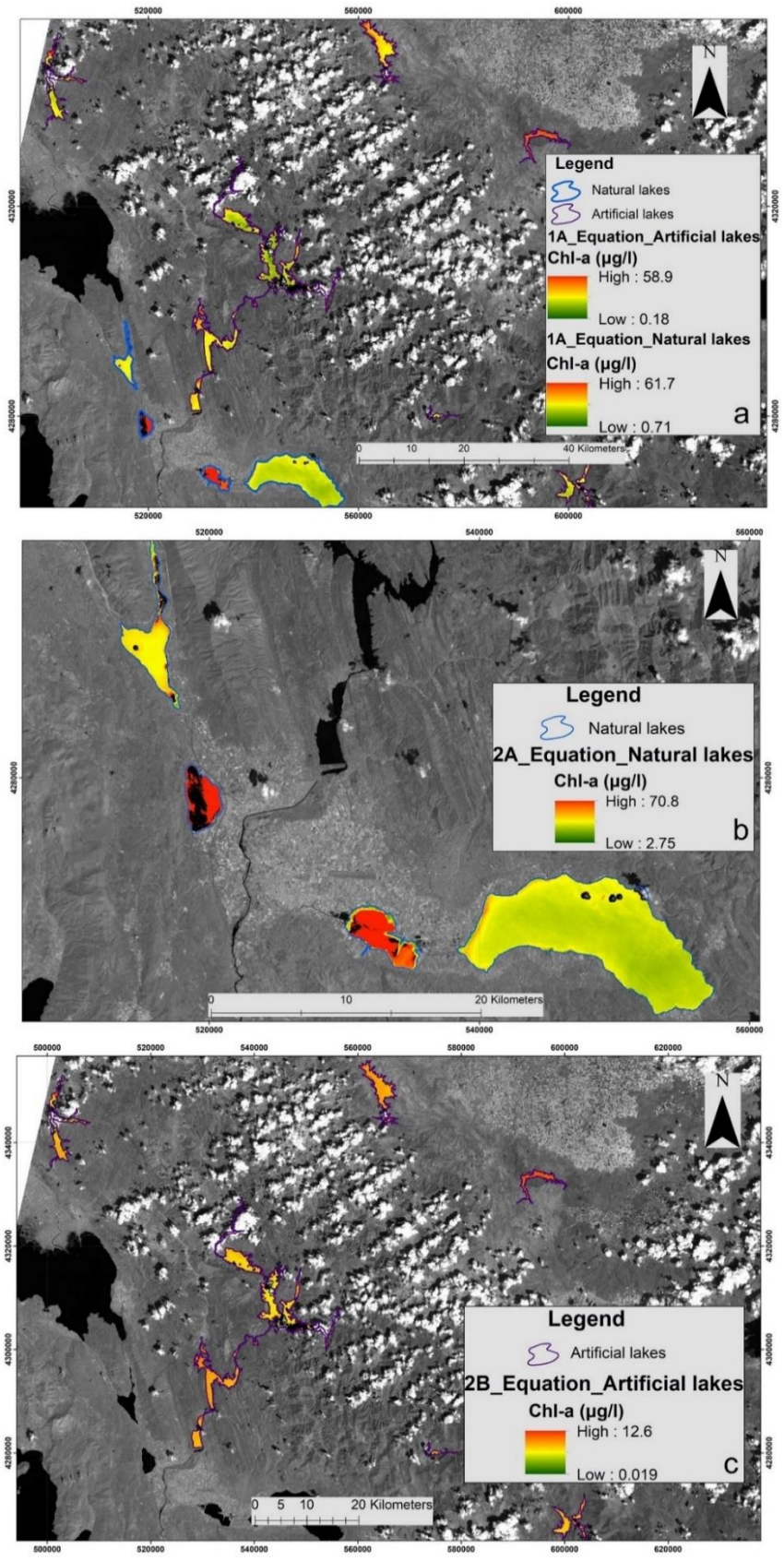


Figure 2.1.3-4. Satellite-derived Chl-*a* maps (on 11/08/2013) of selected lakes after the application of 1A (a), 2A (b) and 2B (c) models (WGS_1984, UTM Zone 34 N Coordinate system).

2.1.4 Discussion

WFD application in Greece concerning Lake Waterbodies yielded so far, a significant Chl-*a* dataset including years 2013-2016 and 2018. Statistical analyses of Chl-*a* measurements indicated that natural lakes presented notably higher concentrations in relation to artificial. Those results indicate a degraded water quality of natural lakes comparing to artificial with some lakes being characterized as vulnerable to eutrophication. Dense algal blooms, provoking high turbidity, are a frequent indicator of lake eutrophication (Schindler et al., 2008). Independent studies on natural lakes have demonstrated a strong correlation between Chl-*a* and total phosphorus concentrations (Sakamoto, 1966; Jones and Bachmann, 1976; Canfield 1979). Canfield (1979) predicted total phosphorus concentrations and trophic states (by measuring Chl-*a* and Secchi depths) in natural and artificial lakes through the development of empirical phosphorus models. Canfield (1979) concluded that while total phosphorus concentrations can be predicted equally well in natural and artificial lakes, prediction of lake trophic state was less reliable in artificial lakes. Relationships between Chl-*a* concentrations to total phosphorus and Secchi disc transparency were less precise in artificial lakes than natural lakes while this difference was attributed to non-algal turbidities. Non-algal turbidities were detected in the artificial lakes and recorded as important water clarity determinants for many of this type's lakes (Canfield and Bachmann, 1981). Based on this hypothesis, phosphorus concentrations in studied lakes may be examined to ascertain the drivers of decreased water clarity and levels of Chl-*a* concentrations in both types of Greek lakes.

Harmonization of Landsat 7 ETM+ and 8 OLI images yielded three Chl-*a* qualitative models including the ratios blue to green and red, red to green and blue, and the ln transformed bands SWIR1 and SWIR2. According to Barrett and Frazier (2016) ratios between either chlorophyll absorption bands (red and blue) or chlorophyll reflectance bands (green and NIR) with either of the two SWIR bands are highlighting the spectrum's part influenced by chlorophyll. Thus, modelled values are better correlated with actual *in-situ* ones.

MLR analysis using the total amount of the available Chl-*a* dataset resulted to a quite reliable assessment model (R=0.65). Next, it was explored what are the most significant parameters affecting the variance of Chl-*a* concentrations in studied lakes and the outcome implied the lakes' characteristics (natural/artificial) and WFD typology followed by the samplings' season. The final models were separately developed for natural (R=0.78) and artificial lakes (R=0.76), with the latter being accompanied by a date difference between *in-situ* and satellite data ranging ±5 days. Those models were proven to be better -based on statistical indices concerning the validation process - in relation to the one yielded from the total amount of the

dataset. This superiority highlights the significance of the information acquisition concerning the studied areas.

Results of this study are in accordance with others similar studies exploring properties of inland waters using either Landsat or other EO spaceborne sensors. Gholizadeh et al. (2016) have conducted a detailed review on water quality parameters that are widely estimated using remote sensing techniques. As authors noted, most Chl-*a* assessment models use a wavelength near 675 nm and 700 nm. Many researchers have developed empirical Chl-*a* algorithms using various but basically common image bands; Nas et al. (2007) used the visible near-infrared (VNIR) and the shortwave infrared (SWIR) of Terra/ASTER for Chl-*a* mapping, presenting a R^2 value 0.86. Zhang and Han (2015) used the coastal, blue, red and green Landsat 8 OLI bands to map Chl-*a* concentrations in Laizhou Bay while Kim et al. (2014) utilized the blue, NIR and the ratio blue to red Landsat 8 OLI bands to measure Chl-*a* in the Fjord of Svalbard, in arctic sea with R^2 value 0.6. Lim and Choi (2015) also used Landsat-8/OLI in order to monitor water quality of Nakdong River in Korea and presented high correlations among Chl-*a* and OLI bands especially the green and NIR bands and the band ratio NIR to green (Pearson's correlation coefficient of -0.7, 0.71 and -0.64, respectively). Bonansea et al. (2018) tried to generate a different Chl-*a* model for different Landsat sensors (5 TM, 7 ETM+ and 8 OLI) in the largest artificial reservoir in Córdoba province (Rio Tercero, Argentina). Overall, they observed that each Landsat sensor can be used to estimate Chl-*a* in the reservoir while the best model for TM sensor included a combination of green, red and NIR band, and the ratio green/red ($R^2 = 0.92$) and for ETM+ sensor ($R^2 = 0.91$) the green and SWIR-1 bands and the ratio red/green.

While several satellite sensors can be used for Chl-*a* determination, mapping Chl-*a* in Case 2 waters is a complicated task since the optical properties are measured based on a compound of dissolved organic matter, dead organic and inorganic particulate matter and phytoplankton (Chl-*a*). Therefore, Chl-*a* determination is characterized by less accuracy as these constituents are not statistically correlated. Taking this shortcoming into account, we tried to use spectral band ratios which decrease irradiance and atmospheric biases in the sensor's signal (Dekker and Peters, 1993) and more than one band, since then the scattering and absorption of Chl-*a* are better studied (Dekker et al., 1991). Furthermore, according to Kloiber et al. (2002), all significant band combinations for chlorophyll include at least one of the short-wave infrared bands, thus SWIR bands were incorporated into this study's analysis in order to produce optimal assessment Chl-*a* models.

A major difference in relation to aforementioned studies is that the study area used in the present study incorporates 50 different lake systems throughout Greece covering a broad geographic area and a wide range of limnological conditions, while the majority of the respective literature focuses

on regional scales and discrete inland water bodies. Large spatial scales require greater computational potential; thus, the release of the Google Earth Engine platform dramatically increased the scale at which earth observation research can take place (Topp et al., 2020). An example of extended study area is the research by Lin et al. (2018) who combined *in-situ* Chl-*a* data from 1157 lakes (2007) with Landsat data and developed a well-validated lake national model (RMSE = 34.9 μg/L), by using machine learning algorithms built into Google Earth Engine. Another example regarding Greece is the study by Kontopoulou et al. (2017) who used Landsat 8 images and WFD Chl-*a* and turbidity datasets concerning 11 lakes for years 2013-2015. They conducted regression analysis by using Matlab scripts and also examined the effect of the time difference between satellite and field data. In that study, an R^2 of 0.78 (log Chl-*a*, n=168) was reported for a time window 0-15 days, while R^2 reached 0.8 (n=39) for a narrower three-day time difference.

Since Chl-*a* concentrations in lakes cannot be accurately determined due to aforementioned restrictions, we consider that the proposed empirical models are reliable and should be applicable to most natural and artificial lakes within Greece. One limitation of empirical models is their restriction to confident assessments only within the range and setting of the input data. This restriction limits their application across spatiotemporal domains (Topp et al., 2020), risk which to a large extent is restrained since training datasets of this study include the majority of Greek lakes and three sampling seasons. However, the general applicability and potential limitations of this approach have not been thoroughly addressed, hence further improvement will be explored as soon as the latest WFD datasets are released.

Furthermore, clearly, there are some factors that should be taken into consideration in this study, affecting the accuracy of Chl-*a* quantification in Greek studied lakes.

1. An uncertainty of accuracy regarding the location of sampling points. Sampling in lakes requires special attention as winds and other external factors (e.g. season, lake depth and changes in water level, ease of proximity) contribute to potential transpositions of sampling sites.
2. The implementation of the DOS1 atmospheric correction method has not been validated in order to assure that atmosphere biases have been completely removed. However, this method is widely used by the EO community and has proven useful when no atmospheric measurements are available.
3. Optimal models have been applied to lake surfaces accrued from lake shapefiles, acquired from the Environment and Energy Ministry's website. Since no classification between land and water

has been conducted, there is the possibility that some pixels, covering land, hinder Chl-*a* quantification with high accuracy.

All in all, it should be noted that EO is recommended to be combined with conventional *in-situ* water sampling in order to achieve high assessment accuracy. Such a synergistic approach in conjunction with cooperation among government and scientists contribute to increased data retrieval, obtained knowledge of the lakes' water quality and by extension to better protection and pollution mitigation measures.

2.1.5 Conclusions

In this study a methodological framework has been proposed for quantifying Chl-*a* pigments using Earth Observation (EO) data from Landsat 7 ETM+ and 8 OLI sensors. Its practical use is evaluated in a total of 50 lake water bodies (natural and artificial) during 2013-2018, consisting the National Lake Network Monitoring of Greece in the context of Water Framework Directive (WFD).

Use of geoinformation technologies, such as of EO and GIS, in combination with conventional field surveying and spatial data analysis methods are the most efficient ways forward for monitoring water quality parameters in lakes. Application of WFD in Greece has resulted in a significant dataset of various water quality parameters concerning, in this case, 50 lake water bodies with different morphological characteristics and other properties. The integration of spectral information from two Landsat sensors and statistical analyses employing principle component analysis and stepwise MLR analyses yielded statistically significant results. Optimal models were developed in this study, separately for natural and artificial lakes, and increased the feasibility of Chl-*a* assessment with high accuracy. The majority of the respective literature focuses on discrete inland water bodies reporting the most accurate and statistically significant models. Monitoring of water quality in large spatial scales though, as in this study, may result in sustainable water resources management even though the models may be statistically weaker.

Since Chl-*a* is the major indicator of trophic state, considered as one of the top water pollution indices related to eutrophication, this study supports WFD application concerning the perpetual water quality monitoring of Greek lakes. WFD application throughout Europe aims at the monitoring of hydromorphological, physicochemical and biological data to assess ecological status of surface waters. Those data include optically active constituents of water that interact with light (e.g. Chl-*a*) and can be measured using remote sensing and also others that lack optical properties. Some examples are pH, dissolved oxygen and nutrient concentrations. Monitoring of those properties, characterized by low signal noise ratio, by using geoinformation technologies-

in particular EO and GIS- is a challenging task and has motivated us to pursue it in the near future, exploiting the WFD monitoring results.

Acknowledgments: Data used in this study have been collected in the framework of the National Water Monitoring Network for lakes, according to the Joint Ministerial Decision 140384/2011, implemented by The Goulandris Natural History Museum, Greek Biotope/Wetland Centre (EKBY). The Network is supervised by the Directorate for the Protection and Management of Water Resources of the Special Secretariat for Waters of the Ministry of Environment and Energy.

2.1.6 References

Alawadi, F. Detection of surface algal blooms using the newly developed algorithm surface algal bloom index (SABI). Proceedings of the International Society for Optics and Photonics **2010**, 7825:1–14.

Alymkulova, B.; Abuduwaili, J.; Issanova, G.; Nahayo, L. Consideration of Water Uses for Its Sustainable Management, the Case of Issyk-Kul Lake, Kyrgyzstan. *Water* **2016**, *8*, 298.

American Public Health Association (APHA). Standard methods for the examination of water and wastewater, 17th ed. American Public Health Association, Washington, D.C., 1989.

Baret, F.; Guyot, G. Potentials and limits of vegetation indices for LAI and PAR assessment. Remote Sens. Environ. **1991**, 35, 161–173.

Barrett, D.; Frazier A. Automated method for monitoring water quality using Landsat imagery. *Water* **2016**, 8(6), 257. https://doi.org/10.3390/w8060257.

Bonansea, M.; Bazán, R.; Ledesma, C.; Rodriguez, C.; Pinotti, L. Monitoring of regional lake water clarity using Landsat imagery. *Hydrology Research* **2015**, 46 (5), 661–670. https://doi.org/10.2166/nh.2014.211.

Bonansea, M.; Rodriguez, C.; Pinotti, L. Assessing the potential of integrating Landsat sensors for estimating chlorophyll-a concentration in a reservoir. *Hydrology Research* **2018**, 49 (5), 1608–1617. https://doi.org/10.2166/nh.2017.116.

Canfield, D. Jr.; Bachmann, R. Prediction of total phosphorus concentrations, chlorophyll a, and secchi depths in natural and artificial lakes. *Canadian Journal of Fisheries and Aquatic Sciences* **1981**, 38 (4): 414-423. https://doi.org/10.1139/f81-058.

Canfield, D.E. Jr. Prediction of total phosphorus concentrations and trophic states in natural and artificial lakes: the importance of phosphorus sedimentation. Retrospective Theses and Dissertations. Iowa State University, 1979.

Congedo, L. Semi-automatic Plugin Documentation. Release 6.2.0.1. https://fromgistors. blogspot.com/p/user-manual.html, 2019. Accessed on 01/06/2019.

Dancey, C.P.; Reidy, J. Statistics without Maths for Psychology. 4th edition, Harlow: Pearson Education, 2007.

Dekker, A.; Peters, S. The use of the thematic mapper for the analysis of eutrophic lakes: a case study in the

Netherlands. *International Journal of Remote Sensing* **1993**, 14 (5), 799–821. https://doi.org/10.1080/01431169308904379.

Dekker, A.G.; Malthus, T.J.; Seyhan, E. Quantitative modeling of inland water quality for high-resolution MSS systems. *IEEE Transactions on Geoscience and Remote Sensing* **1991**, 29 (1), 89–95. https://doi.org/10.1109/36.103296.

Duan, W.; He, B.; Takara, K.; Luo, P.; Nover, D.; Sahu, N.; Yamashiki, Y. Spatiotemporal evaluation of water quality incidents in Japan between 1996 and 2007. *Chemosphere* **2013**, 93 (6), 946–953. https://doi.org/10.1016/j.chemosphere.2013.05.060.

Duan, W.; Takara, K.; He, B.; Luo, P.; Nover, D.; Yamashiki, Y. Spatial and temporal trends in estimates of nutrient and suspended sediment loads in the Ishikari river, Japan, 1985 to 2010. *Science of the Total Environment* **2013**, 461, 499–508. https://doi.org/10.1016/j.scitotenv.2013.05.022.

El-Din, M.S.; Gaber, A.; Koch, M.; Ahmed, R.S.; Bahgat, I. Remote sensing application for water quality assessment in lake Timsah, Suez canal, Egypt. *Journal of Remote Sensing Technology* **2013**, 1 (3), 61-74.

George, D. The airborne remote sensing of phytoplankton chlorophyll in the lakes and tarns of the english lake district. *International Journal of Remote Sensing* **1997**, 18 (9), 1961–1975. https://doi.org/10.1080/014311697217972.

Gholizadeh, M.; Melesse, A.; Reddi, L. A comprehensive review on water quality parameters estimation using remote sensing techniques. *Sensors* **2016**, 16 (8), 1298. https://doi.org/10.3390/s16081298.

Giardino, C.; Bresciani, M.; Cazzaniga, I.; Schenk, K.; Rieger, P.; Braga, F.; Matta, E.; Brando, V.E. Evaluation of multi-resolution satellite sensors for assessing water quality and bottom depth of lake Garda. *Sensors* **2014**, 14 (12), 24116–24131. https://doi.org/10.3390/s141224116.

Gitelson, A.; Mayo, M.; Yacobi, Y.Z.; Parparov, A.; Berman, T. The use of high-spectral-resolution radiometer data for detection of low chlorophyll concentrations in lake Kinneret. *Journal of Plankton Research* **1994**, 16 (8), 993–1002. https://doi.org/10.1093/plankt/16.8.993.

Gitelson, A.A.; Dall'Olmo, G.; Moses, W.; Rundquist, D.C.; Barrow, T.; Fisher, T.R.; Gurlin, D.; Holz, J. A simple semi-analytical model for remote estimation of chlorophyll-a in turbid waters: validation. *Remote Sensing of Environment* **2008**, 112 (9), 3582–3593. https://doi.org/10.1016/j.rse.2008.04.015.

Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. Remote Sens. Environ. **1996**, 58, 289–298.

Jones, J. R.; Bachmann, R. W. Prediction of phosphorus and chlorophyll levels in lakes. *Journal of Water Pollution Control Federation* **1976**, 48 (9), 2176-2182.

Kim, S.I.; Kim, H.C.; Hyun, C.U. High resolution ocean color products estimation in Fjord of Svalbard, Arctic sea using Landsat-8 OLI. *Korean Journal of Remote Sensing* **2014**, 30(6), 809-816.

http://dx.doi.org/10.7780/kjrs.2014.30.6.11.

Kloiber, S.; Brezonik, P.; Olmanson, L.; Bauer, M. A Procedure for regional lake water clarity assessment using Landsat multispectral data. *Remote Sensing of Environment* **2002**, 82 (1), 38-47. https://doi.org/10.1016/S0034-4257(02)00022-6.

Kontopoulou, E.; Kolokoussis, P.; Karantzalos, K. Water quality estimation in greek lakes from Landsat 8 multispectral satellite data. Special Issue: 10th Word Congress on Water Resources and Environment (EWRA2017 - Issue II). European Water 2017, 58, 191-196.

Le, C.; Li, Y.; Zha, Y.; Wang, Q.; Zhang, H.; Yin, B. Remote sensing of phycocyanin pigment in highly turbid inland waters in lake Taihu, China. *International Journal of Remote Sensing* **2011**, 32 (23), 8253–8269. https://doi.org/10.1080/01431161.2010.533210.

Lim, J.; Choi, M. Assessment of water quality based on Landsat 8 operational land imager associated with human activities in Korea. *Environmental Monitoring and Assessment* **2015**, 187, 1–17. https://doi.org/10.1007/s10661-015-4616-1.

Lin, S.; Novitski, L.N.; Qi, J.; Stevenson, R.J. Landsat TM/ETM+ and machine-learning algorithms for limnological studies and algal bloom management of inland lakes. Journal of Applied Remote Sensing 2018, 12 (02), 1. https://doi.org/10.1117/1.JRS.12.026003

Liu, H.Q.; Huete, A.R. A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. IEEE Trans. Geosci. Remote Sens. 1995, 33, 457–465.

Loveland, T.; Dwyer J. Landsat: building a strong future. *Remote Sensing of Environment* **2012**, 122, 22–29. https://doi.org/10.1016/j.rse.2011.09.022.

MacKay, H.; Finlayson, C.M.; Fernández-Prieto, D.; Davidson, N.; Pritchard, D.; Rebelo, L.M. The role of earth observation (EO) technologies in supporting implementation of the Ramsar convention on wetlands. *Journal of Environmental Management* **2009**, 90 (7), 2234-2242. http://dx.doi.org/10.1016/j.jenvman.2008.01.019.

Markogianni, V.; Dimitriou, E.; Karaouzas, I. Water quality monitoring and assessment of an urban mediterranean lake facilitated by remote sensing applications. *Environmental Monitoring and Assessment* **2014**, 186, 5009–5026. https://doi.org/10.1007/s10661-014-3755-0.

Markogianni, V.; Kalivas, D.; Petropoulos, G.; Dimitriou, E. an appraisal of the potential of Landsat 8 in estimating chlorophyll-a, ammonium concentrations and other water quality indicators. *Remote Sensing* **2018**, 10 (7), 1018. https://doi.org/10.3390/rs10071018.

Matthews, M. A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters. *International Journal of Remote Sensing* **2011,** 32 (21), 6855–6899. https://doi.org/10.1080/01431161.2010.512947.

Matthews, M.W.; Bernard, S.; Robertson, L. An algorithm for detecting trophic status (chlorophyll-a), cyanobacterial-dominance, surface scums and foating vegetation in

inland and coastal waters. *Remote Sensing of Environment* **2012**; 124, 637–652. https://doi.org/10.1016/j.rse.2012.05.032.

Mavromati, E.; Kagalou, I.; Kemitzoglou, D., Apostolakis, A.; Tsiaoussi, V. Linkages between physicochemical status and hydromorphology in greek lakes under WFD policy. European Water Resources Association. Special Issue: 10th Word Congress on Water Resources and Environment (EWRA2017 - Issue II). European Water **2017**, 58, 273-279.

Mavromati, E.; Kagalou, I.; Kemitzoglou, D.; Apostolakis, A.; Seferlis, M.; Tsiaoussi, V. Relationships among land use patterns, hydromorphological features and physicochemical parameters of surface waters: WFD lake monitoring in Greece. *Environmental Processes* **2018**, 5 (S1): 139-151. Springer Science and Business Media LLC. doi:10.1007/s40710-018-0315-6.

McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. Int. J. Remote Sens. **1996**, 17, 1425–1432.

McKinna, L.; Werdell, P. Approach for identifying optically shallow pixels when processing ocean-color imagery. *Optics Express* **2018**, 26 (22): A915-A928. https://doi.org/10.1364/oe.26.00a915.

Nas, B.; Karabork, H.; Ekercin, S.; Berktay, A. Assessing water quality in the Beysehir lake (Turkey) by the application of GIS, geostatistics and remote sensing. In Proceedings of the 12th World Lake Conference, Taal 2007, Jaipur, India, 29 October–2 November 2007. p. 646.

Olmanson, L.; Brezonik, P.; Finlay, J.; Bauer, M. Comparison of Landsat 8 and Landsat 7 for regional measurements of cdom and water clarity in lakes. *Remote Sensing of Environment* **2016**, 185, 119–128. https://doi.org/10.1016/j.rse.2016.01.007.

Olmanson, L.G.; Bauer, M.E.; Brezonik, P.L. A 20-year Landsat water clarity census of Minnesota's 10,000 lakes. *Remote Sensing of Environment* **2008**, 112 (11), 4086–4097. https://doi.org/10.1016/j.rse.2007.12.013.

Ozemi, S. L.; Bauer, M.E. Satellite remote sensing of wetlands. *Wetlands Ecology and Management* **2002**, 10: 381-402. https://doi.org/10.1023/A:1020908432489.

Peel, M.C.; Finlayson, B.L.; McMahon, T.A. Updated world map of the Köppen-Geiger climate classification. *Hydrology and Earth System Sciences* **2007**, 11, 1633-1644. https://doi.org/10.5194/hess-11-1633-2007.

Peppa, M.; Vasilakos, C.; Kavroudakis, D. Eutrophication monitoring for lake Pamvotis, Greece, using sentinel-2 data. *ISPRS International Journal of Geo-information* **2020**, 9 (3), 143. https://doi.org/10.3390/ijgi9030143.

QGIS Development Team. QGIS Geographic Information System. Open source geospatial foundation project. **2019**, http://qgis.osgeo.org/. Accessed on 01/06/2019.

Ramsar Convention Bureau. Wetlands values and functions; Ramsar Convention Bureau: Gland, Switzerland, 2001.

Ritchie, J.C.; Cooper, C.M.; Schiebe, F.R. The relationship of MSS and TM digital data with suspended sediments, chlorophyll, and temperature in Moon lake, Mississippi. *Remote Sensing of Environment* **1990**. 33 (2), 137–148. https://doi.org/10.1016/0034-4257(90)90039-O.

Ritchie, J.C.; Zimba, P.V.; Everitt, J.H. Remote sensing techniques to assess water quality. *Photogrammetric Engineering and Remote Sensing* **2003**, 69 (6), 695–704. doi: 10.14358/PERS.69.6.695.

Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring Vegetation Systems in the Great Plains with ERTS; NASA SP-351; NASA:Washington DC, USA, **1974**; Volume 1, pp. 309–317.

Sakamoto, M. Primary production by phytoplankton community in some japanese lakes and its dependence on lake depth. Archiv für Hydrobiologie 1966, 62, 1-28. http://dx.doi.org/10.1371/journal.pone.0095757.

Schindler, W. David; Hecky, E. R.; Findlay, L. D.; Stainton, P. M.; Parker, R. B.; Paterson, J. M.; Beaty, G. K.; Lyng, M.; Kasian, S. E. M. Eutrophication of lakes cannot be controlled by reducing nitrogen input: results of a 37-year whole-ecosystem experiment. Proceedings of the National Academy of Sciences of the United States of America. August 12 2008, 105 (32), 11254-11258. https://doi.org/10.1073/pnas.0805108105.

Tebbs, E.; Remedios, J.; Harper, D. Remote sensing of chlorophyll-a as a measure of cyanobacterial biomass in lake Bogoria, a hypertrophic, saline–alkaline, flamingo lake, using Landsat ETM+. *Remote Sensing of Environment* **2013**, 135, 92–106. https://doi.org/10.1016/j.rse.2013.03.024.

Topp, S.; Pavelsky, T.; Jensen, D.; Simard, M.; Ross, M. Research trends in the use of remote sensing for inland water quality science: moving towards multidisciplinary applications. *Water* **2020**, 12 (1), 169. https://doi.org/10.3390/w12010169.

Xu, H. Modification of Normalized Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery. Int. J. Remote Sens. **2006**, 27, 3025–3033.

Zacharias, I.; Bertachas, I.; Skoulikidis, N.; Koussouris, T. Greek lakes: Limnological overview. *Lakes and Reservoirs: Research and Management* **2002**, 7(1), 55-62. https://doi.org/10.1046/j.1440-1770.2002.00171.x.

Zhang, C.; Han, M. Mapping chlorophyll-a concentration in Laizhou bay using Landsat 8 OLI data. In Proceedings of the 36th IAHR World Congress, The Hague, The Netherlands, 28 June–3 July 2015.

2.2 Modelling of Greek Lakes Water Quality Using Earth Observation in the Framework of the Water Framework Directive (WFD)

Published as: **Markogianni, V.**; Kalivas, D.; Petropoulos, G.P.; Dimitriou, E. Modelling of Greek Lakes Water Quality Using Earth Observation in the Framework of the Water Framework Directive (WFD). Remote Sens. **2022**, 14, 739. https://doi.org/10.3390/rs14030739

Preamble

Given the great importance of lakes in Earth's environment and human life, continuous water quality (WQ) monitoring within the frame of the Water Framework Directive (WFD) is the most crucial aspect for lake management. In this study, Earth Observation (EO) data from Landsat 7 Enhanced Thematic Mapper Plus (ETM^+) and Landsat 8 Operational Land Imager (OLI) sensors have been combined with co-orbital *in-situ* measurements from 50 lakes located in Greece with the main objective of delivering robust WQ assessment models. Subsequently, trophic status classification was conducted herein, calculating Carlson's Trophic State Index (TSI) initially throughout all lakes and then oriented toward natural-only and artificial-only lakes. All in all, the study findings provide important support toward the perpetual WQ monitoring and trophic status prediction of Greek lakes and, by extension, their sustainable management, particularly in cases when ground truth data is limited.

2.2.1 Introduction

Surface freshwater is one of the most essential resources for the terrestrial ecosystem and the predominant source of drinking water on Earth (Whyte et al., 2018). Over the past few decades, climate change and human activities have deteriorated water quality (WQ) (Michalak 2016). Some factors responsible for it include rapid development, as well as changes in land use/land cover (LULC) patterns, industrialization, and urbanization (El-Alem et al., 2012). The close proximity of water reservoirs to settlements may reduce the price of water to consumers. However, it may also prevent the sustainable management of water resources against deteriorating activities and inappropriate disposal of urban sewage generated within drainage basins (Alparslan et al., 2009).

Deterioration of lake systems' WQ has resulted in many lake eutrophication problems; therefore, environmental scientists have tried to monitor, manage, and limit it for more than two decades (Shafique et al., 2003). WQ monitoring is the most crucial aspect for lake management (Bonansea et al., 2018) and particularly includes the monitoring of certain WQ properties through *in-situ* sampling and field work. The aforementioned WQ properties include Chl-*a* concentration, total suspended matter (TSM), Secchi depth (SD), and nutrient concentrations (Moore et al., 2014).

However, conventional WQ measurements and *in-situ* sampling are laborious, costly and time consuming (El-Alem et al., 2012). Moreover, those techniques are characterized by limited ability to provide a synoptic spatiotemporal view of WQ (Giardino et al., 2001; He et al., 2008) since the condition of an entire water body cannot be fully represented. Furthermore, patchy distribution of nutrients, algal blooms, and TSM define those methods as unsuitable for monitoring a large number of water bodies at a regional or national scale (Dekker et al., 1991; Poor 2010).

Recent developments in geoinformation technologies and in particular of Remote Sensing (RS) and Geographical Information Systems (GIS), concerning pollution loads and WQ, offer a number of advantages that practically address the limitations of traditional water sampling (Brivio et al., 2001; Pozdnyakov et al., 2005; Tyler et al., 2006). Among the key advantages of RS is the ability to cover large areas (Chatziantoniou et al., 2017) and to collect spectral information at variable spatial scales and at dramatically lower cost compared to field measurements (Haddad and Harris, 1985).

According to Morel and Gordon (1980), there are three well-documented methodologies to estimate the concentration of WQ elements in inland waters: empirical, semi-empirical, and physical or analytical methodology. Empirical methods attempt to establish relationships between *in-situ* water quality measurements and water leaving radiance measured by the sensor without the precondition of prior understanding of the complex water and light interactions. Those relationships imply effective data improvement but limited transferability (Austin and Petzold, 1981). Moreover, empirical methods incorporate machine learning techniques, which are differentiated by their robust ability to handle complicated non-linear relationships, typical of WQ remote sensing data (Sagan et al., 2020; Topp et al., 2020). Machine learning algorithms include artificial neural networks (ANN), genetic algorithms (GA), support vector machines (SVM), random forest regression trees, and empirical orthogonal functions (Topp et al., 2020). On the other hand, through semi-empirical techniques, spectral and physical knowledge are combined and then correlated to the *in-situ* concentrations. Regarding physical or analytical approaches, the acquisition of certain biogeochemical parameter values (e.g., Chl-*a*, CDOM) is required, as well as inherent and

apparent optical properties, and are based on radiative transfer within the water column. Then, the *in-situ* concentrations are assessed by modeling the reflectance of surface water. Although analytical methods, including fuzzy logic and Principal Component Analysis (PCA), have already been extensively used, empirical and semi-empirical predicting models are still widely utilized (Gholizadeh et al., 2016). Analytical methods' complexity in terms of their theory and calculation difficulties (Gholizadeh et al., 2016) and the non-availability of required detailed spectral information of the optically active water constituents (optical properties, radiometric quantities) have contributed to the maintenance and development of empirical models. This trend is further observed especially in cases where machine learning models are utilized, as most of them reduce overall error and maximize model fit (Topp et al., 2020). However, it should be noted that empirical algorithms are more specific to certain water types, regional or optical (Odermatt et al., 2012).

The classification of waters in Case 1 (oceanic) and Case 2 (coastal regions, rivers, and lakes, refined by (Gordon and Morel, 1983), is characterized by great importance when remote sensing techniques are utilized to monitor their WQ. The distinction between the two cases has some significant effects on the interpretation and modeling of optical data. In particular, according to this classification scheme, the optical properties of Case 1 waters are determined by phytoplankton and co-varying substances, while Chl-*a* is considered a proxy of phytoplankton concentration. This assumption has facilitated the implementation of large-scale optical models and the development of Chl-*a* predicting algorithms for Case 1 waters. On the other hand, single variable models should be abandoned when Case 2 waters should be studied. It is, on the whole, acknowledged that Case 2 waters are more complex than Case 1 concerning their composition and optical properties. Monitoring the WQ of Case 2 waters is a more sophisticated task since phytoplankton, suspended material, yellow substances, and perhaps bottom reflectance vary independently of each other. The main difficulty lies in the fact that the alterations in optical signal and the concentrations of the dissolved constituents are often so small that they hinder the ability to extract reliable information or the optical signal may be affected in a similar way by more than one substance, which results in an inability to discriminate the different materials (Gholizadeh et al., 2016). Moreover, of principal value is the choice of the appropriate wavelengths, as well as their number in a Case 2 adopted algorithm. Hence, given the difficulty that WQ monitoring of Case 2 waters constitutes a multi-variable, non-linear problem, it is more realistic to establish a series of algorithms rather than a single all-purpose one. In this way, more than one algorithm will attempt to capture and solve the problem for all variables and over several and different ranges of concentrations (IOCCG 2000).

In parallel, the Case 1/Case 2 classification can substantially improve remote sensing products when associated with individual optical water types (OWTs). In particular, coastal regions and inland waters are characterized by such optical diversity that any further information about their variability in IOPs and biogeochemical significance would be particularly valuable. Some OWTs can be hypereutrophic waters, turbid waters with high organic content, sediment-laden waters, CDOM-rich waters, or even very clear blue waters. Several hierarchical, partitional, and hybrid clustering techniques have been utilized to further discriminate distinct OWTs within and between Case 1 and Case 2 waters (Spyrakos et al., 2018). After all, a reliable OWT classification optimizes the selection of the finest constituent algorithms when simpler approaches cannot yield reliable results.

Inland waters, and especially lakes, are small water bodies that are not detected by current ocean color satellites, and even though this lack prevents the monitoring and estimation of their WQ components, it has been replenished by the use of Landsat sensors. A recent review of state-of-the-art RS-based approaches by (Gholizadeh et al., 2016) underpins the use of Landsat sensors, TM (Thematic Mapper), MSS (Multi-Spectral Scanner), ETM (Enhanced Thematic Mapper), and OLI (Operational Land Imager) as fairly successful choices to assess the important WQ parameters, including Chl-*a*, SDD, TP, and trophic status (Alparslan et al., 2009; Allan et al., 2011; Giardino et al., 2014; Kim et al., 2014; Markogianni et al., 2014; Markogianni et al., 2018; Markogianni et al., 2020).

RS has been widely demonstrated as an effective solution for detecting the relationship between algae concentration and corresponding nutrients (Hans et al., 2002). Nitrogen (N) and phosphorus (P) are vital micronutrients for algae, while P (existing either in a particulate or dissolved phase) is the key limiting nutrient responsible for eutrophication in most lakes (Correll 1999). In general, special attention should be paid depending on which nutrient is growth limiting, as in one water body the correlation with Chl-*a* might be with N, while in a different water body the correlation might be with P (Sagan et al., 2020). Total phosphorus (TP) estimation via RS has been explored due to its high correlation with optically active constituents (Kutser et al., 1995; Wang et al., 2004; Wu et al., 2010) since it cannot be measured directly using optical RS instruments. The chlorophyll-*a* (Chl-*a*) and TP relationship has been investigated in individual lakes (Smith 1982; Malve and Qian, 2006), and it is well documented to be accompanied by a strong and positive correlation among lakes (Healey and Hendzel 1979; Busse et al., 2006). He et al. (2008) performed routine WQ monitoring on the slightly-polluted Guanting Reservoir in China using Landsat-5 TM and retrieved WQ data with eight variables, namely algae, turbidity, concentrations of chemical oxygen demand (COD), total nitrogen (TN), ammonia nitrogen (NH3–N),

nitrate nitrogen (NO3–N), total phosphorus (TP), and dissolved phosphorus (DP). Their results indicated a statistically significant correlation (10–30% mean relative error) among all estimated parameters and reflectance regression algorithms. Landsat-5 TM data was also used by (Akbar et al., 2010), who predicted TP among other water quality components of different sources across Alberta and managed to classify lakes into four trophic states indicating low to very high productivity. In another study, Song et al. (2006) established both a regression model and an empirical neural network to simulate the relationship between TP and Landsat TM radiances for Chagan Lake, China. As TP is highly correlated to Chl-*a* concentration, and TSM usually reflects TP loading, TP is also closely related to Secchi depth (SD) with an exponential equation according to Carlson's findings (Carlson 1977). Based on the same rationale, (Song et al., 2012) estimated TP empirically through associated Chl-*a*, TSM, and Secchi depth across three reservoirs in Indiana, US, with R^2 values between *in-situ* and spectral data ranging from 0.55 to 0.72.

Water clarity, commonly reflected by SD, is reduced by the increased presence of suspended sediment, organic matter, and zooplankton (Carlson 1977). The stimulating production of algae in a lake usually originates from increased nutrients, in particular, phosphorus (Busse et al., 2006). As the algae and suspended inorganic matter increase in a lake, the depth to which light can penetrate (Fuller et al., 2004) is reduced. Therefore, SD is often used as a trophic state indicator (Carlson 1977). In general, there are two methodologies followed to retrieve SDT based on remote sensing data. Empirical approach estimating SD through regression analysis and semi-analytical approach retrieving SD based on an underwater visibility theory (Jiang et al., 2019). Regarding empirical models, reflectance at the red spectrum has been almost globally used to retrieve water clarity (Baban 1993; Nelson et al., 2003; Wu et al., 2008; McCullough et al., 2012; Hicks et al., 2013) since increased brightness is accompanied by decreased water clarity (Matthews 2011). Moreover, further studies have also documented the usefulness of spectral response of the blue, green, and near-infrared spectral bands in combination with *in-situ* measurements of SD and Chl-*a* concentrations in predicting water clarity for inland lakes (Olmanson et al., 2001; Fuller et al., 2004). It should also be noted that semi-analytical methods are superior to empirical ones mainly concerning the reliability of results and the fact that no *in-situ* data are required afterwards for recalibrating the retrieval algorithm. On the other hand, those approaches require the utilization of a spectroradiometer and the collection of *in-situ* measured Rrs spectra including the radiance of skylight, radiance from a standard gray board, and the total upwelling radiance from the water (Jiang et al., 2019).

Since water clarity has long been proven to interact with nutrient availability and Chl-*a* concentrations within lakes (Carlson 1977; Megard et al., 1980), remote sensing studies frequently use it to assess overall lake

trophic status (oligotrophic, mesotrophic, or eutrophic) (Peckham et al., 2006; Olmanson et al., 2008). WQ monitoring programs (such as WFD) have been implemented worldwide to acquire large datasets of several WQ parameters, while several methods (such as cluster and discriminant analysis) have been efficiently utilized to manage those complex data and interpret the underlying patterns of trophic status. However, these methods need continuous *in-situ* measurements, while the classical and most widely used method to characterize a lake's trophic status is Carlson's Trophic State Index (TSI) (Carlson 1977). This approach includes equations employing Secchi depth, Chl-*a*, and TP measurements (Nauman 1929).

The hereby adopted methodological scheme includes the implementation of stepwise multiple regression (MLR) analyses among *in-situ* measurements and satellite data. I*n-situ* data concern Secchi depths and TP concentrations along 50 lakes, included in the National Lake Monitoring of Greece (WFD), and since the majority of those data were recorded during 2013–2015, images of sensors Landsat 7 ETM+ and Landsat 8 were the exclusive choice for the implementation of this research. According to a previous study conducted by the authors (Markogianni et al., 2020), a principal component analysis (PCA) indicated that the variance of Chl-*a* concentrations of the same lakes was affected by whether the lakes were natural or artificial, while the rest of the tested parameters were the climatic type, WFD typology, and the sampling season. Hence, based on those PCA's results, hereby MLR analyses concerned: (a) all *in-situ* measurements of TP and Secchi depth during 2015–2016 and 2013–2016, respectively, and (b) *in-situ* TP and Secchi depth datasets of the same years, including natural-only and artificial-only lakes. Correlation analyses were additionally conducted to explore and detect the existing interrelationships among TP, Chl-*a* concentrations, and SD of monitored lakes and improve the effectiveness of the WQ assessment models by indicating further significant predictors. Subsequently, Chl-*a* regression models developed by Markogianni et al. (2020), and hereby established TP and Secchi depth's models were utilized to calculate the water trophic index of the studied lakes.

In purview of the above and taking advantage of the large *in-situ* dataset derived from the application of National Lake Monitoring in Greece (WFD), the present study aims to: (1) explore the complicated relationships among TP, Chl-*a* concentrations, and Secchi depth measurements throughout 50 lakes, substantially representing Case 2 waters, (2) generate accurate quantitative TP and Secchi depth models by incorporating satellite images with concurrent *in-situ* measurements, and (3) derive the Carlson Trophic Index for assessing water trophic state spatially over all monitored waterbodies.

2.2.2 Study area

The study area includes 50 lakes, natural and artificial (Figure 2.2.2-1; Table 2.2.2-1). These waterbodies comprise the National Monitoring Network of Waters in Greece, which is implemented by the Goulandris Natural History Museum, Greek Biotope/Wetland Centre (EKBY). More information about the general characteristics of the monitored lakes can be found at the study conducted by Markogianni et al. (2020) or more detailed data can be retrieved from the EKBY's site (Goulandris Natural History Museum, Greek Biotope/Wetland Centre; http://biodiversity-info.gr/index.php/el/lakes-data#!IMGP4731; accessed date 5 February 2020).

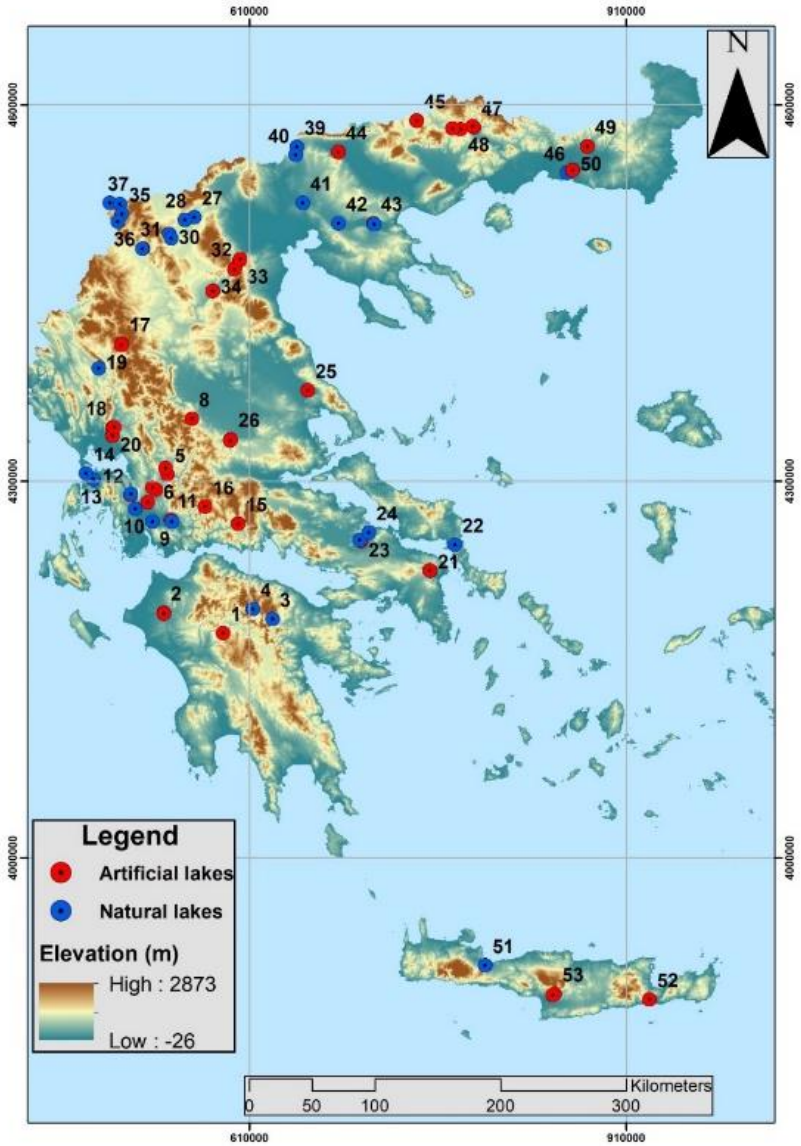


Figure 2.2.2-1. National Lake Monitoring Network in Greece (numbers of sampling stations coincide with the numbers presented in Table 2.2.2-1).

Table 2.2.2-1. Main characteristics of the lakes included in the National Lake Monitoring Network in Greece (WFD) (Mavromati et al., 2018; Markogianni et al., 2020).

No	National Name Station	Surface (km^2)	(N)atural/ (A)rtificial	Mean depth (m)	No	National Name Station	Surface (km^2)	(N)atural/ (A)rtificial	Mean depth (m)
1	Lake Ladona	-	A	-	28	Lake Petron	11.91	N	3.1
2	Lake Pineiou	19.64	A	15.1	29	Lake Zazari	2.98	N	3.95
3	Lake Stymfalia	-	N	1.31	30	Lake Cheimaditida	9.82	N	1.01
4	Lake Feneou	0.47	A	10.5	31	Lake Kastorias	30.87	N	3.7
5	Lake Kremaston	68.43	A	47.2	32	Lake Sfikias	3.96	A	23.2
6	Lake Kastrakiou	25.58	A	33.2	33	Lake Asomaton	2.46	A	20.8
7	Lake Stratou	7.02	A	9.6	34	Lake Polyfytou	63.49	A	22.4
8	Lake Tavropou	21.46	A	15.0	35	Lake Mikri Prespa A	-	N	3.95
9	Lake Lysimacheia	10.87	N	3.5	36	Lake Mikri Prespa B		N	-
10	Lake Ozeros	10.57	N	3.8	37	Lake Megali Prespa A	-	N	17
11	Lake Trichonida	93.53	N	29.6	38	Lake Megali Prespa B		N	-
12	Lake Amvrakia	13.14	N	23.4	39	Lake Doirani 1	33.25	N	4.6
13	Lake Voulkaria	7.38	N	0.96	40	Lake Doirani 2		N	-
14	Lake Saltini	-	N	-	41	Lake Pikrolimni	6.30	N	1.2
15	Lake Mornou	17.50	A	38.5	42	Lake Koroneia	-	N	3.8
16	Lake Evinou	2.68	A	31.5	43	Lake Volvi	70.36	N	12.3
17	Lake Pigon Aoou	11.44	A	20.8	44	Lake Kerkini	-	A	2.19
18	Lake Pournariou	19.28	A	29.8	45	Lake Leukogeion	0.83	A	4.05
19	Lake Pamvotida	21.82	N	5.3	46	Lake Ismarida	-	N	0.9
20	Lake Pournariou II	0.56	A	11.7	47	Lake Platanovrysis	2.99	A	26.4
21	Lake Marathona	2.17	A	15.8	48	Lake Thisavrou	13.43	A	38.4
22	Lake Dystos	-	N	-	49	Lake Gratinis	0.80	A	14.2
23	Lake Yliki	19.96	N	20.1	50	Lake N. Adrianis	-	A	-
24	Lake Paralimni	9.96	N	2.99	51	Lake Kourna	-	N	15
25	Lake Karlas	-	A	0.9	52	Lake Bramianou	-	A	10.1
26	Lake Smokovou	-	A	-	53	Lake Faneromenis	0.33	A	9.98
27	Lake Vegoritida	47.67	N	26.52					

2.2.3 Materials and Methods

2.2.3.1 Data acquisition

2.2.3.1.1 In-situ data

Data used in this study were collected in the framework of the Greek Water Monitoring Network for lakes (WFD). All data is freely accessible and was downloaded from the EKBY's site (Goulandris Natural History Museum, Greek Biotope/Wetland Centre (http://biodiversity-info.gr/index.php/el/lakes-data#!IMGP4731; in Greek). The network incorporates 50 lakes, natural and reservoirs. At the majority of the lakes, only one sampling station is detected, except for trans-boundary lakes (Megali Prespa, Mikri Prespa, and Doirani), where two sampling stations are located (Table 2.2.2-1; Figure 2.2.2-1). From the total of 53 sampling sites, there are 27 surveillance and 26 operational ones. Surveillance stations operate in water bodies of good status, for a certain period of time (one year in every monitoring cycle), while operational stations are monitored on a monthly or seasonal basis, in water bodies which fail to achieve good status (Markogianni et al., 2020). The selected data used herein includes the Secchi depth measurements in several dates from 2013 up to 2018 and TP concentrations from 2015 up to 2018 throughout the monitored lake stations. Secchi depth measurements were conducted with a Secchi disk, measuring the transparency of water while *in-situ* Chl-*a* data was already available in the framework of our last study (Markogianni et al., 2020).

Particularly, Chl-*a* concentrations were measured from 2013 to 2018 and determined spectrophotometrically (Method 10200 H; APHA 1989). TP concentrations include all inorganic, organic and dissolved forms of phosphorus and the available dataset incorporates measurements analyzed during the years 2015, 2016 and 2018. During years 2013 & 2014 (i.e. since the beginning of the WFD), analysis of orthophosphates resulted in low concentrations, lower than the quantitation limit (LOQ) of the respective adopted method, hence no measurement was available during this period. Therefore, the following years (i.e. 2015, 2016, 2018) analyses of orthophosphates were replaced by Total phosphorus ones, which resulted in the acquirement of actual measurements during this period.

Further investigation of *in-situ* data included a seasonal statistical analysis by incorporating dates of same season of all lakes during the monitored years. The seasons were determined as: summer (June, July, and August), autumn (September, October and November), winter (December, January and February) and spring (March, April, and May) while more information about the sampling periods, sampling and analysis methodologies can also be found in EKVY' site.

Exploratory statistics among the Secchi depth measurements of 2013, 2014, 2015, 2016 and 2018 and TP concentrations of 2015, 2016 and 2018 were

calculated incorporating the estimation of mean, median, standard deviation and min-max. Skewness, Kurtosis and the Kolmogorov-Smirnov and Shapiro-Wilk tests were conducted to explore the data normality. Furthermore, SPSS Statistical Package (v. 24.0) was used to group and categorize the under studied WQ parameters based on the sampling's season, year and whether the lakes are natural or artificial. Moreover, correlation matrix among simultaneous *in-situ* measurements of TP, Chl-*a* and Secchi depths was conducted to explore their existent interrelationships and further contribute to indicating the most significant predictors.

2.2.3.2 Exploratory Statistical Analyses

Secchi depths throughout the monitored Greek lakes were measured during the years 2013, 2014, 2015, 2016 and 2018 (Table 2.2.3-1). Minimum values ranged from 0.03 (2014, 2015) to 0.2 m (2013) while maximum ones from 11 (2015) up to 15.5 m (2018). Mean values of Secchi depth are similar during all years and equal to around 3.2 m. Secchi depths are presented higher in artificial than in natural lakes while the highest values are observed during summer months for both natural and artificial lakes (Figure 2.2.3-1a). The temporal distribution of Secchi depths was categorized on the criterion of whether the lakes are artificial or natural; values are also higher in artificial lakes during all sampling years with some exceptions (e.g. Trichonida Lake; Figure 2.2.3-1b).

Table 2.2.3-1. Summary of descriptive statistics of *in-situ* Secchi depth values during years 2013-2016 and 2018.

Secchi depth (m) in Year:	N	Min	Max	Mean	Std. Deviation	Skewness	Kurtosis
2013	134	.20	14.0	3.1	2.8	1.5	3.3
2014	125	.030	14.0	3.8	3.1	.9	.2
2015	140	.030	11.0	3.2	2.6	.8	-.2
2016	64	.050	15.0	3.03	3.2	1.7	3.1
2018	314	.100	15.5	3.04	2.7	1.4	2.4
all years	777	.03	15.5	3.2	2.8	1.3	1.7

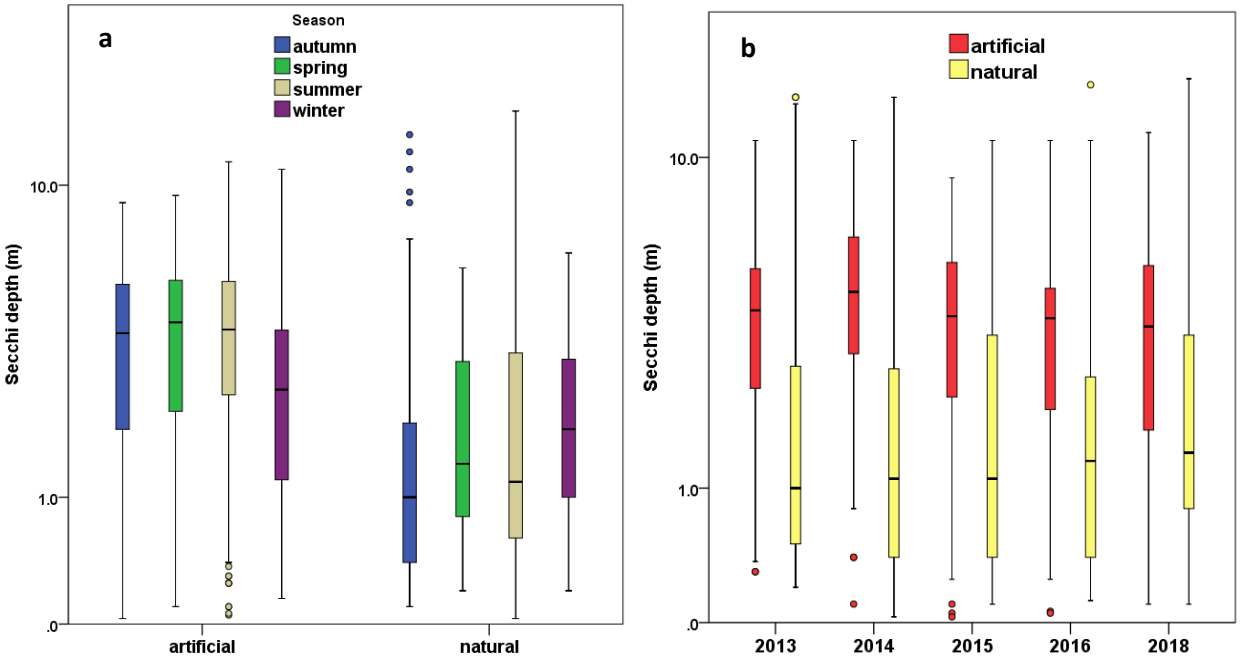


Figure 2.2.3.2-1. Boxplots presenting basic statistics of Secchi depths (a) grouped by the lake's nature and categorized by the sampling season, and (b) grouped by sampling year and categorized by the lake's nature.

Total measurements of TP concentrations are 370, including years 2015, 2016 and 2018 (Table 2.2.3-2). Minimum TP values are similar during all years (around 0.01 mg/l) while maximum values are increasing during the years. Same tendency is reflected based on average values with the mean TP value of 2018 to be double compared to the respective value of 2016. Higher TP concentrations are detected in natural lakes, particularly during autumn sampling months while water sampling analysis in summer revealed the highest TP concentrations in artificial lakes (Figure 2.2.3-2a). As far as the yearly distribution of TP concentrations in Greek lakes is concerned, it is confirmed that natural lakes are more affected by TP pollution sources than the artificial ones with an increasing tendency throughout the years (Figure 2.2.3-2b).

Table 2.2.3-2. Summary of descriptive statistics of *in-situ* TP concentrations during years 2015-2016 and 2018.

Total phosphorus (mg/l) in Year:	N	Min	Max	Mean	Std. Deviation	Skewness	Kurtosis
2015	169	.01	4.2	.14	.56	6.7	45.4
2016	69	.02	5.1	.23	.8	5.5	29.9
2018	132	.02	13.3	.48	1.8	4.9	25.98
all years	370	.01	13.3	.28	1.2	6.9	54.2

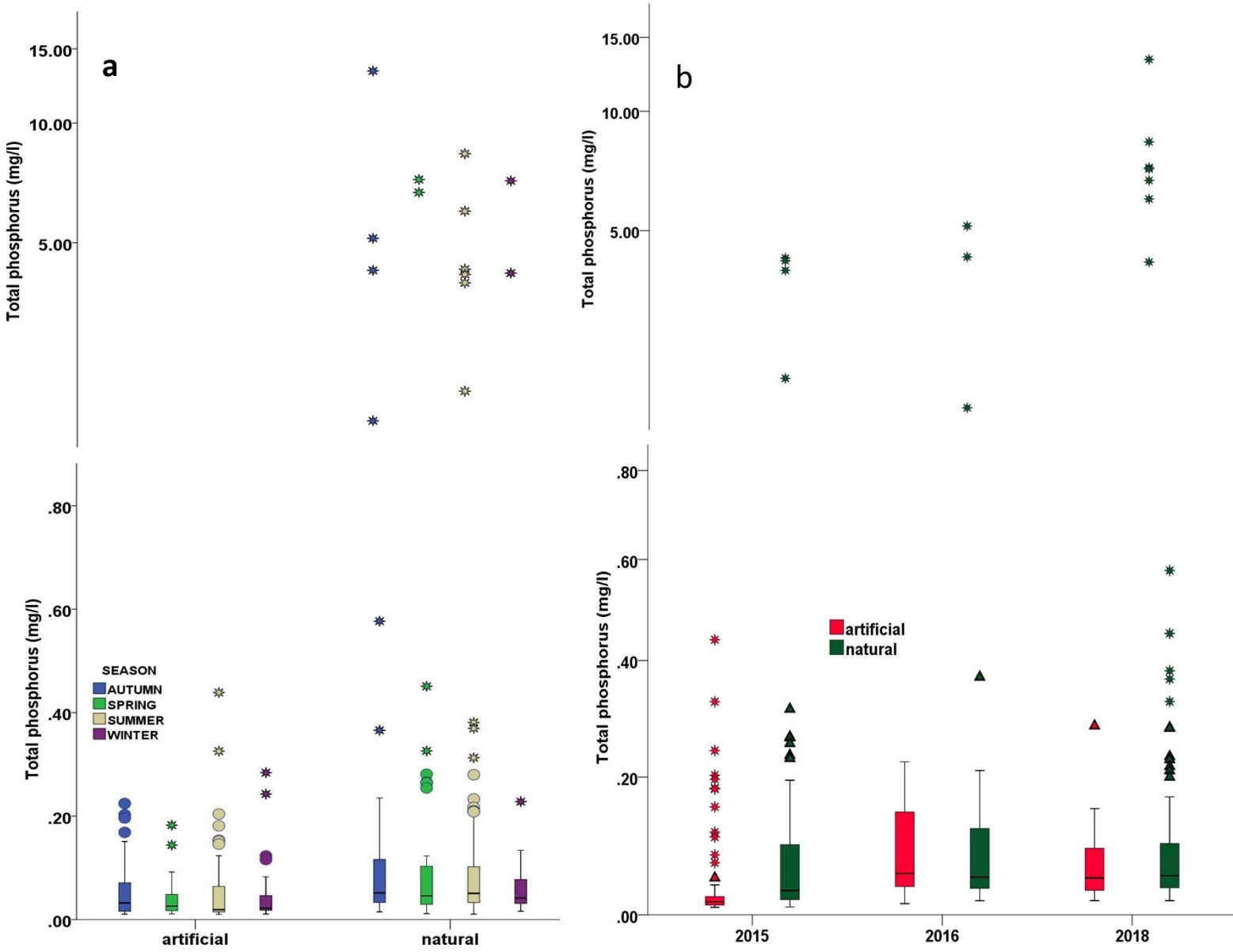


Figure 2.2.3-2. Boxplots presenting basic statistics of TP concentrations (a) grouped by the lake's nature and categorized by the sampling season, and (b) grouped by sampling year and categorized by the lake's nature (top and bottom panels illustrate the range of *in-situ* TP measurements in a logarithmic scale divided in 0-0.8 mg/l and 1-15 mg/l extents, respectively).

2.2.3.3 EO Data Acquisition & Pre-Processing

Landsat 8 OLI and Landsat 7 ETM+ images used herein covered the 50 monitored lakes throughout Greece. These data had been previously acquired in the framework of our previous study (Markogianni et al., 2020). In particular, a 2013–2016 and 2018 time series of 296 Landsat images— with a

mean time window between the satellite overpass and the *in-situ* measurements equal to 4 days—were downloaded from the USGS (United States Geological Survey) Data Centre (https://earthexplorer.usgs.gov/ (accessed date 5 February 2020)) for Chl-*a* estimations. More specifically, total *in-situ* Chl-*a* data include 702 measurements, and the time window between sampling and satellite dates ranges from -21 to 17 days. Moreover, since not all monitored WQ parameters were sampled on simultaneous dates, Secchi depth data were aligned with a total of 304 images (2013–2018) and the TP concentrations with 122 images (2015–2018), including some newly downloaded extra images. Secchi depth measurements are equal to 578, and the time window difference ranges between -16 to 19 days with a mean time gap of approximately 4 days.

As far as the TP measurements are concerned, 268 total values were recorded during the years 2015, 2016, and 2018, accompanied by satellite images with overpass dates ranging from 21 to 14 days before and after the field work, respectively and the mean time gap is equal to 4 days. Moreover, it should be noted that the majority of the satellite images have been used for the monitoring of more than one of the studied WQ parameters, and the statistical analysis eventually included those that met certain criteria (e.g., images that portrayed lakes with mean depth higher than 5 m; images of dates coincident with sampling dates of all the three parameters Chl-*a*; TP; Secchi depth for the TSI calculation, etc.).

Concerning the great time window between sampling and satellite dates in some cases, it should be noted that only a few images are temporally far from the field work's date. It has been proven that a time-window up to ±7 days yields reasonable results and is not considered a problem when lake water quality, especially in non-tidal systems, is monitored (Kloiber et al., 2002; Hellweger et al., 2004; Chu et al., 2018). Therefore, concerning the Chl-*a* training dataset (general model), only 15.4% of records surpassed the ±7 days' time gap. The respective percentages for Secchi depth and TP training datasets are 13.2% and 15%, while 50% and 71% of those records, respectively, constitute artificial lakes that have been separately elaborated in a restricted time gap of ±5 days. The percentages are similarly low concerning the development of WQ models for natural lakes. The Chl-*a* natural model was developed by employing 12 out of 85 records (14.1%) with a date difference higher than ±7 days from the satellite overpass, while the Secchi natural model included 5 out of 65 (7.7%) records characterized by the same time window. As far as the TP natural model is concerned, only 2 out of 29 measurements have been aligned with images acquired at dates greater than ±7 days from the sampling date. Given the low percentage rates of those records utilized in the development of WQ models, it is assumed that their effect is insignificant on the models' performance and prediction accuracy.

The pre-processing steps that were adopted herein are identical to the ones described in our earlier study (Markogianni et al., 2020). More particularly, semi-automatic classification plugin (SCP) of the free and open-source cross-platform desktop geographic information system Q-GIS v. 3.6.3-Noosa was employed to perform: (a) conversion of images from digital numbers (DN) to top-of-atmosphere reflectance (TOA), (b) atmospheric correction by using the DOS1 method (applied to all bands except for thermal ones), and (c) the creation of a band stack set for each image. The band stack set of L7 ETM+ includes bands B1 (blue), B2 (green), B3 (red), B4 (NIR), B5 (SWIR1), and B7 (SWIR2), while L8 incorporates bands B2 (blue), B3 (green), B4 (red), B5 (NIR), B6 (SWIR1), and B7 (SWIR2).

Since 2003, sensor ETM+ has acquired and delivered data with gaps caused by Scan Line Corrector (SLC) failure. In order to retrieve the data that concurred with those line gaps, several calculations were conducted by employing focal statistics through ArcMap. Those line gaps are approximately 205 m in length on the vertical axis, and in combination with the spatial resolution of the Landsat sensor (30 m), the mean value within a circle of 7 cells was determined among several trials as the most optimal neighborhood to include the coincident sampling station everywhere within this line. Through the focal statistics tool, an output raster (focal raster) for each input one (satellite band) was calculated, and then the Con and IsNull functions were applied (Equation (2.2.3-1)) in order only the no-values cells to be replaced while the rest preserved their values.

Con (IsNull(Satellite band with gaps), (focalRaster), (Satellite band with gaps)) (2.2.3-1)

The implementation of the DOS1 atmospheric correction method was not validated in order to ensure that atmosphere biases were completely removed. However, this method is widely used by the EO community (Barrett and Frazier 2016; Japitana and Burce 2019) and proved useful when no atmospheric measurements are available and correcting historical imagery. In the framework of the effort of (Doña et al., 2014) to develop WQ empirical algorithms across certain Spanish lakes and ponds, they evaluated three different atmospheric correction methods (DOS; ATCOR3; MODTRAN5). Those methods were applied to Landsat 7 ETM+ bands, and the results indicated that the DOS method performed better than the others, reporting the lowest errors.

Moreover, to further ensure the use of only cloud-free pixels over the sampled lakes, the Cloud Masking QGIS plugin (https://smbyc.github.io/CloudMasking; accessed date 10 March 2020) was used. By using this tool, clouds, cloud shadow, cirrus, aerosols, and ice/snow

were masked for all Landsat images using the combination of the Fmask and Blue Band processes.

2.2.3.4 Statistical approach

2.2.3.4.1 Establishment of relationships between Landsat data, Secchi depths and TP

The hereby available *in-situ* data include Secchi depth and TP lake measurements, recorded in the framework of the WFD application in Greece during the years 2013-2016 and 2018. Especially *in-situ* data of 2018 was used as an independent validation dataset for both of the WQ elements. Visible (blue, green and red), NIR and SWIR spectral bands, combined with their ratios, additions, subtractions, ln- and log-transformations were employed in multiple combinations including also transformations from the respective scientific literature (Table 2.2.3-3) with basic aim to explore and develop statistically significant relationships between them and *in-situ* Secchi depths and TP measurements of coincident dates. Figure 2.2.3-3 illustrates the discrete methodological steps followed herein regarding the two *in-situ* datasets indicated by numbers (1) and (2) for Secchi depths and TP values, respectively.

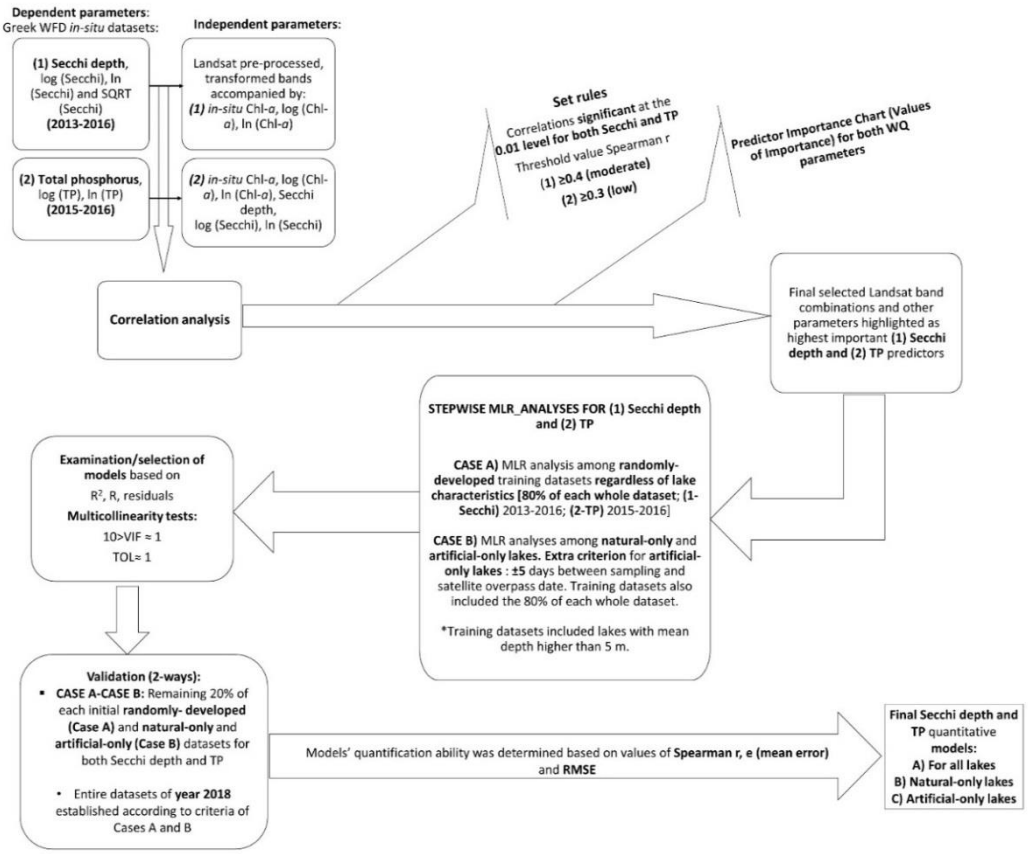


Figure 2.2.3-3. Flow diagram describing the methodology followed regarding the WQ models' establishment and validation.

As far as the Secchi depth dataset (1) is concerned, a correlation analysis between several band transformations and Secchi depths' measurements, as well as log, ln and SQRT Secchi depths was conducted, including previously published band combinations (Table 2.2.3-3). In those independent variables, *in-situ* Chl-*a* was also included as this parameter has been previously proved to affect lake water clarity (Song et al., 2012). Correlation analysis results and the selection of certain significant predictors of Secchi depth was determined based on specific rules. Setting initially a threshold value of the significant correlations at the 0.01 level and a Spearman value equal or higher than ±0.4 (which indicates moderate relationship according to Dancey and Reidy (2007), resulted in the distinction of the initial wide group of Secchi depth's predictors. Furthermore, those predictors were also enriched and confirmed based on the results of the predictor importance chart (IBM SPSS software Statistics v. 23.0, Armonk, NY, USA). This chart indicates the relative importance of each predictor in estimating a model while the predictor importance relates to the importance of each predictor in making a prediction, not whether or not the prediction is accurate.

Additional criteria including multi-collinearity and values of tolerance factor, variance inflation factor (VIF) and R^2 were also applied to explore statistical performance and residuals and resulted in a subset of the initial predictors. According to Markogianni et al. (2020), a factor analysis was implemented to obtain an indication of underlying common factors (components) that explain the interrelationships among Chl-*a* concentrations, lake nature (natural/artificial), sampling season and climatic type. The rotated component matrix results indicated that the lake characteristics (natural/artificial), followed by the sampling season were the variables that mostly affect the variance of Chl-*a* concentrations in the same -as in this work- studied lakes during the same period (2013-2018).

In the effort of the authors (Markogianni et al., 2020) to further enhance the efficiency of Chl-*a* regression models, a confined time window of ±5 days between field measurements and satellite overpass was used, in cases with statistically significant results. Those results were indeed further improved when artificial lakes were the case, regardless of the sampling season. Hence, since a) the herein research concerns the same lakes being monitored during the same period and b) the ultimate goal is the assessment of their trophic status, it was decided to conduct MLR analyses based on the same rationale as in (Markogianni et al., 2020). Consequently, the two basic scenarios employed, concern: Case A) MLR analysis among attributes originating from a randomly- developed training dataset.

Total Secchi depth measurements were divided in training and validation datasets including 80% (228 out of 286 Secchi depth measurements) and 20% of the entire dataset, respectively. This analysis constitutes an effort

to develop a Secchi depth quantitative model for lakes when no information is available (e.g. regardless the sampling season, natural/artificial etc.), and, Case B) MLRs analysis focused separately on attributes concerning natural-only or artificial-only lakes, with the latter being accompanied by a time window of ±5 days between sampling and satellite date as proposed by (Markogianni et al., 2020). Furthermore, the addition of Chl-*a* values in the possible Secchi depth's predictors had as a result to further shorten the initial total *in-situ* dataset, as only the records of dates characterized by simultaneous sampling of Secchi depths and Chl-*a* were included in the analysis.

The same methodology was also adopted in the TP concentrations (**2** in Figure 2.2.3-3). *In-situ* dataset of TP is narrower than the one concerning Secchi depths, as it includes only values sampled during the years 2015 and 2016 (dataset of 2018 was utilized as an independent validation dataset). Correlation analysis was also conducted among satellite band transformations (Table 2.2.3-3) and *in-situ* TP values while *in-situ* Chl-*a* concentrations, Secchi depths and their logarithmic transformations were also included since they have been proven to interact and affect TP concentrations in lakes (Kutser et al., 1995; Wang et al., 2004; Wu et al., 2010). Due to fewer TP available measurements, the threshold value of Spearman r was set to 0.3 to avoid the loss of possible significant TP predictors and the proposed TP predictors were also confirmed by the significance predictor chart. Hence, according to Figure 2.2.3-3, multiple datasets with simultaneous measurements of all the three parameters or combinations of them were established and constituted the randomly made datasets (Case A). Those multiple datasets were further divided in training (80% of each total records) and validation ones (the rest 20%). Then, concerning MLRs analyses of B case (Figure 2.2.3-3), the aforementioned datasets were further divided in cases including natural-only and artificial-only lakes (with data acquisition time window of ±5 days), while they were additionally separated in training (80%) and validation (20%) datasets, respectively.

Training datasets regarding the *in-situ* Secchi depths and TP include measurements of 2013-2016 and 2015-2016, respectively. It should also be noted that training and validation datasets contained lakes with mean depth higher than 5 m to surely avoid the bottom reflectance noise (McKinna and Werdell, 2018). In particular McKinna and Werdell (2018) recommended that any pixel with a water-column depth of 5 m or less should be characterized as optically shallow and omitted from the analysis in order to avoid any unwanted optically shallow effects apparent in satellite -derived products; WQ models and their efficiency in our case. Final results of the MLRs should be the development of rigorous quantitative algorithms regarding: a) Secchi depth and TP for all the lakes, b) Secchi depth and TP for natural lakes, c) Secchi depth and TP for artificial lakes.

Table 2.2.3-3. Published band combinations utilized in remotely estimating TP and Secchi depth values.

Reference	Parameters	Band combinations and sensors
Lim and Choi, 2015	TP	Blue, Green, Red, NIR, NIR/Green (L8)
Song et al., 2006; Song et al., 2012	TP	Blue, Green, Red, and NIR (L5)
Wu et al., 2010	Ln (TP)	Blue, Red/Green, Blue/Red (L5)
Alparslan et al., 2009	TP	Blue, Green, Red, NIR, SWIR1 and SWIR2 (L5)
Baban 1993	(1) TP (2) Secchi depth	(1) Red, Green, Red/Blue, (Green + Red)/2, Green^2, (Blue + Green)/2 (L5) (2) Red/Blue, Red^2, Blue, (Blue+Green)/2, (Blue + Red)/2 (L5)
Isenstein and Park, 2014	(1) SQRT (TP) (2) Secchi depth	(1) Red, SWIR2 (L7 ETM+) (2) LOGRed, LOGSWIR2 (L7 ETM+)
Chen and Quan, 2012	Phosphorus	Blue, Green, Red, NIR (L5)
Huang et al., 2015	LOG (P)	NIR/Visible light (GOCI)
Moses et al., 2014	(1) Phosphates (2) TP	(1) Red, MIR (2) Red IRS P6 (LISS III)
Shafique et al., 2003	TP	LOG (Green/Red to NIR), (CASI)
Allan et al., 2011	(1) Secchi depth (m) (2) LN Secchi depth	(1) Blue/Red, (Blue-Red)/Green, LN [(Blue-Red)/Green] (L7 ETM+) (2) NIR, (Blue-Red)/Green, LN Red
Brezonik et al., 2005	LN Secchi depth	Blue, Blue/Red (L5)
Choubey 1998	Secchi depth	Blue, Green, Red (IRS-1A)
Zhou et al., 2021	Secchi Depth	Green, Red, Blue, Vegetation red edge (B5), Water Vapour (Sentinel 2)
Ohammad and Alsahli, 2021	Secchi depth	Green, Blue (MODIS-Aqua)
Kratzer et al., 2019	Secchi depth	Blue, Red (MERIS)

2.2.3.5 Validation approach

WQ quantitative models were validated in two ways. The basic statistical metric selected to verify efficiency is the Spearman's (r) correlation coefficient which was selected based on the Kolmogorov-Smirnov and Shapiro-Wilk tests of normality. Additionally, the mean error (e) and the Root-Mean-Square Error (RMSE) indices were also applied. Initially, each

validation dataset, including the 20% of the total values during years 2013-2016 for Secchi depths and 2015-2016 for TP, respectively, constituted the first validation process (the rest 80% were used as training datasets). Then those values were linked with the respective images in order to acquire the predicted parameters' values and further assure the good performance of the selected models. The second validation process included the utilization of the independent *in-situ* datasets sampled during the year 2018 (Figure 2.2.3-3).

2.2.3.6 Carlson's Trophic State Index (TSI) and validation

Carlson's Trophic State Index (TSI) is the most widely used tool for characterizing a lake's health or its trophic state while the latter is defined as the biological reaction of water bodies to nutrient additions (Nauman 1929). Carlson's method (Carlson 1977) uses Secchi depth in meter, a logarithmic transformation (Ln) of chlorophyll-*a* concentration in microgram per liter, and total phosphorus measurements in microgram per liter while it concerns an index represented as a numerical scale to categorize lakes into classes related to their trophic status.

Equations (2.2.3-2), (2.2.3-3) and (2.2.3-4), derived from Carlson (1977), have been widely used to compute the TSIs according to TP, Chl-*a* and SD, respectively, while an average (Equation 2.2.3-5) is estimated to produce the final trophic state as follows:

$$TSI(TP) = 10 * [6 - \frac{LN\left(\frac{48}{TP}\right)}{LN(2)}] \quad (2.2.3\text{-}2)$$

$$TSI(Chla) = 10 * [6 - (2.04 - \left(0.68 * \frac{\ln(Chla)}{LN(2)}\right))] \quad (2.2.3\text{-}3)$$

$$TSI(SDT) = 10 * [6 - \left(\frac{LN(SDT)}{LN(2)}\right)] \quad (2.2.3\text{-}4)$$

$$TSI(average) = [TSI(TP) + TSI(Chla) + TSI(SDT)]/3 \quad (2.2.3\text{-}5)$$

The trophic status classification system categorizes lakes as oligotrophic (TSI value<30), mesotrophic (TSI value 40–50), eutrophic (TSI value 60-70), and hypereutrophic (TSI value>70; Table 2.2.3-4) and since the scale of the index is arithmetic, it can describe trophic changes and a larger number of transitional individual lake classes (e.g. oligotrophic-mesotrophic, mesotrophic-eutrophic).

Table 2.2.3-4. Carlson's trophic state index values and classification of lakes (Carlson 1977; Prasad and Siddaraju, 2012).

TSI values		Trophic Status	Attributes
< 40	< 30	Oligotrophic	Transparent water
	30-40	Oligotrophic-Mesotrophic	
41-50	41-48	Mesotrophic	Higher turbidity, higher algae abundance and macrophytes
	49-50	Mesotrophic-Eutrophic	
51-70	51-60	Mesotrophic-Eutrophic	
	61-70	Eutrophic	Usually blue-green algae blooms
>70		Hypereutrophic	Extreme blue-green algae blooms

Based on these equations, the *in-situ* TSI for all the cases accompanied by available simultaneous *in-situ* measurements of TP, Chl-*a* and Secchi depths were also calculated. In the framework of the study conducted by Markogianni et al. (2020), through the harmonization of Landsat 7 ETM+ and 8 OLI images, three Chl-*a* quantitative models were developed including the ratios of blue to green and red, red to green and blue, and the ln-transformed bands SWIR1 and SWIR2. Those models were established based on the same period and same lakes as the ones developed herein; equation 2.2.3-6 concerns the calculation of Chl-*a* concentrations across all lakes while equations 2.2.3-7 and 2.2.3-8 regard the Chl-*a* assessment of natural-only and artificial-only lakes, respectively. Hence, taking into consideration those Chl-*a* models, we calculated TSI (Chl-*a;* Eq. 2.2.3-3) by using the equation being established regardless the lake characteristics (Equation 2.2.3-6), TSI (Chl-*a*) of natural lakes by employing the respective equation (Equation 2.2.3- 7) and TSI (Chl-*a*) of reservoirs by using the Chl-*a* equation respectively developed (Equation 2.2.3-8). Then we used the hereby developed models concerning the TP and Secchi depths for the calculation of satellite derived TSI (TP; Eq. 2) and TSI (SDT; Eq. 2.2.3-4), respectively. After implementing equations (2.2.3-2), (2.2.3-3), (2.2.3-4) and (2.2.3-5), satellite-derived TSI values have been calculated and trophic state classification has been conducted initially for the cases concerning all the lakes and then separately for the natural-only and artificial-only cases (by using the independent models). Validation of satellite TSI was carried out based on statistical analysis and the resulted deviation from the respective *in-situ* TSI values.

$$logChla = 3.599 - 0.63 * \left(\frac{blue}{red}\right) - 2.183 * \left(\frac{lnred}{lnswir2}\right) \quad (2.2.3\text{-}6)$$

$$logChla = 4.443 - 1.421 * \left(\frac{blue}{green}\right) - 3.454 * \left(\frac{lnred}{lnswir2}\right) + 1.304 * \left(\frac{red}{green}\right) (2.2.3\text{-}7)$$

$$logChla = 2.919 - 2.011 * \left(\frac{lnred}{lnswir1}\right) + 1.449 * \left(\frac{red}{green}\right) - 1.441 * \left(\frac{lnred}{lnblue}\right) (2.2.3\text{-}8)$$

2.2.4 Results

2.2.4.1 Secchi depth and Total phosphorus Quantitative models for Greek lakes

2.2.4.1.1 Secchi depth models

Spearman r values that resulted from the correlation analysis among all available band transformations and Secchi depth values, log, ln and SQRT Secchi depth values ranged from -0.56 to +0.56. In total 74 band transformations have been elaborated in correlation analysis and those parameters are provided in the Appendix (Table 1). Correlation matrix in combination with the predictor importance chart (IBM SPSS software Statistics v. 23.0, Armonk, NY, USA) indicated the highest important predictors. Values of importance for the same variables varied depending on the dependent parameter (Secchi, SQRTSecchi etc.), some variables were common for all the Secchi transformations (Table 2.2.4-1) whereas each Secchi transformation (e.g. SQRT, LOG, LN) indicated also some different variables that were important concerning their prediction. Those variables/predictors were further inserted in several combinations in numerous stepwise linear regressions. Application of multi-collinearity tests (i.e. Variance Inflation Factor-VIF with values higher than 1 and less than 10 and Tolerance higher than 0.1) and R^2 values indicated the optimal Secchi quantitative models which included as dependent variables the ln-, log- and SQRT Secchi transformations, with the latter proven to be the most satisfactory (Equation 2.2.4-1; Table 2.2.4-2). The selected $SQRT(Secchi)_{general}$ model incorporated ratios of bands blue, red and green from the visible spectrum and the second band from the short-wave infrared part of spectrum while collinearity statistics suggested an absence of autocorrelation.

Table 2.2.4-1. Common variables with the highest value of importance concerning the prediction of Secchi, SQRTSecchi and LOG/LN Secchi, derived from the predictor importance chart.

	Value of Importance		
Variable	**Secchi**	**SQRT(Secchi)**	**LOG-LN(Secchi)**
Green/SWIR1	0.014	0.008	0.011
LOG(Blue/Red)	0.033	0.044	0.041

(Blue -Red)/ (Blue + Red)	0.034	0.045	0.042
LN Green/LN Blue	0.035	0.041	0.045
Red / Blue	0.035	0.045	0.046
LOG Blue /LOG Green	0.037	0.043	0.047
LN((Blue-SWIR2)/(Green-SWIR1))	0.039	0.032	0.038
(Blue -Red)/ Green	0.046	0.054	0.050
Blue + Red + Red /Blue	0.046	0.050	0.047
Green/Blue	0.052	0.052	0.058
(Blue -Green)/ (Blue +Green)	0.056	0.055	0.059
LOG (Blue /Green)	0.056	0.055	0.059

Table 2.2.4-2. Regression analysis statistics and $Secchi_{general}$ model's summary.

Model	R	R^2	Adjusted R^2	Std. Error of the Estimate	Durbin-Watson
$Secchi_{general}$	**0.74**	**0.54**	**0.54**	**0.46**	2.24
Predictors: (Constant), Blue+Red+Red/Blue, LN Green/LN SWIR2					

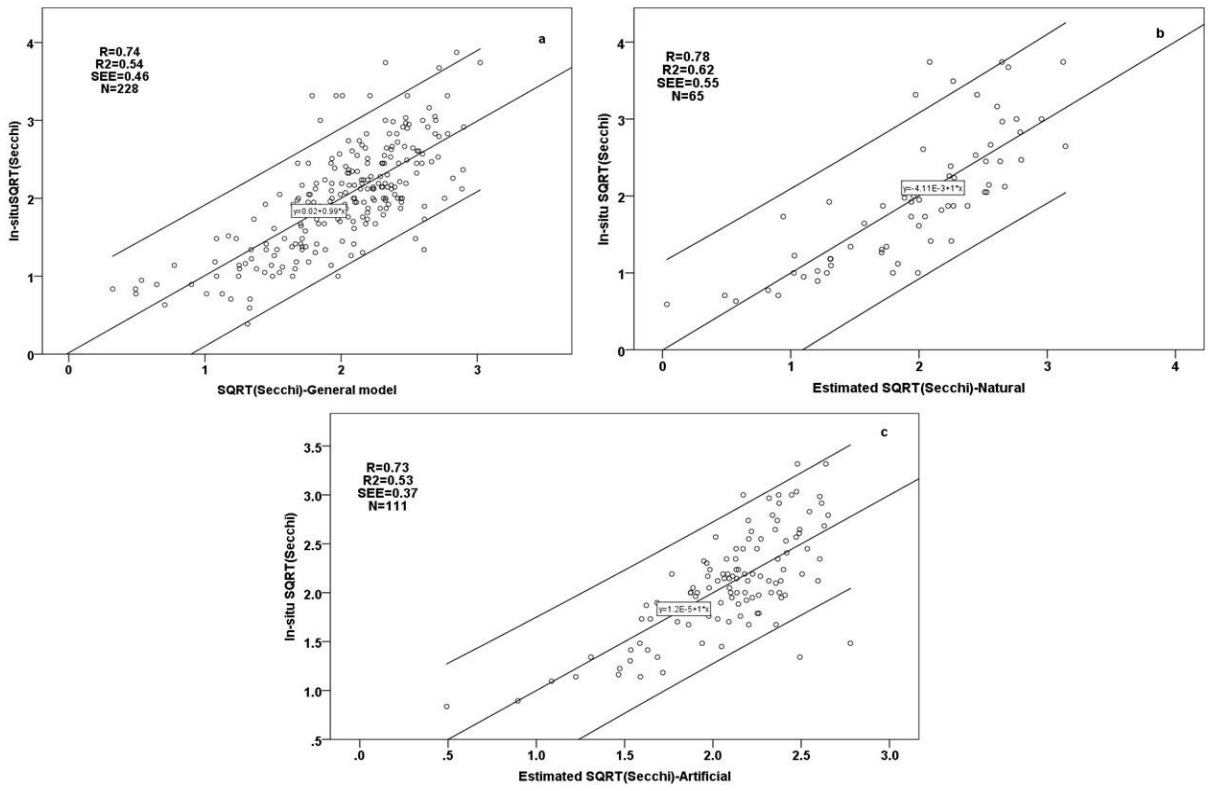


Figure 2.2.4-1. Scatter plots between *in-situ* and estimated SQRT Secchi depths derived from a) General model, b) model established for natural lakes and c) model established for reservoirs (lines set at confidence intervals 95%).

To ensure that further independent special models are essential to be developed for natural and artificial lakes to gain higher accuracy, the $Secchi_{general}$ model (Equation 2.2.4-1) was also separately applied to natural-only and artificial-only lakes. Even though some statistical indices were acceptable, special models for the different type of lakes proved to perform better compared to the general one, particularly concerning artificial lakes. Statistical and verification results derived from the application of the general model to natural and artificial lakes are presented at the validation section.

$$SQRT(Secchi)_{general} = 1.215 - 2.479 * \left(blue + red + \frac{red}{blue}\right) + 3.394 * (\frac{lngreen}{lnswir2}) \quad (2.2.4\text{-}1)$$

Subsequently, after the conduction of multiple MLR analyses employing separately natural-only and artificial-only lakes, the SQRT Secchi transformation proved as well to perform better in both cases and reflected adequate and reliable Secchi depths. It should be noted that autonomous elaboration of natural and artificial lakes signified the log-chl-*a* transformation as a Secchi predictor accompanied by high beta coefficient, especially for natural lakes (Table 2.2.4-3). Hence, the models that met the aforementioned criteria and were finally selected to calculate Secchi depth in natural (Equation 2.2.4-2) and artificial (Equation 2.2.4-3) lakes, included except for the logchl-*a*, visible bands as well: red, green and blue while equation 2.2.4-3 (artificial lakes) incorporated additionally the SWIR1 band.

$$SQRT(Secchi)_{natural} = 1.172 - (1.003 * logchl - a) - (1.031 * logred) \quad (2.2.4\text{-}2)$$

$$SQRT(Secchi)_{artificial} = 3.927 - 1.365 * \left(\frac{green}{blue}\right) - 0.318 * \left(\frac{red}{swir1}\right) - 0.361 * logchl - a \quad (2.2.4\text{-}3)$$

Table 2.2.4-3. Regression analysis statistics and $Secchi_{natural}$ and $Secchi_{artificial}$ models' summaries.

Scenario/Model	R	R^2	Adjusted R^2	Std. Error of the Estimate	Durbin-Watson
Secchi natural	**0.78**	**0.6**	**0.59**	**0.55**	2.14
Secchi artificial	**0.73**	**0.53**	**0.51**	**0.37**	2.12
Predictors natural: (Constant), Log Chl-*a*, Log Red Predictors artificial: Green/Blue, Red/SWIR1, Log Chl-*a*					

2.2.4.1.2 Total phosphorus models

The correlation matrix among all variables including all the lakes with mean depth higher than 5 meters resulted in slightly weaker correlations than those regarding Secchi depths. In this case, Spearman threshold value was reduced to ±0.3 to discriminate and incorporate more phosphorus variables/predictors. In total 69 band transformations have been elaborated in correlation analysis and those parameters are provided in the Appendix (Table 2). Furthermore, coefficient of determination among phosphorus, chlorophyll-*a* and Secchi depths were very high with values equal to 0.85 and -0.84, respectively. Optimal predictors with Spearman values higher than ±0.3 were further enriched and confirmed based on the calculation of their significance according to the significance predictor chart. Final selected predictors (Table 2.2.4-4) were inserted in manifold stepwise MLRs. The insertion of Chl-*a* and Secchi depth data as independent variable in MLRs, improved the results and yielded some statistically acceptable models employing some of those predictors.

Table 2.2.4-4. Common variables with the highest value of importance concerning the prediction of TP and LOG/LN TP, derived from the predictor importance chart.

	Value of Importance	
Variable	TP	LOG-LN (TP)
Red/SWIR1	0.2672	0.3283
Green/SWIR1	0.2296	0.2973
LN Green /LNSWIR1	0.1308	
Green / Red	0.1249	0.1525
LOG Chl-*a*	0.0953	0.1848
LOG (Red / Green)	0.0776	
LN Red /LN Green	0.0344	
LN Secchi		0.1315

Among the most optimal models, Equation 2.2.4-4 is the one selected for TP quantification in Greek lakes, employing except for the Chl-*a*, the band ratio of Ln-Red and Ln-SWIR1 bands. Both predictors are accompanied by equally high beta coefficient values while Durbin-Watson's statistic test is fully acceptable (Table 2.2.4-5).

$$LogTPgeneral = -1.425 + 0.452 * logChla - 0.573 * (\frac{lnred}{lnswir1}) \quad (2.2.4\text{-}4)$$

Table 2.2.4-5. Regression analysis statistics and $TP_{general}$ model's summary.

Model	**R**	**R^2**	**Adjusted R^2**	**Std. Error of the Estimate**	**Durbin-Watson**
$LogTP_{general}$	**0.85**	**0.73**	**0.71**	**0.18**	2.34
Predictors: (Constant), LogChl-*a*, LN Red/LN SWIR1					

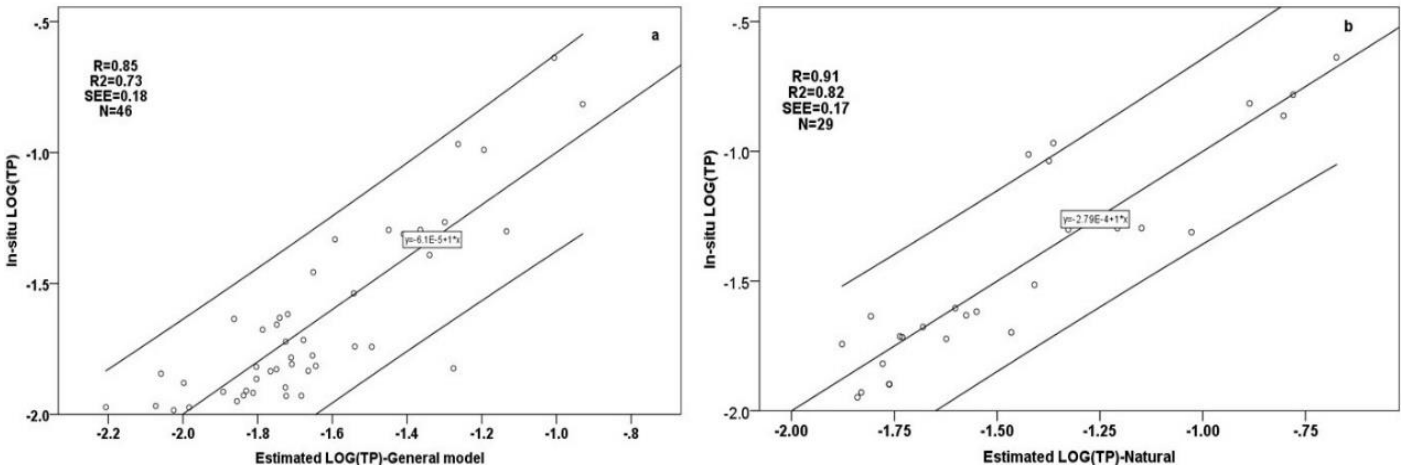


Figure 2.2.4-2. Scatter plots between *in-situ* and estimated LOG TP values derived from a) General model, b) model established for natural lakes (lines set at confidence intervals 95%).

MLRs that concerned artificial lakes, resulted in weak models characterized by poor statistical performance -R^2 values ranged from 0.13 up to 0.3- while TP models concerning the natural lakes managed to deliver highly acceptable results based on given statistical indices. Since no special model was delivered for TP quantification in artificial lakes, the $LogTP_{general}$ model's further efficiency was explored by applying it on natural and artificial lakes (dataset of 2018) while the results are presented in the validation section. Concerning natural lakes, the log Secchi proved to be a strong TP predictor, followed by the band ratio of green and red (Equation 2.2.4-5). The best quantitative $TP_{natural}$ model is characterized by high Pearson's and coefficient of determination values while no autocorrelation problem is detected (Table 2.2.4-6).

$$LogTPnatural = -0.633 - (0.704 * logSecchi) - 0.392 * (\frac{green}{red}) \quad (2.2.4\text{-}5)$$

Table 2.2.4-6. Regression analysis statistics and $TP_{natural}$ model's summary.

Model	R	R^2	Adjusted R^2	Std. Error of the Estimate	Durbin-Watson
$LogTP_{natural}$	**0.91**	**0.82**	**0.81**	**0.17**	1.9
Predictors: (Constant), LogSecchi, Green/Red					

2.2.4.2 Models' validation

This section presents the results of the analysis concerning the evaluation of the general models' ($Secchi_{general}$, $TP_{general}$) performance after their application separately on natural-only and artificial-only lakes. Since those

two general models were developed based on data of 2013-2016, their validation was conducted based only on data of 2018. Regarding the Spearman value, all correlations selected and presented are significant at the 0.01 level.

2.2.4.2.1 Secchi depth models

$Secchi_{general}$ model was developed based on 228 cases and validated twice based on 55 and 105 cases, respectively while RMSE values are similar, equal to 1.6 m (1st validation) and 1.7 m (2018 validation), respectively (Table 2.2.4-7). Application of $Secchi_{general}$ model on natural lakes resulted in acceptable and reliable results and similar values regarding the examined statistical indices as those derived by the $Secchi_{natural}$ model. Nevertheless, Spearman and RMSE values of $Secchi_{natural}$ model (2018 data) are better than the general's one, hence $Secchi_{natural}$ is selected as the most optimum to quantify Secchi depths in natural lakes.

As far as the artificial lakes are concerned, $Secchi_{artificial}$ model is selected compared to the $Secchi_{general}$, since both average residuals and the RMSE values are lower (-0.2 compared to 0.5 m and RMSE 1.4 m compared to 1.7 m).

Table 2.2.4-7. Statistical indices used to validate the Secchi selected algorithms (**correlation significant at the 0.01 level (two-tailed). RMSE—root-mean-square error. * Values concern Secchi depths in m).

	1st validation (20%)					2nd validation (2018 data)				
Models	**Spearman r**	**Average *in-situ****	**Average satellite***	**Average residuals (m)**	**RMSE (Secchi; *m*)**	**Spearman r**	**Average *in-situ****	**Average satellite***	**Average residuals (m)**	**RMSE (Secchi; *m*)**
$Secchi_{general}$ Training dataset N=228	.78** N=55	4.45	4.7	-0.24	1.6	.58** N=105	4.2	3.8	0.4	1.7
$Secchi_{general}$ applied on natural						.65** N=42	3.8	3.6	0.15	1.7
$Secchi_{general}$ applied on artificial						.51** N=48	4	3.5	0.5	1.7
Secchi natural Training dataset N=65	.95** N=24	3.1	3.3	-0.21	1.1	.73** N=24	3.2	3.5	-0.3	1.1
Secchi artificial Training dataset N=111	.62** N=23	3.9	4.2	-0.24	0.89	.56** N=39	4.7	4.8	-0.2	1.4

2.2.4.2.2 Total phosphorus models

$LogTP_{general}$ model performed well concerning both validation procedures (Table 2.2.4-8). High Spearman values derived from datasets of 12 and 31 cases respectively, and similar average *in-situ* and satellite TP values characterize both validations. RMSE is 0.008 mg/l (1st validation) and 0.03 mg/l (validation of 2018). Application of the $LogTP_{general}$ model on artificial lakes yielded as well acceptable results since RMSE equals to 0.03 mg/l.

Concerning the application of the general model on natural lakes, it is clear that the special developed model for natural lakes is superior since the values of average residuals and RMSE are quite lower (-0.003 compared to 0.01 mg/l; RMSE=0.03 compared to 0.08 mg/l). Moreover, higher Spearman value (0.68) and larger size of the validation dataset (n=47) indicate as well the advantage of this model in assessment of TP concentrations in natural lakes.

Table 2.2.4-8. Statistical indices used to validate the TP selected algorithms (**correlation significant at the 0.01 level (two-tailed). RMSE—root-mean-square error. *All values concern TP in mg/l.

	1st validation (20%)					**2nd validation (2018 data)**				
Models	**Spearman r**	**Average *in-situ****	**Average satellite***	**Average residuals (mg/l)**	**RMSE (TP; mg/l)**	**Spearman r**	**Average *in-situ****	**Average satellite***	**Average residuals (mg/l)**	**RMSE (TP; mg/l)**
$LogTP_{general}$ Training dataset **N=46**	.71** **N=12**	0.02	0.02	0.001	0.008	.81** N=31	0.08	0.08	0.002	0.03
$LogTP_{general}$ applied on **natural**						.55** N=38	0.09	0.08	0.01	0.08
$LogTP_{general}$ applied on **artificial**						.86** N=11	0.06	0.07	-0.02	0.03
$LogTP_{natural}$ Training dataset **N=29**	.93** **N=7**	0.034	0.04	-0.008	0.02	.68** N=47	0.07	0.08	-0.003	0.03

Furthermore, Chl-*a*, Secchi depth and TP maps of selected lakes were created after application of the Chl-*a* (Figure 2.2.4-3) algorithms derived by Markogianni et al. (2020) and the herein developed Secchi (Figure 2.2.4-4) and TP algorithms (Figure 2.2.4-5). The Landsat 8 OLI satellite image of 11 August 2013 was used in order to produce the satellite-derived spatial distribution of the studied WQ parameters of this day, while the respective *in situ* values of

those lakes were sampled with -2 and +5 days (Chl-*a*) and -5 and +5 days (Secchi depth) of difference from the aforementioned date while there is no available *in situ* data for their TP concentrations.

Application of the Secchi general model resulted in Secchi depth values ranging from 0.000002 to 8.2 m and from 0.000005 to 30.2 m for natural and artificial lakes, respectively (Figure 2.2.4-4a). Application of the Secchi natural model in natural lakes, yielded Secchi depths ranging between 0.000001 and 7.8 m (Figure 2.2.4-4b), while the Secchi artificial model resulted in Secchi depth's values varying from to 0.05 to 8.4 m (Figure 2.2.4-4c), as far as the artificial lakes are concerned. Secchi general model (Equation 2.2.4-1) was applied by using the aforementioned band combinations while Secchi natural (Equation 2.2.4-2) and Secchi artificial (Equation 2.2.4-3) models were applied including the respective Chl-*a* equations specially designed for the natural (Equation 2.2.3-7; Markogianni et al., 2020) and artificial (Equation 2.2.3-8; Markogianni et al., 2020) lakes, respectively.

Concerning the application of TP general model which also includes Chl-*a,* Equation 2.2.4-4 was used whereas TP model of natural lakes employed the Secchi natural model in order to be applied. TP general model resulted in values ranging from 0.0008 to 0.85 mg/l and from 0.002 to 0.12 mg/l for natural and artificial lakes, respectively (Figure 2.2.4-5a). Values of specially designed TP model for natural lakes vary from 0.016 to 19 mg/l while only a few values are higher than 0.2 mg/l (Figure 2.2.4-5b). Furthermore, since the variance of TP estimated values is small, it was decided to present those values by grouping them in classes as stretching values resulted in low quality of results' presentation. Furthermore, it should be noted that all parameters' values have been converted in actual units e.g. Chl-*a* in μg/l, Secchi depth in meters and TP in mg/l to facilitate the understanding and the comparison among the concentrations.

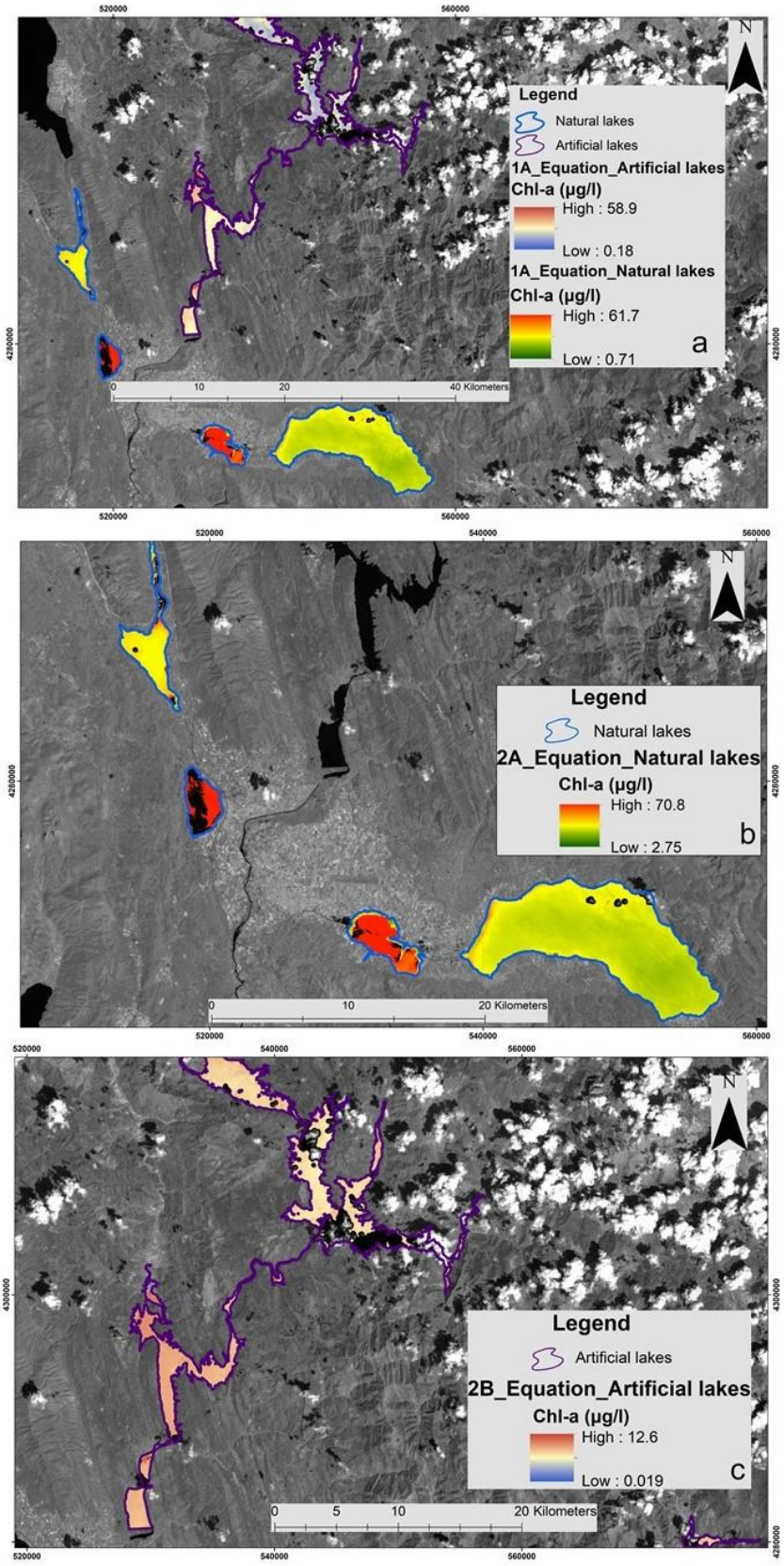


Figure 2.2.4-3. Satellite-derived Chl-a maps (on 11 August 2013) of selected lakes after the application of General-(a), Natural (b), and Artificial (c) models (WGS_1984, UTM Zone 34 N Coordinate system), derived by Markogianni et al. (2020).

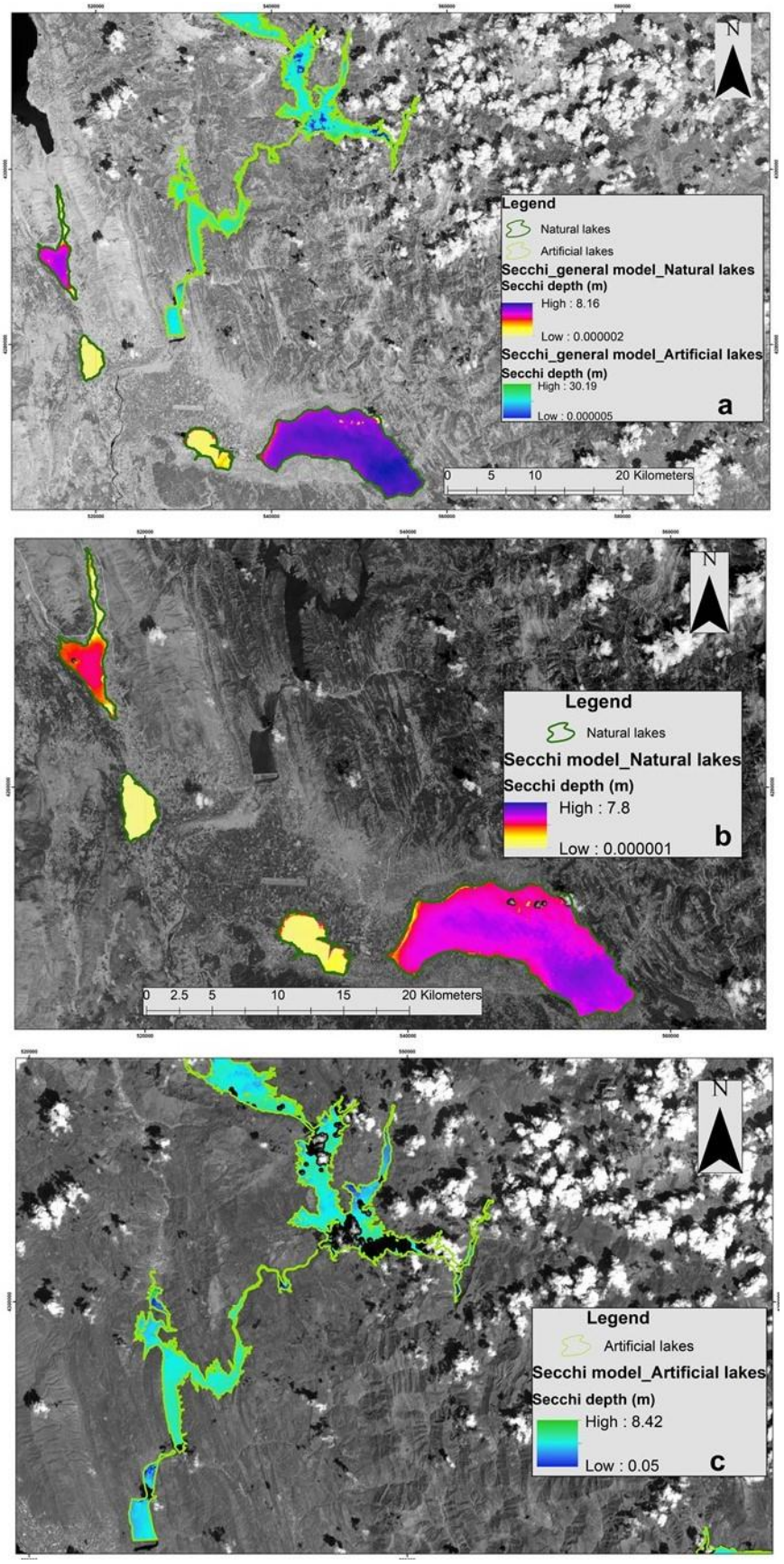


Figure 2.2.4-4. Satellite-derived Secchi maps (on 11 August 2013) of selected lakes after the application of Secchi General-(a), Secchi Natural (b), and Secchi Artificial (c) models (WGS_1984, UTM Zone 34 N Coordinate system).

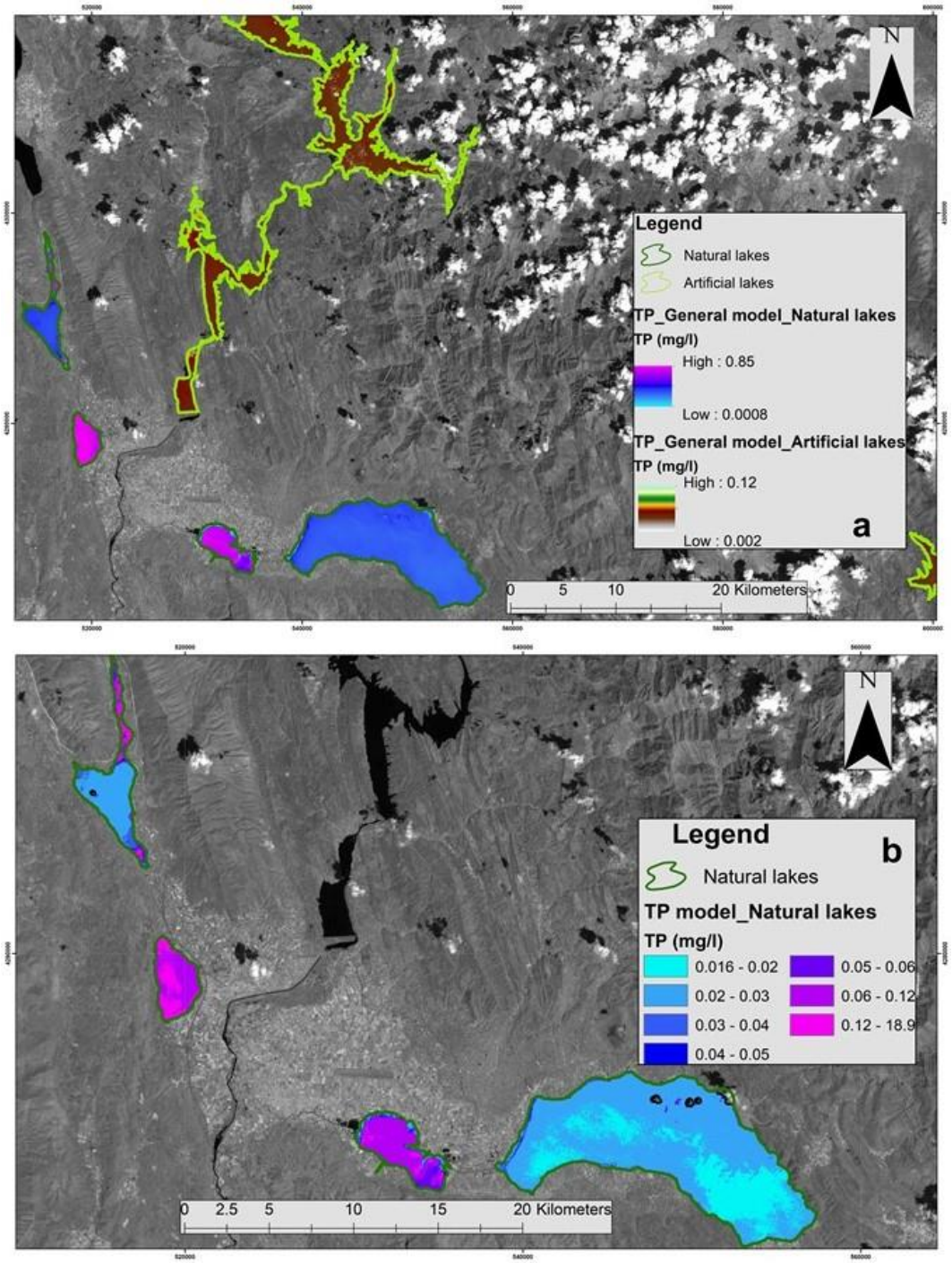


Figure 2.2.4-5. Satellite-derived TP maps (on 11 August 2013) of selected lakes after the application of TP General-(a**),** and TP natural (b) models (WGS_1984, UTM Zone 34 N Coordinate system).

2.2.4.3 Satellite derived assessment of trophic status of Greek lakes based on Carlson's Trophic State Index

2.2.4.3.1 Evaluation of the lake trophic status's assessment based on the whole dataset

Calculation of both types of TSI using *in-situ* only data and models based on satellite data concerns the attributes that were accompanied by available simultaneous *in-situ* measurements of TP and Chl-*a* concentrations and Secchi depth measurements (176 total cases). Since *in-situ* available TP data are those analyzed during the years 2015-2016 and 2018, both calculated TSIs concern the same period. Furthermore, the application of equations (2.2.3-2), (2.2.3-3), (2.2.3-4) and (2.2.3-5) concerning both the *in-situ* only data and the models resulted in categorizing the under-study attributes (and by extension the lakes) in 5 classes regarding their trophic status (Table 2.2.4-9). The main difference is that *in-situ* measurements indicated 1 eutrophic case and none hypereutrophic while remote sensing detected 2 hypereutrophic cases and none eutrophic. In both analyses, cases that are characterized as oligotrophic are the majority of the entire dataset and cases with a tendency to mesotrophy and mesotrophic ones occupy the next positions.

Table 2.2.4-9. *In-situ* and satellite derived TSIs' s frequencies and percentages of all cases.

WHOLE DATASET	**TSI (*in-situ*)**	**TSI (satellite)**	**TSI (*in-situ*)**	**TSI (satellite)**
	Frequency		**Valid Percent**	
1 (Oligotrophic)	92	124	52.3	70.5
2 (Oligotrophic-Mesotrophic)	42	30	23.9	17
3 (Mesotrophic)	26	15	14.8	8.5
4 (Mesotrophic-Eutrophic)	15	5	8.5	2.8
5 (Eutrophic)	1	-	0.6	-
6 (Hypereutrophic)	-	2		1.1
Total	176	176	100.0	100.0

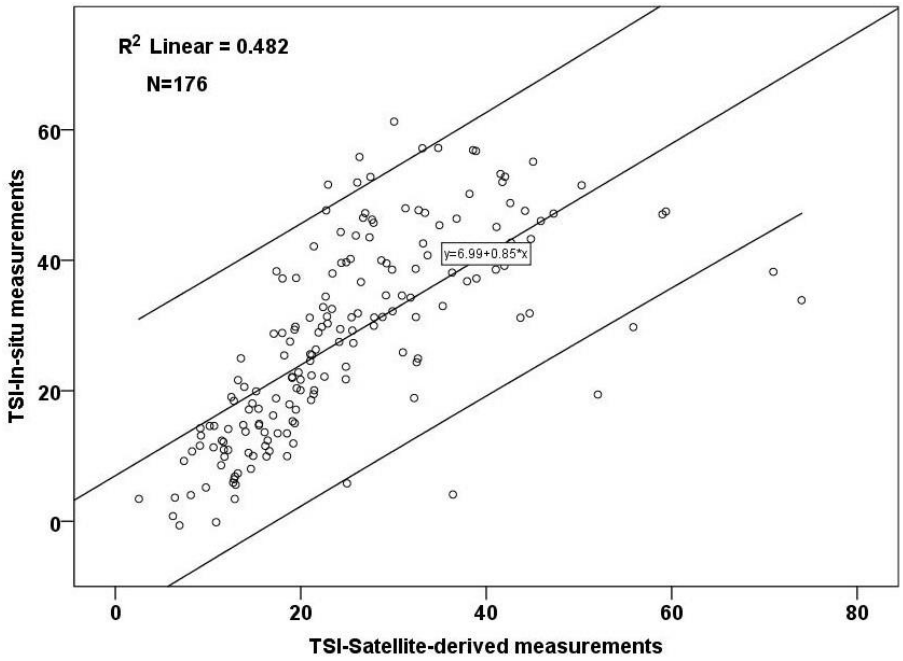


Figure 2.2.4-6. Scatter plot between *in-situ* and satellite-derived TSI values, based on the whole dataset (lines set at confidence intervals 95%).

Further statistical analysis suggested that 103 out of 176 attributes (58.5 %) were identically classified based on the two TSI calculations while 50 cases out of 73 that were classified differently, were allied to the right previous or next class (-1, +1) in relation to the *in-situ* results. Furthermore, attributes that were misclassified in 3 or 4 classes away from the *in-situ* ones are in total 8, which correspond to a 4.5 % of the misclassified dataset. Considering the mean depth of the lakes (Figure 2.2.4-7a), it is proven that cases concerning deeper lakes (> 5 m) were more successfully classified than the shallow ones verifying the effect of the bottom reflectance as an obstacle in the remote sensing elaboration. Records belonging to natural lakes were the majority of those that were either identically (56 out of 112) or by-one-class misclassified (40 out of 112; Figure 2.2.4-7b).

Concerning the sampling season (Figure 2.2.4-7c), results of summer months resulted in 68 attributes that were identically classified by both TSI calculations, while 40 out of 120 were misclassified in the previous or next trophic status class. All records regarding spring- monitored lakes were identically classified while remote sensing concerning autumn season indicated a slight weakness in properly classifying the trophic status of lakes compared to summer. The total number of 176 attributes is divided in 49 sampled in autumn, 7 sampled in spring and 120 cases sampled in summer months. Those sampling dates are accompanied by 26 images during autumn, 6 images during spring and 51 images during summer months.

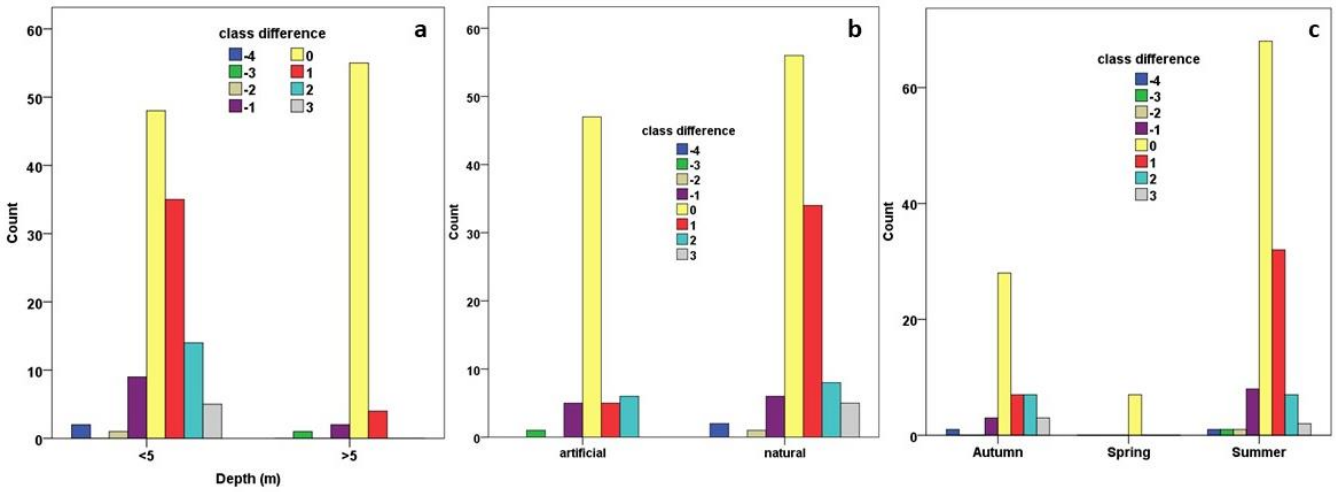


Figure 2.2.4-7. Count of satellite-classified/misclassified cases concerning all monitored lakes grouped by (a) the lake's mean depth, (b) lakes' nature and (c) sampling season. Numbers from -4 up to 3 represent the class deviation between the satellite and in-situ derived TSIs while 0 indicates no differentiation. Positive and negative signs represent the direction of the deviation from oligotrophy to hyper-eutrophy and vice versa, respectively based on the corresponding in-situ TSI value (reference value).

2.2.4.3.2 Evaluation of the lake trophic status assessment concerning natural and artificial lakes

The calculation of *in-situ* TSI values of records belonging to natural lakes categorized them in 5 trophic status classes (1-5), while satellite TSI resulted in 6 classes (1-6), characterizing 5 cases as hypereutrophic (Table 2.2.4-10; Figure 2.2.4-8). Furthermore, the majority of those attributes were characterized as oligotrophic and oligotrophic-mesotrophic based on both calculations while one case was classified as eutrophic by both calculations.

Table 2.2.4-10. *In-situ* and satellite derived TSIs' s frequencies and percentages of cases belonging to natural lakes.

NATURAL LAKES	**TSI (*in-situ*)**	**TSI (satellite)**	**TSI (*in-situ*)**	**TSI (satellite)**
	Frequency		**Valid Percent**	
1 (Oligotrophic)	50	59	44.6	52.7
2 (Oligotrophic-Mesotrophic)	35	29	31.3	25.9
3 (Mesotrophic)	14	11	12.5	9.8
4 (Mesotrophic-Eutrophic)	12	7	10.7	6.3
5 (Eutrophic)	1	1	0.9	0.9
6 (Hypereutrophic)	-	5	-	4.5
Total	112	112	100.0	100.0

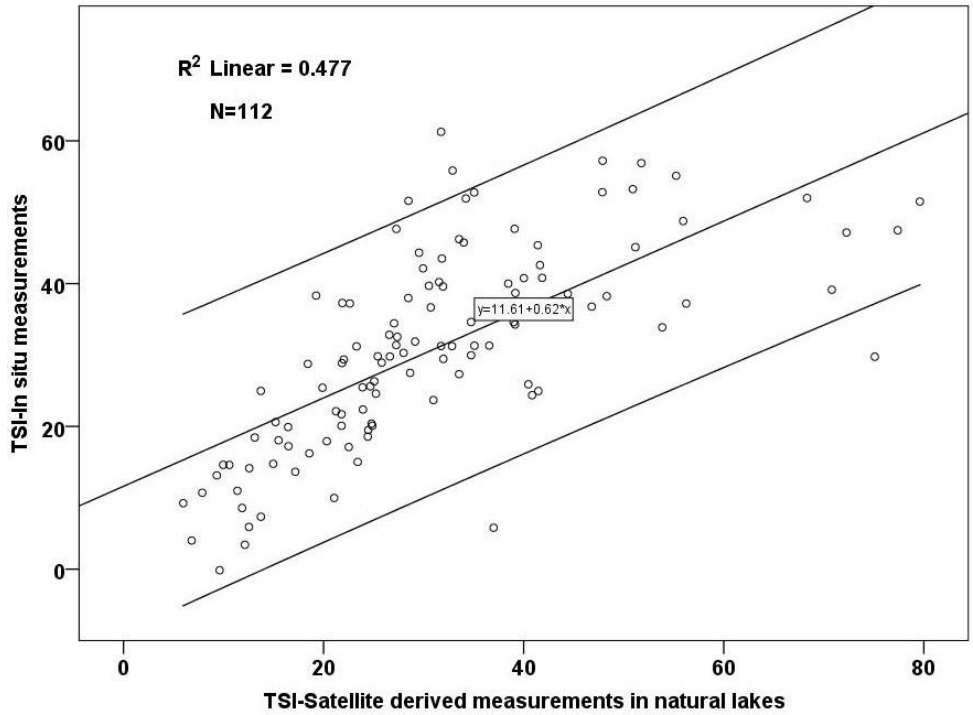


Figure 2.2.4-8. Scatter plot between *in-situ* and satellite-derived TSI values, of natural lakes (lines set at confidence intervals 95%).

From the total of 112 records concerning the natural lakes, 66 of them were identically classified in the same class while the 46 that presented differences concern mostly cases that were misclassified by only one class (30 out of 46). Furthermore 6 cases out of 46 were misclassified by three or four classes away from the respective *in-situ* ones. Trophic status classification of deep natural lakes (average depth> 5 m) was in particular successful since 22 out of 27 cases were identically classified according to both TSI calculations and the rest of 5 cases were misclassified by only one class deviation (Figure 2.2.4-9a). Trophic status classification of shallower natural lakes was also satisfactory since 44 out of 85 cases have no difference regarding their classification, 25 were misclassified by only one class deviation while 10 cases were misclassified by 2-classes from the respective *in-situ* ones.

As far as the water sampling seasons are concerned (Figure 2.2.4-9b), calculation of satellite derived average TSI during summer months was also proved successful since 52 out of 80 cases presented no difference compared to respective *in-situ* TSI while 18 presented misclassifications by one category deviation. Furthermore, all of 5 cases concerning spring-monitored lakes were identically classified based on both *in-situ* and satellite TSI values. The calculation of TSI throughout the natural lakes was based on the acquirement of 19 images while 4 and 36 images were used for calculating the spring and summer TSI, respectively including 5 and 80 attributes.

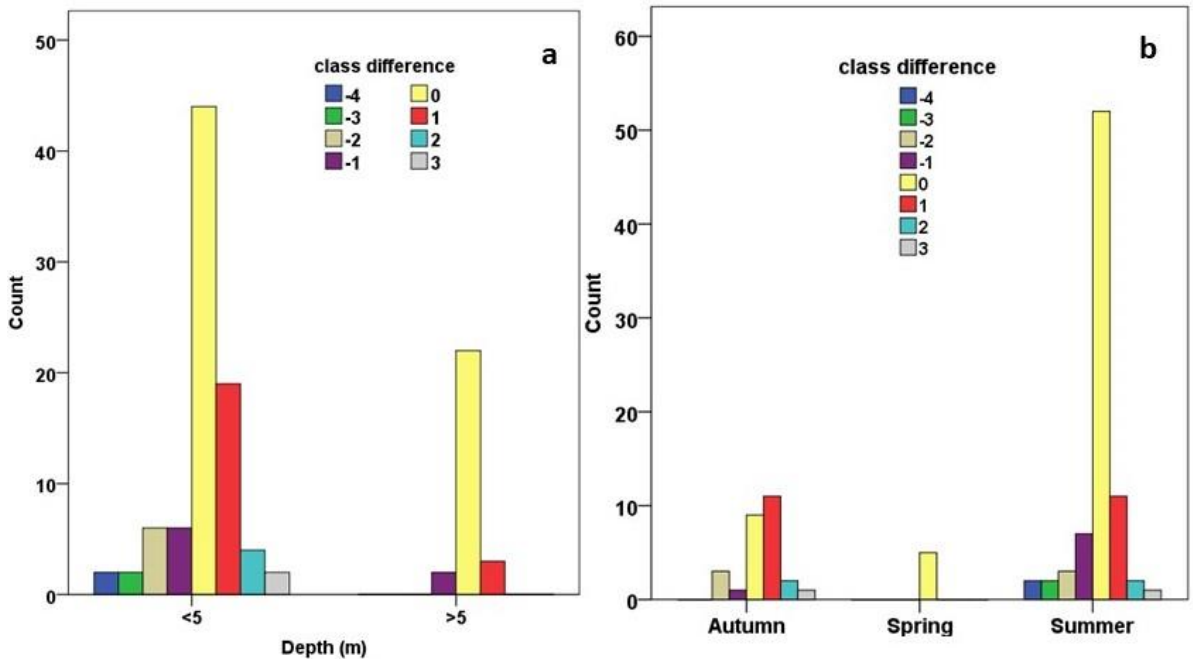


Figure 2.2.4-9. Count of satellite-classified/misclassified cases belonging to natural lakes grouped by (a) the lake's mean depth and (b) sampling season. Numbers from -4 up to 3 represent the class deviation between the satellite and *in-situ* derived TSIs while 0 indicates no differentiation. Positive and negative signs represent the direction of the deviation from oligotrophy to hyper-eutrophy and vice versa, respectively based on the corresponding *in-situ* TSI value (reference value).

As far as the artificial lakes are concerned, both *in-situ* and satellite TSI calculations resulted in similar trophic status classifications and identical classes (1-4; Table 2.2.4-11). The majority of records concerning the artificial lakes are characterized as oligotrophic and oligotrophic-mesotrophic while 3 cases were classified as mesotrophic-eutrophic based on both TSI values (*in-situ* only, models).

Table 2.2.4-11. *In-situ* and satellite derived TSIs' s frequencies and percentages of cases belonging to artificial lakes.

ARTIFICIAL LAKES	**TSI (*in-situ*)**	**TSI (satellite)**	**TSI (*in-situ*)**	**TSI (satellite)**
	Frequency		**Valid Percent**	
1 (Oligotrophic)	42	48	65.6	75
2 (Oligotrophic-Mesotrophic)	7	9	10.9	14.1
3 (Mesotrophic)	12	4	18.8	6.3
4 (Mesotrophic-Eutrophic)	3	3	4.7	4.7
Total	**64**	**64**	**100.0**	**100.0**

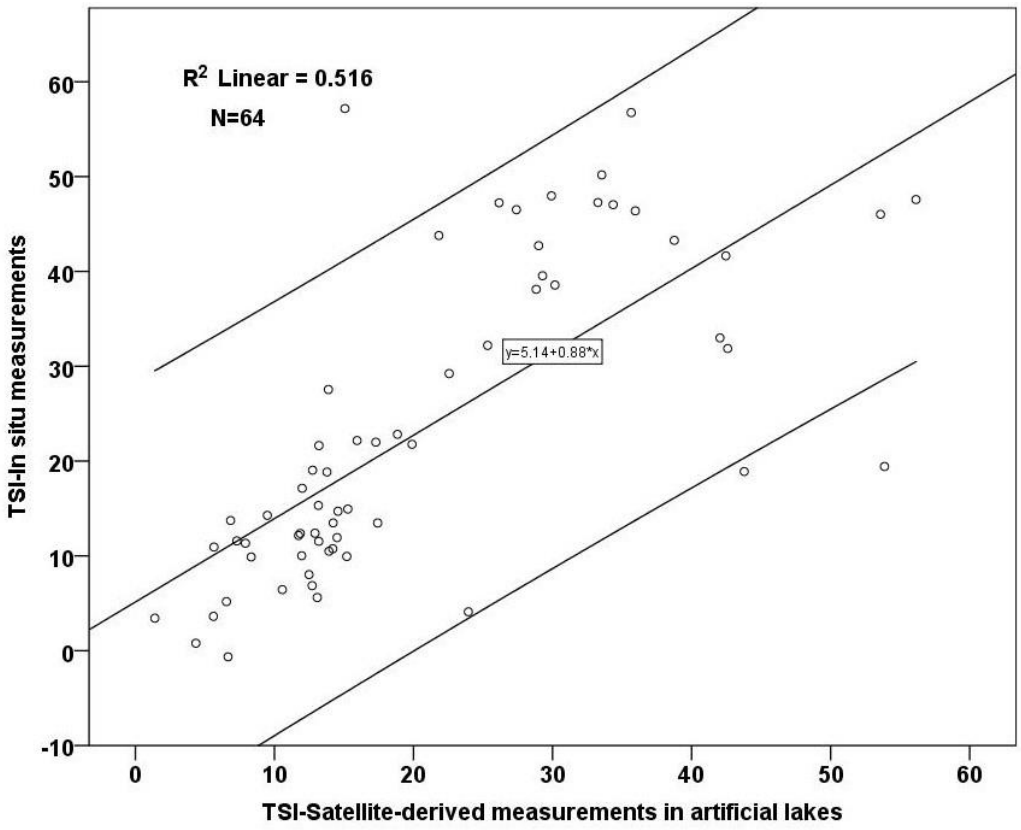


Figure 2.2.4-10. Scatter plot between in-situ and satellite-derived TSI values, of artificial lakes (lines set at confidence intervals 95%).

Regarding the trophic status misclassifications of artificial lakes, only 21 out of 64 records were misclassified and particularly 12 out of 21 were classified in categories that deviated only 1 class away from the respective *in-situ* ones. Observing artificial lakes based on their mean depth (Figure 2.2.4-11a), it is proven that attributes regarding deeper artificial lakes were successfully classified concerning their trophic status since 33 out of 35 presented no classification differentiation and two (2) of them were misclassified in classes that deviated 3 and 2 classes from the *in-situ* ones, respectively. Additionally, cases belonging to shallower artificial lakes were also satisfactory classified as 10 out of 29 showed no differentiation and 12 were misclassified by one class difference.

Observing classification of artificial lakes based on the sampling season (Figure 2.2.4-11b), it is clear that not only summer trophic status classifications are successful (26 out of 40 cases presented no differentiation) but also TSI calculations during spring and autumn seasons managed to classify records with high accuracy. TSI classification throughout the artificial lakes during autumn was conducted by using 15 Landsat images, while 2 and 28 images were used for spring and summer seasons, respectively.

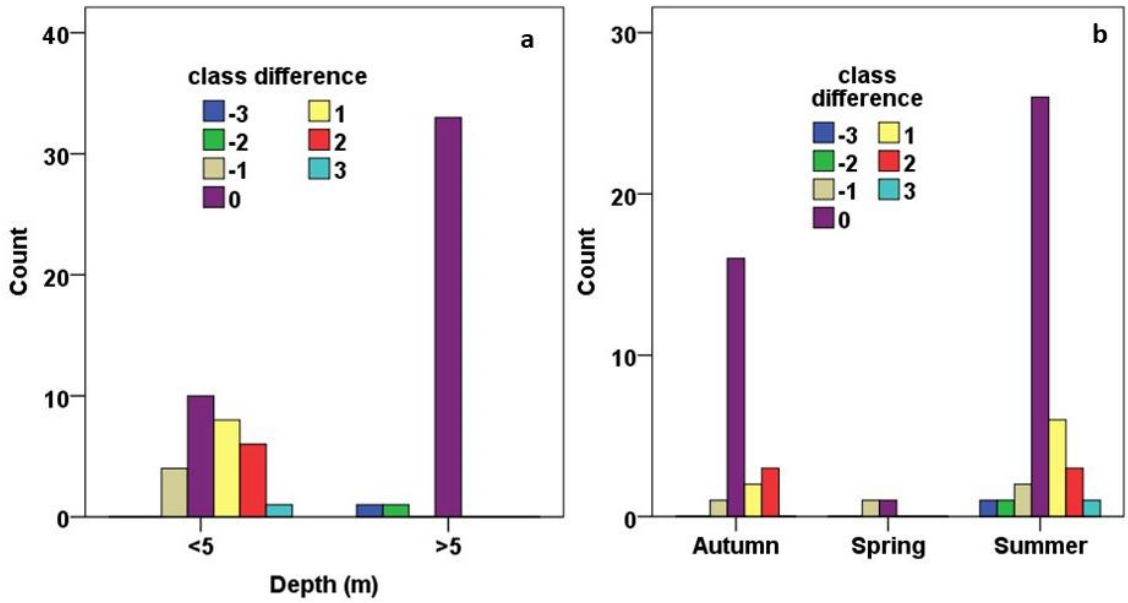


Figure 2.2.4-11. Count of satellite-classified/misclassified cases belonging to artificial lakes grouped by (a) the lake's mean depth and (b) sampling season. Numbers from -3 up to 3 represent the class deviation between the satellite and *in-situ* derived TSIs while 0 indicates no differentiation. Positive and negative signs represent the direction of the deviation from oligotrophy to hyper-eutrophy and vice versa, respectively based on the corresponding *in-situ* TSI value (reference value).

2.2.5 Discussion

Increasing human activities and industrialization have dramatically degraded lake water quality (Zheng et al., 2021). Therefore, implementation of WFD in Greece, as well as in other European countries, has as a main aim to ensure sustainable management of lakes. Use of geoinformation technologies - and in particular of RS and GIS - with conventional *in-situ* water samplings have been proven as the most efficient, cheap and reliable tool to monitor WQ parameters in lakes. WFD has been implemented in Greece at least the last seven years while numerous *in-situ* measurements of WQ elements provide valuable means to scientists and public authorities to assess and monitor Greek lake WQ. In particular *in-situ* measurements of Secchi depths and TP concentrations combined with Landsat data have been utilized in this study framework to assess trophic status of monitored Greek lakes.

The significance of lakes' nature concerning the constituents' variance

Exploratory statistical analysis of the available datasets indicated higher Secchi depth values in artificial rather than in natural lakes during all sampling years (2013-2018) whereas the highest TP concentrations were detected in natural lakes, illustrating accumulating TP loadings and an increasing tendency throughout the years (2015-2018). Moreover,

Markogianni et al. (2020) reported that natural lakes presented also notably higher Chl-*a* concentrations in relation to reservoirs. The present study findings are also in accordance with those reported in other similar studies. For example, (Søballe et al., 1992) documented that chlorophyll-*a* concentrations tend to be lower in reservoirs than in natural lakes because higher inorganic turbidity and high flushing rates (low hydraulic residence times) in reservoirs limit the development of phytoplankton biomass. In this way, higher Secchi depth values in artificial lakes indicate clearer water. This is once again interpreted by a higher presence of non-algal turbidities in this type of lakes compared to natural (Canfield and Bachmann, 1981). Concerning the TP values, it should be noted that artificial lakes lose nutrients (in particular P) through settling in a downstream direction. The sampling station's location plays a major role in WQ monitoring. One of the main differences between artificial and natural lakes is that artificial characteristically exhibit a trophic gradient (Søballe et al., 1992) as it may grade from eutrophic (in its upper reaches) to oligotrophic (close to the dam) (Virginia 2007).

Correlation matrix among *in-situ* measurements of monitored WQ parameters throughout all lakes resulted in high and positive correlation between TP and Chl-*a* (0.85) and high negative relationship between Secchi depth with TP and Chl-*a* with values of coefficient of determination equal to -0.84 and -0.83, respectively. This finding agrees with results reported in other studies studying natural and artificial lakes around the world e.g. (Canfield and Bachmann, 1981; Canfield and Hodgson, 1983). For most lakes, chlorophyll *a* was highly correlated with SD, phosphorus was directly correlated with chlorophyll *a* and inversely correlated with SD. This is mainly due to the fact that increases in nutrient concentrations (in particular TP) result directly in higher algal growth (Chl-*a* concentration) and decreased water transparency (Secchi depth) (Virginia 2007). Additional explanation to the fact that Secchi depth is decreased with increasing TP concentration, was given by (Heiskary and Wilson, 2005), who proved that a proportion of phosphorus may be linked to suspended particles resulted from soil erosion and carried through river's downslope.

MLR analysis and resulted proxies of studied WQ parameters

MLRs analyses among *in-situ* Sechi depth measurements and Landsat 7 ETM+ and 8 OLI data yielded three (3) optimal Secchi estimation models concerning the assessment of Secchi depth of all lakes ($Secchi_{general}$), natural ($Secchi_{natural}$) and artificial ($Secchi_{artificial}$) ones. The $Secchi_{general}$ model incorporated a combination of bands blue, red, green and SWIR2 while models developed for natural and artificial lakes were accompanied by the insertion of logchl-*a* as a significant Secchi predictor. The $Secchi_{general}$ model

was also independently applied to natural and artificial lakes to further explore its effectiveness regarding the nature of lakes. The abovementioned model proved to perform better concerning the natural lakes than the reservoirs since water transparency in artificial lakes is notably influenced by non-algal sources of turbidity. This rationale is equally supported by (Lind 1986), who documented that the use of Chl-*a* to estimate Secchi depth is inappropriate for waters where even moderate amounts of non-algal turbidity are present. On the other, (Lorenzen 1980) proposed to take into consideration this type of turbidity when reservoirs are evaluated. However, many scientists argue that Secchi depth data are calibrated for each lake or reservoir, hence they may be used for WQ monitoring. Numerous algorithms have been developed for Secchi depth assessment. Relevant literature is enriched with studies that demonstrated strong relationships between Landsat data and *in-situ* Secchi depths by employing mostly the blue, green, red, NIR bands and their ratios of the visible spectrum (Allee and Johnson, 1999; Olmanson et al., 2001; Giardino et al., 2001; Olmanson et al., 2008) while in the framework of this paper we also tried and managed to combine other water quality indicators and remotely sensed spectral reflectance. Even more models based on Landsat series data have been empirically developed to map SD for inland and coastal waters (Olmanson et al., 2008; Doña et al., 2014; Page et al., 2019). However, in contrast to our work, those studies utilized calibration and validation datasets sampled from one, two or a few lakes within a small geographical region, failing to generate a uniform model for the systematic assessment of SD at a greater scale (Zhang et al., 2021). On the other hand, Zhang et al. (2021) constructed a general SD power function model (based on red band) established on extensive *in-situ* SD and Landsat reflectance from 225 China lakes, exploring SD spatial variation from 1986 to 2018. This study in agreement to ours, not only performed regression-related efforts but also confirmed that Landsat series data can result in an accurate long-term estimation of the SD. Another effort to develop a 20-year water clarity census on a broad regional and spatial scale has been conducted by Olmanson et al. (2001) who studied over 10500 lakes of Minnesota state. In particular, a regression model incorporating the blue and red bands of several Landsat series (4 MSS, 7 ETM+, 5 TM) demonstrated that satellite imagery is an accurate method to assess water clarity over a long period of time. Moreover, one of the latest studies that developed a unified model mapping global lake clarity using Landsat imagery was conducted by (Dekker et al., 1996). In the framework of this research, the combination of trained *in-situ* SD data (3586 data points; 2235 lakes across the world) and match-up Landsat images (TOA; L5-TM; L7-ETM+; L8-OLI) were used to establish various regression models. The proposed model based on the blue/green and red/blue bands demonstrated its applicability to monitor SD in inland bodies across the globe

and its stability to variations in time and space of the optical properties of lakes.

MLRs analyses among TP concentrations and Landsat band transformations yielded statistically weak models whereas further insertion of *in-situ* Chl-*a* and Secchi depth data improved the results. A general TP assessment model with application on all lakes was produced including the logarithmic transformation of Chl-*a* and the band ratio of Ln-Red and Ln-SWIR1 bands with reliable values of tested statistical indices. The fact that none statistically acceptable model was generated for artificial lakes may partly be attributed to the time lag that has been observed for phytoplankton to consume TP in this type of lakes. This fact makes the relationship between TP and Chl-*a* or SD more complicated (Song et al., 2012) in reservoirs and further limnological research is needed to additionally penetrate into the functions of those lakes' system. On the other hand, as far as the natural lakes are concerned, Secchi depth proved to be a strong TP predictor. The TP model developed for natural lakes incorporated also the ratio of green and red bands and was accompanied by a high value of coefficient of determination. The weakness of MLRs to produce an optimal TP model for artificial lakes urged us to further explore the efficiency of the TP general model on artificial and natural lakes as well. Application of the $TP_{general}$ model on artificial lakes (2018 data) yielded acceptable results, fact that characterizes it as reliable enough to be used at this type of lakes. On the other hand, special developed TP model for natural lakes was superior compared to the general one based on basic statistical indices (Spearman and RMSE values).

Uusitalo et al. (2000) suggested that TP could not be assessed using RS techniques because it represents dissolved constituents and it is characterized by weak optical characteristics and low signal noise ratio. Nevertheless, it has been investigated based on its high correlation with optically active constituents (Song et al., 2006; Busse et al., 2006) such as phytoplankton (Baban 1993) and Secchi depth (Ritchie et al., 1990). Furthermore, data from Landsat series, among many other satellite sensors, has been widely used for TP assessment in inland waters and especially lakes (Kutser et al., 1995; Wu et al., 2010). Lim and Choi (2015) selected a MLR model (R=0.57) using blue, green, red and NIR Landsat 8 bands to estimate TP among other WQPs in Nakdong River with weak accuracy. Further TP studies have detected similar correlations between the NIR band and the 3 visible bands (blue, green, and red) and Chl-*a* (Lathrop 1993; Lillesand and Kiefer 2000). Another research that utilized SWIR data for the assessment of phosphate concentrations in Akkulam–Veli Lake, Kerala, India is the one conducted by (Moses et al., 2014). They produced an equation (R^2= 0.5) accompanied except for the red band also by the MIR (middle infrared; band that followingly was replaced by the SWIR) (Buiteveld et al., 1994).

The contribution of SWIR bands in WQ monitoring of Case 2 waters

The results accrued by the herein MLRs analyses and the observed weight of SWIR bands regarding the calculation of Secchi depth and TP concentrations constitute a topic that needs further exploration and explanation. The main interpretation is based on the fact that lakes belong to Case 2 waters which are optically complex. Since those waters are influenced also by inorganic and yellow substances- except for phytoplankton and related particles- it is well recognized that sediment reflectance exceeds the absorptive properties in the NIR and SWIR wavelengths (IOCCG 2000; Ouma et al., 2020) and the standard algorithms in use today in Case 1 waters (especially for chlorophyll retrieval from satellite data), break down (IOCCG 2000). Furthermore, according to Moses et al. (2014), in cases where there is even a small quantity of impurities, significant changes are caused in the refractive index of a substance with substances containing more polarizer groups. Hence, since, for example, TP is a pollutant with more polarity, it changes the refractive index of water which in turns changes the reflectance of NIR and MIR in water. In accordance with this theory, there are several studies that have widely used SWIR bands concerning the monitoring of WQ elements in Case 2 waters. Barrett and Frazier (2016) studied WQ of lakes in eastern Oklahoma and indicated the existence of a relationship between SWIR reflection and algae/plant production by including at least one of the short-wave infrared bands (SWIR) in all of their significant band combinations for chlorophyll-*a*. The SWIR band of a Sentinel 2A/MSI image was proven once again important for Chl-*a* estimation (R^2 = 0.7) in Chebara Dam (Kenya) (Kontopoulou et al., 2017) and in particular a second-order polynomial fit was found to be suitable using the reflectance from the difference between the green (B3) and the SWIR-1 (B11) band. Furthermore, Tripathi and Patil (2004) studied 11 representative lakes of Greece (included in our dataset) regarding their Chl-*a* concentrations and managed to establish high correlations between the red and SWIR bands of Landsat 8 images. He et al. (2008) also generated a Chl-*a* three-variable predictive model employing green and SWIR-1 bands and the ratio red/green using EMT+ sensor (R^2 =0.91) in Río Tercero reservoir (Argentina).

Lakes' TSI classification and exploration of the factors affecting its accuracy

In the framework of this study, assessment models of the studied WQPs (TP, Secchi depth) were developed. Then, the Carlson's Trophic State Index (TSI) was applied to assess the trophic status initially of all studied lakes and afterwards separately of natural and artificial ones. TSI can be successfully monitored for lakes using satellite techniques and this methodology has been documented in numerous studies (Papoutsa et al., 2014; Membrillo-Abad et

al., 2016; Rivani and Wicaksono, 2018). Trophic status classification based on satellite-derived TSI of all the cases was coincident with the respective *in-situ* at a percentage of 58.5 % while the 28.5 % of the misclassified cases concerned a deviation at only one (1) trophic class. Satellite TSI calculation independently of cases regarding natural and artificial lakes yielded results that were highly coincident with the *in-situ* derived classes (58.9% and 67.2 %, respectively). Considering the mean depth and the nature of the lakes, deeper (> 5 m) and natural lakes were more successfully classified compared to shallow and artificial ones. Deeper lakes are less affected by the bottom reflectance, fact that is once more verified based on the hereby findings. Light bottom reflection in shallow waters may be a result of the above-water remotely sensed reflectance spectra, hence it cannot be very reliable. Therefore, the estimation of WQPs in shallow waters should be validated using *in-situ* data (Chen et al., 2007).

Concerning the higher TSI misclassification in artificial lakes, it should be noted that TP and Secchi depth are far more variable in reservoirs than in natural lakes (Canfield and Bachmann, 1981). Most models are developed with the assumption that phosphorus is the primary factor limiting algal growth (Kimmel et al., 1990). Nevertheless, there are other nutrients, such as nitrogen, or other factors (e.g. incident light) that may also limit algal production, particularly in reservoirs (Virginia, 2007). The above-mentioned rationales in combination with the fact that in this study TP concentration of artificial lakes has been assessed based on the $TP_{general}$ model, may partly explain the fact that TSI evaluation is less robust in those impoundments.

A significant aspect concerning the contribution of the present study lies in the fact that the study area includes 50 different lake systems of varied chemistry, trophic level, from different regions of Greece and WQ elements collected over different seasons. WQ assessment models have been developed concerning a wide range of limnological conditions with emphasis on whether the lakes are natural or artificial, deep (> 5 m mean depth) or shallow. WQ empirical models are priceless means for trophic status classification for the majority of Greek lakes, especially when *in-situ* data are limited. In addition to their proven predictive performance, it should be noted that- based on the validation processes- they exhibited spatial and temporal stability to variations of the optical properties of the lakes. Furthermore, according to (Loveland and Dwyer 2012), Landsat OLI and ETM+ have similar wavelength ranges and based on the results yielded by (Song et al., 2012), excellent consistency was also found between those sensors in the blue, green, red and NIR regions. Hence, the hereby developed models also accommodates the spectral configuration differences among the used Landsat sensors. However, those empirical models are accompanied by several restrictions such as the accuracy of sampling points' geolocation and the incorporation of many sampling seasons, while the latter plays a crucial role

on TP loadings and Secchi depth values. Moreover, an additional and deeper limnological research is needed mostly oriented towards the primary limiting factors of Chl-*a* production and the predominant sources of turbidity (algal/non-algal), particularly in reservoirs. A wider limnological research would provide valuable information about the lake-wide stratification effects, water movement and other ecosystem-interaction effects on lake water quality, especially for the areas than cannot be accessed and sampled.

2.2.6 Conclusions

This study developed an approach of modelling Greek lakes' water quality by combining EO data (Landsat 7 ETM+ and 8 OLI) with *in-situ* measurements of TP and Secchi depths derived from the application of WFD in Greece. Furthermore, based on our previous study (Markogianni et al., 2020) and the derived Chl-*a* empirical models, WQ assessment models developed herein contribute to the evaluation of the trophic status of all monitored lakes (N=50; National Lake Monitoring Network), by applying the Carlson's trophic index.

Stepwise MLR analyses incorporated, except for Landsat reflectance bands, *in-situ* measurements of water constituents that according to the relevant literature play a role as a proxy of other WQ parameters. Even though estimation of non-optically active constituents of WQ remains a complex challenge for remote sensing, those enhanced analyses managed to explore and highlight the most significant predictors of TP and Secchi depth's values of all lakes but also separately of artificial and natural ones.

According to recent respective literature, even though physical and bio-optical models are considered more robust, they require deep knowledge, collection and parameterization of certain spectral features. Furthermore, even deep learning approaches (belonging to empirical/non-linear methods) still hide issues regarding the appropriate balance between the depth of network and the computational efficiency (Sagan et al., 2020). On the other hand, empirical methods (mostly linear approaches) have the benefit of being easy to implement and straightforward for data processing and in some cases as in (Brewin et al., 2015) proved to outperform a range of bio-optical methods when applied to regional datasets. Based on this perspective, empirical separate models' development (general, natural, artificial) for the assessment of certain WQ parameters (TP, Secchi depth) provides a great opportunity to water resources managers to gain information at any time about the trophic status of any lake in Greece. A reliable prediction of lake trophic status, as the one proposed herein, will further support the monitoring of eutrophication and the drivers of its dynamics, especially

nowadays that lakes are undergone the dual influence of human activities and climate change.

Current approaches for modelling WQ elements in lakes have limited transferability (in space and time). The hereby delivered WQ models may be applicable and deliver fairly acceptable results in lakes outside Greece. However, even though there is a strong possibility those models to be effective only within the borders of Greece, eutrophication has been evolved into such a growing public concern that its investigation and monitoring is considered essential and important even at a country level. In this way, this study supports the aims of WFD and facilitates the continuous water quality monitoring of Greek lakes.

The present study can be extended in different directions; the ultimate goal is the development of a robust tool monitoring WQ parameters in various scales and of a direct and reliable assessment of trophic status for all Greek lakes. However, future work initially includes the harmonization of Sentinel and Landsat images with main aims the investigation of the performance of the hereby developed models if combined with Sentinel images and the minimization of the large time windows (>±7 days) between *in-situ* and satellite data. Moreover, based on the continuous operation of WFD in Greece, at least until 2023, ongoing quality control tests will be conducted to further improve those models' efficiency. Furthermore, since the DOS1 atmospheric correction method has not been validated, one more key priority future action is the application of alternative atmospheric correction methods with principle goal the exploration of their wider effect on models' predictive ability. Additional to hereby utilized methodology, and given the nature of the available data which is non-parametric, authors intend to employ in the near future non-linear methods. Those methods offer, according to literature, great potential for WQ parameter estimation and a sensitivity analysis among several empirical methods would contribute to better understanding of WQ constituents' behavior and possibly to their more accurate assessment. The authors hope that successfully accomplishing all the aforementioned research tasks, on condition the continuous updating of wide WQ datasets, will provide the best opportunity for researchers and public authorities to guide and eventually manage to take sustainably public safety decisions and effective protection measures.

Acknowledgments: Authors wish to thank for the freely shared data used in this study. Those data were collected in the framework of the National Water Monitoring Network for lakes, according to the Joint Ministerial Decision 140384/2011, implemented by The Goulandris Natural History Museum, Greek Biotope/Wetland Center (EKBY). The network is supervised by the Directorate for the Protection and Management of Water Resources of the Special Secretariat for Waters of the Ministry of Environment and Energy.

2.2.7 References

Akbar T. A.; Hassan Q. K.; Achari G. A remote sensing based frame work for predicting water quality of different source waters. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci* **2010**, 34 (3), ISPRS, Calgary, Canada.

Allan, M.G.; Hamilton, D.P.; Hicks, B.J.; Brabyn, L. Landsat remote sensing of chlorophyll a concentrations in central north island lakes of new zealand. *Int. J. Remote Sens.* **2011**, 32, 2037–2055.

Allee, R.; Johnson, J. Use of satellite imagery to estimate surface chlorophyll a and Secchi disc depth of Bull Shoals Reservoir, Arkansas, USA. *Int. J. Remote Sens.* **1999**, 20, 1057–1072.

Alparslan, E.; Coskun, H.G.; Alganci, U. Water quality determination of Küçükçekmece Lake, Turkey by using multispectral satellite data. Sci. World J. 2009, 9, 1215–1229.

American Public Health Association (APHA). Standard Methods for the Examination of Water and Wastewater, **1989**, 17th ed.; American Public Health Association: Washington, DC, USA.

Austin, R. W.; Petzold, T. J. Water colour measurements. **1981**. In J. Gower (Ed.), Oceanography from space, 239–256. New York: Plenum.

Baban, S.M.J. Detecting water quality parameters in the norfolk broads, U.K., using landsat imagery. *Int. J. Remote Sens.* **1993**, 14, 1247–1267.

Barrett, D.C.; Frazier, A.E. Automated Method for Monitoring Water Quality Using Landsat Imagery. *Water* **2016**, *8*, 257. https://doi.org/10.3390/w8060257.

Bartram, J.; Balance, R. Water quality monitoring: a practical guide to the design and implementation of freshwater quality studies and monitoring programmes, 1996, World Health Organization & United Nations Environment Programme, CRC Press.

Bhatti, A. M.; Nasu, S.; Takagi, M.; Nojiri, Y. Assessing the potential of remotely sensed data for water quality monitoring of Coastal and Inland waters. *Research Bulletin of Kochi University of Technology* **2008**, 5(1), 201–207.

Bonansea, M.; Rodriguez, C.; Pinotti, L. Assessing the potential of integrating Landsat sensors for estimating chlorophyll-a concentration in a reservoir. Hydrology Res. 2018, 49, 1608–1617.

Brewin R.J.; Sathyendranath S.; Müller D.; Brockmann C.; Deschamps, P.Y.; Devred, E., R. Doerffer, N. Fomferra, B. Franz, M. Grant, S. Groom. The Ocean Colour Climate Change Initiative: III. A round-robin comparison on in-water bio-optical algorithms. *Remote Sens. Environ*. **2015**, 162, pp. 271-294.

Brezonik, P. L., Menken, K., & Bauer, M. E. Landsat-based remote sensing of lake water quality characteristics, including chlorophyll and colored dissolved organic matter (CDOM). *Lake Reserv Manag* ***2005***, 2, 373-382.

Brivio, P.A.; Giardino, C.; Zilioli, E. Determination of chlorophyll concentration changes in Lake Garda using an image-based radiative transfer code for Landsat TM images. *Int. J. Remote Sens*. **2001**, 22, 487–502.

Buiteveld, H.; Hakvoort, J. H. M.; and Donze, M. The optical properties of

pure water. *SPIE Ocean Optics XII* **1994**, 2258, 174–183.

Busse, L.B.; Simpson, J.C.; Cooper, S.D. Relationships among nutrients, algae, and land use in urbanized southern California streams. *Can. J. Fish. Aquat. Sci.* **2006**, 63, 2621–2638.

Canfield, D. E.; Hodgson L. M. Prediction of Secchi disc depths in Florida lakes: Impact of algal biomass and organic color. *Hydrobiologia* **1983**, 99: 51-60.

Canfield, D., Jr.; Bachmann, R. Prediction of total phosphorus concentrations, chlorophyll a, and secchi depths in natural and artificial lakes. *Can. J. Fish. Aquat. Sci.* **1981**, 38, 414–423.

Carlson, R.E. A trophic state index for lakes. *Limnol. Oceanogr.* **1977**, 22, 361–369.

Chen, J., Quan W. Using Landsat/TM imagery to estimate nitrogen and phosphorus concentration in Taihu Lake, China. IEEE *J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **2012**, 5(1), 273-280. https://doi.org/10.1109/JSTARS.2011.2174339

Chen, Q.; Zhang, Y.; Hallikainen, M. Water quality monitoring using remote sensing in support of the EU water framework directive (WFD): A case study in the Gulf of Finland. *Environ. Monit. Assess.* **2007**, 124, 157–166.

Choubey, V. K. Laboratory experiment, field and remotely sensed data analysis for the assessment of suspended solids concentration and secchi depth of the reservoir surface water, *Int. J. Remote Sens.* **1998**, 19(17), 3349-3360, DOI: 10.1080/014311698214037

Chu, Hone-Jay; Kong, Shish-Jeng; Chang, Chih-Hua. Spatio-temporal water quality mapping from satellite images using geographically and temporally weighted regression. *Int J Appl Earth Obs Geoinf.* **2018**, 65, 1-11, ISSN 0303-2434, https://doi.org/10.1016/j.jag.2017.10.001.

Correll, D. L. Phosphorus: A rate limiting nutrient in surface waters. Poult. Sci. **1999**, 78, 674–82.

Dancey, C.P.; Reidy, J. Statistics without Maths for Psychology, **2007,** 4th ed.; Pearson Education: Harlow, UK.

Dekker, A. G.; Malthus, T. J.; Seyhan, E. Quantitative modeling of inland water quality for high resolution MSS system. *IEEE Trans. Geosci Remote Sens.* **1991**, 29, 89–95.

Dekker, A.; Zamurović-Nenad, Ž.; Hoogenboom, H.; Peters, S. Remote sensing, ecological water quality modelling and in situ measurements: a case study in shallow lakes. *Hydrolog. Sci. J.* **1996**, 41(4), 531–547.

Doña, C., Sánchez M. J., Caselles V., Jose Antonio Domínguez, and Antonio Camacho, Empirical Relationships for Monitoring Water Quality of Lakes and Reservoirs Through Multispectral Images, IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens. **2014**, 7(5).

El-Alem, A.; Chokmani, K.; Laurion, I.; El-Adlouni, S. E. Comparative analysis of four models to estimate chlorophyll-a concentration in case-2 waters using MODerate resolution imaging spectroradiometer (MODIS) imagery. Remote Sens. 2012, 4(8), 2373–2400. doi:10.3390/rs4082373.

Fuller, L.M.; Aichele, S.S.; Minnerick, R.J. Predicting Water Quality by Relating Secchi-Disk Transparency and Chlorophyll a Measurements to Satellite Imagery for Michigan Inland Lakes, August **2002**; US Department of the Interior, US Geological Survey: Denver, CO, USA, 2004.

Gholizadeh, M.; Melesse, A.; Reddi, L. A comprehensive review on water quality parameters estimation using remote sensing techniques. *Sensors* **2016**, 16, 1298.

Giardino, C.; Bresciani, M.; Cazzaniga, I.; Schenk, K.; Rieger, P.; Braga, F.; Matta, E.; Brando, V.E. Evaluation of multi-resolution satellite sensors for assessing water quality and bottom depth of lake garda. *Sensors* **2014**, 14, 24116–24131.

Giardino, C.; Pepe, M.; Brivio, P.A.; Ghezzi, P.; Zilioli, E. Detecting chlorophyll, Secchi disk depth and surface temperature in a sub-alpine lake using Landsat imagery. *Sci. Total Environ*. **2001**, 268, 19–29.

Gordon, H.R. and Morel, A. Remote Assessment of Ocean Color for Interpretation of Satellite Visible Imagery. A Review, Lecture Notes on Coastal and Estuarine Studies, **1983**, R. T. Barber, N. K. Mooers, M. J. Bowman and B. Zeitzschel (eds.), Springer-Verlag, New York, 114 p.

Haddad, K.D.; Harris B.A. Use of remote sensing to assess estuarine habitats. Pages 662-675 in O. T. Magoon; H. Converse; D. Miner; D. Clark and L. T. Tobin, eds Coastal zone **1985**. Proceedings of the 4th symposium on coastal and ocean management. American Society of Civil Engineers, New York.

Hans, H.; Haan, J.; Jordans, R.; Vos, R.; Peters, S.; Rijkeboer, M. Towards airborne remote sensing of water quality in The Netherlands—validation and error analysis. ISPRS J. *Photogramm. Remote Sens.* **2002**, 57(3), 171–183.

He W.; Chen S.; Liu X.; Chen J. Water quality monitoring inland water body through remote sensing – A case study of Guanting Reservoir in Beijing, China. *Front. Environ. Sci. Engin. China* **2008,** 2 (2), 163–71.

Healey, F. P.; Hendzel, L. L. Indicators of phosphorus and nitrogen deficiency in five algae in culture. *Can. J. Fish Aquat. Res.* **1979**, 36, 1364–1369.

Heiskary, S.; Wilson, B. Minnesota Lake Water Quality: Developing Nutrient Criteria, 3rd ed.; Minnesota Pollution Control Agency: St. Paul, MN, USA, **2005**.

Hellweger, F.L.; Schlosser, P.; Lall, U.; Weissel, J.K. Use of satellite imagery for water quality studies in New York Harbor. *Estuarine, Coastal Shelf Sci.* **2004**, 61, 3, 437-448, ISSN 0272-7714, https://doi.org/10.1016/j.ecss.2004.06.019.

Hicks, B.J.; Stichbury, G.A.; Brabyn, L.K.; Allan, M.G.; Ashraf, S. Hindcasting water clarity from Landsat satellite images of unmonitored shallow lakes in the Waikato region, New Zealand. *Environ. Monit. Assess.* **2013**, 185, 7245–7261.

Huang, C.; Guo, Y.; Yang, H.; Li, Y.; Zou, J.; Zhang, M.; Lyu, H.; Zhu, A.; Huang, T. Using remote sensing to track variation in phosphorus and its interaction with chlorophyll-a and suspended sediment. IEEE *J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **2015**, 8, 4171–4180.

IOCCG Remote Sensing of Ocean Colour in Coastal, and Other

Optically-Complex, Waters. **2000**, Sathyendranath, S. (ed.), Reports of the International Ocean-Colour Coordinating Group, No. 3, IOCCG, Dartmouth, Canada.

Isenstein, E. M., Park, M.-H. Assessment of nutrient distribution in Lake Champlain using satellite remote sensing. *J Environ Sci* **2014**, 26, 1831-1836.

Japitana, M.; Burce, M. A Satellite-based Remote Sensing Technique for Surface Water Quality Estimation. *Eng. Technol. Appl. Sci. Res.* **2019**, 9, 3965–3970.

Jiang D., Matsushita B., Setiawan F., Vundo A. An improved algorithm for estimating the Secchi disk depth from remote sensing data based on the new underwater visibility theory, *ISPRS Journal of Photogrammetry and Remote Sensing* **2019**, 152, 13–23.

Kim, S.I.; Kim, H.C.; Hyun, C.U. High resolution ocean color products estimation in Fjord of Svalbard, arctic sea using Landsat-8 oli. *Korean J. Remote Sens.* **2014**, 30.

Kimmel B.L.; Lind O.T.; Paulson L.J. Reservoir primary production K.W.
Thorton, B.L. Kimmel, F.E. Payne (Eds .), Reservoir Limnology. Ecological Perspectives, John Wiley and Sons, New York, 1990.

Kloiber, S. M.; Brezonik, P. L.; Olmanson, L. G.; and Bauer M. E. A procedure for regional lake water clarity assessment using Landsat multispectral data. *Remote Sens. Environ.* **2002**, 82, 38-47.

Kontopoulou, E. ; Kolokoussis, P. ; Karantzalos, K., Water quality estimation in Greek lakes from Landsat 8 multispectral satellite data. 10th World Congress of the European Water Resources Association (EWRA) on Water Resources and Environment (EWRA2017), Athens, Greece, 5-9 July **2017**. European Water 2017 No.58 pp.191-196 ref.12

Kratzer, S.; Kyryliuk, D.; Edman, M.; Philipson, P.; Lyon, S.W. Synergy of Satellite, In Situ and Modelled Data for Addressing the Scarcity of Water Quality Information for Eutrophication Assessment and Monitoring of Swedish Coastal Waters. *Remote Sens.* **2019**, *11*, 2051. https://doi.org/10.3390/rs11172051.

Kutser, T.; Arst, H.; Miller, T.; Käärmann, L.; Milius, A. Telespectrometrical estimation of water transparency, chlorophyll-a and total phosphorus concentration of Lake Peipsi. *Int. J. Remote Sens.* **1995**, 16, 3069–3085.

Lathrop, R. Landsat Thematic Mapper monitoring of turbid inland water quality. *Photogramm. Eng. & Rem. S.* **1992**., 58(4), 465–470.

Lillesand, T. M.; Kiefer R. W. Remote sensing and image interpretation. 4th ed. John Wiley and Sons, New York, NY, **2000**, 724pp.

Lim, J.; Choi, M. Assessment of water quality based on Landsat 8 operational land imager associated with human activities in Korea. *Environ Monit Assess* **2015**, 187, 384, https://doi.org/10.1007/s10661-015-4616-1.

Lind, O.T. The effect of non-algal turbidity on the relationship of Secchi depth to chlorophyll a. *Hydrobiologia* ***1986***, 140, 27–35.
https://doi.org/10.1007/BF00006726

Lorenzen, M. W. The use of chlorophyll-secchi disk relationships. *Limnol. Oceanogr.* **1980**. 25: 371- 372.
Loveland, T.R., Dwyer, J.L. Landsat: building a strong future. *Remote Sens. Environ.* **2012**, 122 (1), 22–29.

Malve, O.; Qian, S. S. Estimating nutrients and chlorophyll a relationships in Finnish lakes. *Environ. Sci. Technol.* **2006**, 40, 7848–7853.
Markogianni, V.; Dimitriou, E.; Karaouzas, I. Water quality monitoring and assessment of an urban Mediterranean lake facilitated by remote sensing applications. *Environ. Monit. Assess.* **2014**, 186, 5009–5026.

Markogianni, V.; Kalivas, D.; Petropoulos, G.; Dimitriou, E. An appraisal of the potential of Landsat 8 in estimating chlorophyll-*a*, ammonium concentrations and other water quality indicators. *Remote Sens.* **2018**, 10, 1018.

Markogianni, V.; Kalivas, D.; Petropoulos, G.P.; Dimitriou, E. Estimating Chlorophyll-*a* of Inland Water Bodies in Greece Based on Landsat Data. *Remote Sens.* **2020**, *12*, 2087. https://doi.org/10.3390/rs12132087.

Matthews, M.W. A current review of empirical procedures of remote sensing in Inland and near-coastal transitional waters. *Int. J. Remote Sens.* **2011**, 32, 6855–6899.

Mavromati, E.; Kagalou, I.; Kemitzoglou, D.; Apostolakis, A.; Seferlis, M.; Tsiaoussi, V. Relationships among land use patterns, hydromorphological features and physicochemical parameters of surface waters: WFD lake monitoring in Greece. *Environ. Process.* **2018**, 5 (Suppl. S1), 139–151.

McCullough, I.M.; Loftin, C.S.; Sader, S.A. Combining lake and watershed characteristics with Landsat TM data for remote estimation of regional lake clarity. *Remote Sens. Environ.* **2012**, 123, 109–115.

McKinna, L.; Werdell, P. Approach for identifying optically shallow pixels when processing ocean-color imagery. *Opt. Express* **2018**, 26, A915–A928.
Megard, R.O.; Settles, J.C.; Boyer, H.A.; Combs, W.S. Light, Secchi Disks, and Trophic States. *Limnol. Oceanogr.* **1980**, 25, 373–377.

Membrillo-Abad, Alejandra-Selene; Torres-Vera, Marco-Antonio; Alcocer, Javier; Prol-Ledesma, Rosa Ma.; Oseguera, Luis A.; Ruiz-Armenta, Juan Ramón. Trophic State Index estimation from remote sensing of lake Chapala, México.*Rev. mex. cienc. Geol* **2016**, 33(2),183-191. ISSN 2007-2902.

Michalak, A. Study role of climate change in extreme threats to water quality. Nature 2016, 535, 349–350. https://doi.org/10.1038/535349a.

Moore, T. S.; Dowell, M. D.; Bradt, S.; Verdu A. R. An optical water type framework for selecting and blending retrievals from bio-optical algorithms in lakes and coastal waters, *Remote Sens. Environ.* **2014**, *143*, 97-111, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2013.11.021.

Morel, A.; Gordon, H. R. Report of the working group on water colour. Boundary-Layer *Meteorol.* **1980**, 18, 343-355

Moses A.; S., Janaki, L.; Joseph, S.; Kizhur K., R. Determining the spatial variation of phosphorus in a lake system using remote sensing techniques, *Lakes Reserv.: Res. Manag.*

2014, 19(1), 24-36. doi: 10.1111/lre.12054.

Nauman E. The Scope of chief problems of regional limnology. *Int.ReviewGes.Hydrobiol.* **1929,** 21:423.

Nelson, S.A.C.; Soranno, P.A.; Cheruvelil, K.S.; Batzli, S.A.; Skole, D.L. Regional Assessment of lake water clarity using satellite remote sensing. *J. Limnol.* **2003**, 62, 27–32.

Odermatt, D., Gitelson, A. Vittorio, E.V., Schaepman, M. Review of constituent retrieval in optically deep and complex waters from satellite imagery. *Remote Sensing of Environment* **2012,** 118, 116-126, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2011.11.013.

Ohammad M.M. Alsahli, Majid Nazeer. Spatiotemporal variability of secchi depths of the North Arabian Gulf over the last two decades. *Estuar. Coast. Shelf Sci. 2021*, 260, 107487, ISSN 0272-7714, https://doi.org/10.1016/j.ecss.2021.107487.

Olmanson, L.G.; Bauer, M.E.; Brezonik, P.L. A 20-year Landsat water clarity census of Minnesota's 10,000 lakes. *Remote Sens. Environ.* **2008**, 112, 4086–4097.

Olmanson, L.G.; Kloiber, S.M.; Bauer, M.E.; Brezonik, P.L. Image Processing Protocol for Regional Assessments of Lake Water Quality; University of Minnesota: St. Paul, MN, USA, **2001**.

Ouma O. Yashon, Noor Kimutai, Herbert Kipkemoi. Modelling Reservoir Chlorophyll-a, TSS, and Turbidity Using Sentinel-2A MSI and Landsat-8 OLI Satellite Sensors with Empirical Multivariate Regression. *J. Sens.* ***2020***, vol. 2020, Article ID 8858408. https://doi.org/10.1155/2020/8858408

Page, B.P., Olmanson, L.G., Mishra, D.R. A harmonized image processing work- flow using Sentinel-2/MSI and Landsat-8/OLI for mapping water clarity in opti- cally variable lake systems. *Remote Sens. Environ*. **2019**, 231, 111284.

Papoutsa, C.; Akylas, E.; Hadjimitsis, D. Trophic State Index derivation through the remote sensing of Case-2 water bodies in the Mediterranean region. *cent.eur.j.geo.* **2014,** 6**,** 67–78. https://doi.org/10.2478/s13533-012-0161-4

Peckham, S.D.; Chipman, J.W.; Lillesand, T.M.; Dodson, S.I. Alternate stable states and the shape of the lake trophic distribution. *Hydrobiologia* **2006**, 571, 401–407.

Poor, N. D. Effect of lake management efforts on the trophic state of a subtropical shallow lake in Lakeland, Florida, USA. *Water Air Soil Pollut.* **2010**, 207, 333–347.

Pozdnyakov, D.; Shuchman, R.; Korosov, A.; Hatt, C. Operational algorithm for the retrieval of water quality in the Great Lakes. *Remote Sens. Environ*. **2005,** 97, 352–370.

Prasad, A.D.; Siddaraju. Carlson's Trophic State Index for the assessment of trophic status of two lakes in Mandya district. *Adv. Appl. Sci. Res.* **2012**, *3*.

Ritchie, J. C.; Cooper, C. M.; Schiebe, F. R. The relationship of MSS and TM digital data with suspended sediments, chlorophyll, and temperature in Moon Lake, Mississippi. *Remote Sens. Environ*. **1990**, 33(2), 137–148.

Rivani A.; Wicaksono P. Water Trophic Status Mapping of Tecto-

Volcanic Maninjau Lake during Algae Bloom using Landsat 8 OLI Satellite Imagery. *IEEE International Conference on Aerospace Electronics and Remote Sensing Technology (ICARES)*, Bali, **2018**, pp. 1-7, doi: 10.1109/ICARES.2018.8547055.

Sagan Vasit, Peterson T. K., Maimaitijiang M., Sidike P., Sloan J., Greeling A., B., Samar Maalouf, Craig Adams. Monitoring inland water quality using remote sensing: potential and limitations of spectral indices, bio-optical simulations, machine learning, and cloud computing, *Earth-Science Reviews* **2020**, 205, 103187, ISSN 0012-8252, https://doi.org/10.1016/j.earscirev.2020.103187.

Shafique, N A.; Fulk, F A.; Autrey B C.; Flotemersch, J E. Hyperspectral remote sensing of water quality parameters for large rivers in the Ohio river basin. Presented at Ohio River Basin Consortium for Research and Education, November 5-7, 2003.

Marietta, O. H. Smith, V. H. The nitrogen and phosphorus dependence of algal biomass in lakes: An empirical and theoretical analysis. *Limnol. Oceanogr*. **1982**, 27(6), 1101– 1112.

Søballe, D. M.; Kimmel, B. L.; Kennedy, R. H.; Gaugush, R. F. Reservoirs. In C. T. Hackney, S. M. Adams, & W. H. Martin (Eds.), *Biodiversity of the Southeastern United States: Aquatic communities* ***1992*** (pp. 421–474). New York: Wiley.

Song K.; Zhang B.; Wang Z.; Li F.; Duan H.; Guo Y. Water TOC and TP concentration estimation using Landsat TM data with empirical algorithms in Chagan lake, China. IEEE International Conference, Geosciences and Remote Sensing Symposium **2006**, IGRASS.

Song, K.; Li, L.; Li, S; *et al.* Hyperspectral Remote Sensing of Total Phosphorus (TP) in Three Central Indiana Water Supply Reservoirs. *Water Air Soil Pollut* **223,** 1481–1502 (2012). https://doi.org/10.1007/s11270-011-0959-6

Spyrakos, Evangelos ; O'Donnell, Ruth; Hunter, Peter D.; Miller, Claire; Scott, Marian; Simis, Stefan G.H.; Neil, Claire; Barbosa, Claudio C.F.; Binding, Caren E.; Bradt, Shane; Bresciani, Mariano; Dall'Olmo, Giorgio; Giardino, Claudia ; Gitelson, Anatoly A.; Kutser, Tiit; Li, Lin; Matsushita, Bunkei; Martinez-Vicente, Victor; Matthews, Mark W.; Ogashawara, Igor; Ruiz-Verdú, Antonio; Schalles, John F.; Tebbs, Emma; Zhang, Yunlin; Tyler, Andrew N. / Optical types of inland and coastal waters. In: *Limnology and Oceanography* **2018**, 63(2), 846-870.

Topp, S.; Pavelsky, T.; Jensen, D.; Simard, M.; Ross, M. Research trends in the use of remote sensing for inland water quality science: moving towards multidisciplinary applications. *Water* **2020**, 12 (1), 169. https://doi.org/10.3390/w12010169.

Tripathi, N.K.; Patil, A.A. Spectral characterization of aquatic nutrients and water quality parameters in marine environment. *Bibliogr. Inform*. **2004**, 15, 25–31.

Tyler, A.N.; Svab, E.; Preston, T.; Presing, M.; and Kovacs, W.A. Remote sensing of the water quality of shallow lakes: a mixture modeling approach to quantifying phytoplankton in water characterized by high suspended sediment. *Int. J. Remote Sens*. **2006,** 27, 1521–1537.

Uusitalo, R.; Yli-Halla, M.; Turtola, E. Suspended soil as a source of

potentially bioavailable phosphorus in surface runoff waters from clay soils. *Water Res.* **2000**, 34, 2477–2482.

Virginia Water Resources Research Institute and State University Blacksburg. Nutrients in lakes and reservoirs-Aliterature review for use in nutrient criteria development. VWRRC Special report SR34-2007. Virginia, **2007**. Downloaded from: http://www.vwrrc.vt.edu (Accessed on 10/10/2020).

Wang, Y. P.; Xia, H.; Fu, J.; Sheng, G. Y. Water quality change in reservoirs of Shenzhen, China: Detection using LANDSAT/TM data. *Sci. Total Environ.* **2004**, 328, 195–206.

Wu, C. F.;Wu, J. P.; Qi, J. G.; Zhang, L. S.; Huang, H. Q.; Lou, L. P.; et al. Empirical estimation of total phosphorus concentration in the mainstream of the Qiantang River in China using Landsat TM data. *Int. J. Remote Sens.* **2010**, 31(9), 2309–2324.

Wu, G.; De Leeuw, J.; Skidmore, A.K.; Prins, H.H.T.; Liu, Y. Comparison of MODIS and Landsat TM5 images for mapping tempo–spatial dynamics of Secchi disk depths in Poyang Lake National Nature Reserve, China. *Int. J. Remote Sens.* **2008**, 29, 2183–2198.

Zhang Yibo, Zhang Yunlin, Shi Kun, Zhou Yongqiang, Li Na, Remote sensing estimation of water clarity for various lakes in China, *Water Research* ***2021***, 192, 116844, ISSN 0043-1354, https://doi.org/10.1016/j.watres.2021.116844.

Zheng, L.; An, Z.; Chen, X.; Liu, H. Changes in Water Environment in Erhai Lake and Its Influencing Factors. *Water* **2021**, *13*, 1362. https://doi.org/10.3390/w13101362

Zhou Yadong, Liu Hui, He Baoyin, Yang Xiaoqing, Feng Qi, Kutser Tiit, Chen Feng, Zhou Xinmeng, Xiao Fei, Kou Jiefeng. Secchi Depth estimation for optically-complex waters based on spectral angle mapping - derived water classification using Sentinel-2 data, *Int. J. Remote Sens.* ***2021***, 42:8, 3123-3145, DOI: 10.1080/01431161.2020.1868606.

2.3 Landsat-based Lake Water Quality Monitoring: How Transferable are the WQ Algorithms to Sentinel 2 images?

2.3.1 Introduction

This chapter's section basic aim is to explore initially whether Landsat-based empirical WQ algorithms can be efficiently applied to Sentinel 2 images and then whether the combined use of multi-sensor data improves the algorithms' prediction accuracy. Furthermore, independently from whether there is some improvement or not, the ultimate goal of this combined approach is to decide whether multi-sensor images could be used with at least equally reliable results as those accrued from only-one sensor's utilization.

In previous chapters, the methodology regarding the development of WQ models has been fully described. In particular, image data from Landsat 7 ETM+ and 8 OLI were combined with simultaneous *in-situ* WQ data during 2013-2016 while afterwards the implementation of MLR analyses resulted in the generation of quantitative models of Chl-*a*, Secchi disk depths and Total Phosphorus concentrations of 50 Greek lake water bodies. Proposed WQ models have been developed separately for natural-only and artificial-only lakes while *in-situ* dataset of year 2018 was used to further validate their efficiency. In this dissertation, Landsat 7 ETM+ and 8 OLI sensors have been incorporated in an effort to increase the temporal range of available and useful data, since the majority of *in-situ* measurements were recorded during 2013-2015 and together the sensors provide four (4) satellite images for every 32 days (Pedreros-Guarda et al., 2021). However, we investigated whether there is a possibility to arise issues from the combined use of different Landsat sensors. Based on relevant literature review, Landsat 7 ETM+ and Landsat 8 OLI images have similar spatial resolution (30 m), are statistically comparable and homogeneous over WQ sample sites (Wang et al., 2020) while both have similar spectral band placements (Table 2.3.2-1) for the Blue (ETM+ band 1, 0.45–0.52 μm; 8 OLI band 2, 0.45– 0.51 μm) and Green bands (ETM+ band 2, 0.52–0.60 μm; 8 OLI band 3: 0.53–0.59 μm). Differences are particularly observed in the NIR (ETM+ Band 4, 0.76–0.90 μm; 8 OLI Band 5, 0.85–0.88 μm) and to a lesser extent in Red bands (ETM+ Band 3, 0.63–0.69 μm; 8 OLI Band 4, 0.64–0.67 μm) (Deutsch et al., 2018). Furthermore, an effort trying to assess CDOM concentrations and water clarity in oligotrophic lakes and reservoirs of Minnesota (Olmanson et al. 2016) by using Landsat 7 ETM+ and 8 OLI sensors, indicated once again that are generally comparable.

The successful launch of Sentinel 2 in June 2015 and the simultaneous provision of image data with those of Landsat 8 OLI offered great opportunities for long term high-frequency WQ monitoring (Mandanici and Bitelli, 2016) through building time-series. The Sentinel-2 mission carries two

satellites, Sentinel-2A and Sentinel-2B. They are both equipped with identical Multispectral Instruments (MSI) capable of acquiring data at 13 bands at different spatial resolutions (between 10 m and 60 m) while the revisit frequency of each satellite is 10 days.

According to Deutsch et al. (2018) the transferability of WQ algorithms across sensors remains poorly examined, while a number of conceptual and technical challenges may accrue originating from their orbital, spatial and spectral differences. Towards this direction, Sentinel 2 images of 2018 with concurrent dates with those of field measurements were utilized to facilitate a WQ models' efficiency evaluation and comparison with respective Landsat's validation results.

Additionally, another effort has been made to improve WQ models' quantification capability through the combined use of Landsat (7 ETM+, 8 OLI) and Sentinel 2 images, while the selection of each image for each case was based on the acquisition date depending on the corresponding sampling one. In this way, when Sentinel 2 sensor is combined especially with the Landsat 8 OLI, the revisit time is significantly reduced to 2-3 days globally (Li et al., 2021). Mandanici and Bitelli (2016) highlighted some potentials and challenges deriving from the joint use of Landsat and Sentinel 2 sensors; they observed a significant match between the corresponding spectral bands, however differences in the recorded radiometric values were also present. What is important though, concerning those differences, is the application and the approach adopted to implement multi-sensor time series analyses. On one hand, may empirical approaches based on multispectral indices be more affected by the problem (Werff and Meer, 2016) but when methods and processing are applied separately on every image and the training is also independent, as in our case, results are less affected (Mandanici and Bitelli, 2015; 2016). The independent elaboration of Landsat and Sentinel 2 images in the framework of this study, did not require the implementation of a resampling procedure, which is essential mostly in change detection analyses, based on the different spatial resolution of the two sensors (Landsat 30 vs. Sentinel 10 m).

2.3.2 Methodology

2.3.2.1 In-situ/Remote-sensing data and pre-processing

The selected *in-situ* data used herein includes Chlorophyll-*a*, Secchi depth measurements and TP concentrations of 2018 in different dates throughout the monitored lake stations (surveillance and operational). The available *in-situ* WQ dataset includes 136 Chl-*a* measurements, 218 Secchi disk depth values and 88 TP concentrations which are freely accessible and were downloaded from the EKBY's site (Goulandris Natural History Museum,

Greek Biotope/Wetland Centre (http://biodiversity-info.gr/index.php/el/lakes-data#!IMGP4731; in Greek).

Forty-nine (49) Landsat 7 ETM+ and 8 OLI images of 2018 were already downloaded from the USGS (United States Geological Survey) Data Centre (https://earthexplorer.usgs.gov/) in the context of our previous study (Markogianni et al., 2022). Moreover, forty-four (44) Sentinel 2 images of 2018, with concurrent dates to sampling ones, up to ± 7 days, were also downloaded from the Copernicus open access hub (https://scihub.copernicus.eu/dhus/#/home). Some Landsat and Sentinel 2 images have been used twofold or more depending on the studied WQ element (Figure 2.3.2-1) and the sampling date connected to more than one parameter. The majority of *in-situ* WQ data and by extension of satellite images, are detected during summer months whereas Secchi depths were measured during the whole year (Figure 2.3.2-1). The selection of the Sentinel 2 imagery was based on the tiling grid which is available by the ESA (https://sentinel.esa.int/web/sentinel/missions/sentinel-2/data-products) as a KML file, providing unique IDs for each tile (100*100 km^2 ortho-images in UTM/WGS84 projection).

What is important concerning the pre-processing process is that images from both sensors underwent the same process, to minimise possible discrepancies originating from the correction (Mandanici and Bitelli, 2016). Hence, Sentinel 2 images were subjected to the same pre-processing procedure as the Landsat ones (fully described by Markogianni et al., 2022) and more particularly, they were imported in the semi-automatic classification plugin (SCP) of the free and open-source cross-platform desktop Quantum Geographic Information System (Q-GIS), v. 3.6.3-Noosa to perform: (a) conversion of images from digital numbers (DN) to top-of-atmosphere reflectance (TOA), (b) atmospheric correction by using the DOS1 method (applied to all bands except for thermal ones), and (c) the creation of a band stack set for each image. The band stack set and and bandwidths of each satellite sensor is presented in Table 2.3.2-1.

Table 2.3.2-1. Band stack sets and bandwidths (BW) of Landsat 7 ETM+, 8 OLI and Sentinel 2 sensors.

Sensor	Blue	BW (µm)	Green	BW (µm)	Red	BW (µm)	NIR	BW (µm)	SWIR1	BW (µm)	SWIR2	BW (µm)
L7 ETM+	B1	0.441-0.514	B2	0.519-0.601	B3	0.631-0.692	B4	0.772-0.898	B5	1.547-1.749	B7	2.064-2.345
L8 OLI	B2	0.452-0.512	B3	0.533-0.590	B4	0.636-0.673	B5	0.851-0.879	B6	1.566-1.651	B7	2.107-2.294
Sentinel 2	B02	0.458-0.523	B03	0.543-0.578	B04	0.650-0.680	B08	0.785-0.899	B11	1.565-1.655	B12	2.100-2.280

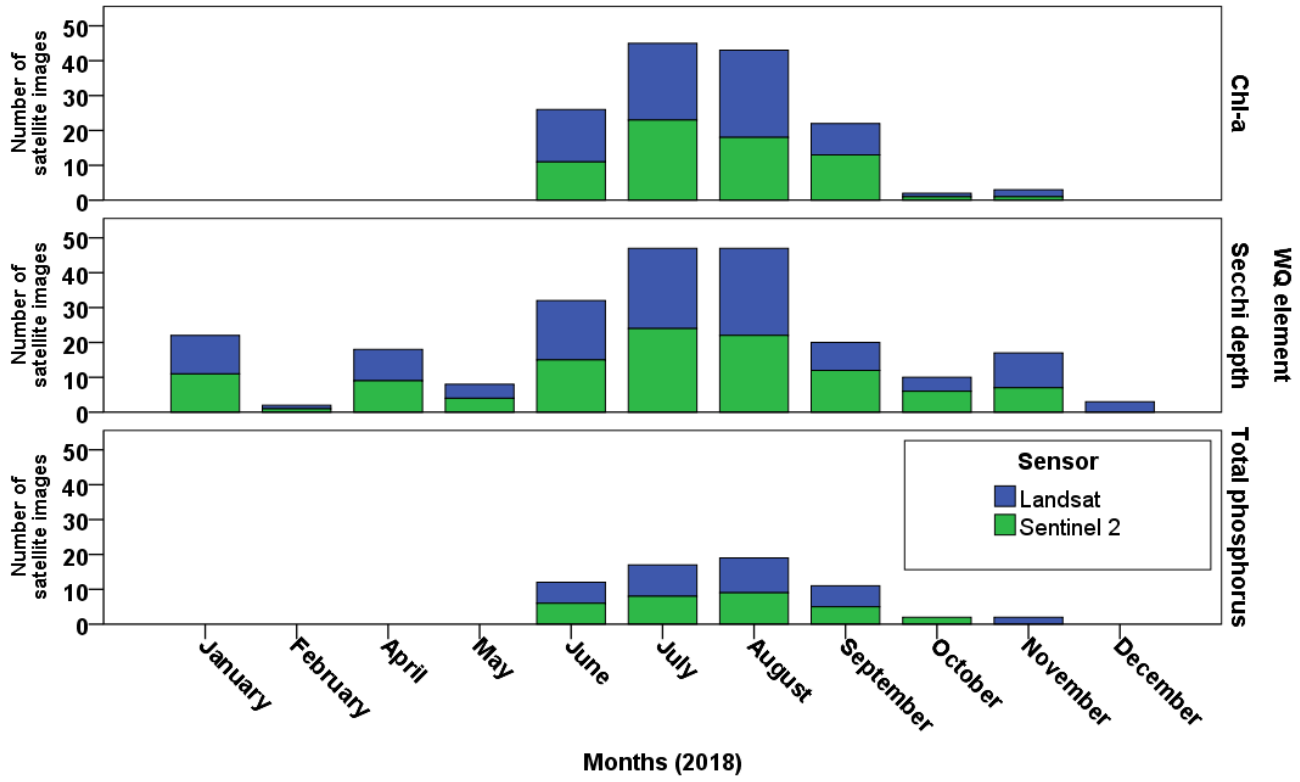


Figure 2.3.2-1. Number of Landsat 7 ETM+/8 OLI and Sentinel 2 images of 2018, grouped by studied WQ element.

2.3.2.2 Comparison of sensors' performance and validation

WQ models developed by Markogianni et al., 2020 (Chl-*a*); 2022 (Secchi depth; TP) were initially applied to Landsat 7+ETM/ 8 OLI and then to Sentinel 2 images of 2018, based on the corresponding *in-situ* dataset. Application of WQ models includes Chl-*a* models (General- Equation 2.3.2-1; Natural- model-Equation 2.3.2-2; Artificial model -Equation 2.3.2-3), Secchi depth models (General-Equation 2.3.2-4; Natural model- Equation 2.3.2-5; Artificial model -Equation 2.3.2-6) and TP models (General-Equation 2.3.2-7; Natural model -Equation 2.3.2-8). Concerning the total phosphorus WQ element, it should be noted that no statistically strong model was delivered for TP quantification in artificial lakes (Markogianni et al., 2022).

$$\log Chla = 3.599 - 0.63 * \left(\frac{blue}{red}\right) - 2.183 * \left(\frac{\ln red}{\ln swir2}\right) \quad (2.3.2\text{-}1)$$

$$\log Chla = 4.443 - 1.421 * \left(\frac{blue}{green}\right) - 3.454 * \left(\frac{\ln red}{\ln swir2}\right) + 1.304 * \left(\frac{red}{green}\right) \quad (2.3.2\text{-}2)$$

$$\log Chla = 2.919 - 2.011 * \left(\frac{\ln red}{\ln swir1}\right) + 1.449 * \left(\frac{red}{green}\right) - 1.441 * \left(\frac{\ln red}{\ln blue}\right) \quad (2.3.2\text{-}3)$$

$$SQRT(Secchi)_{general} = 1.215 - 2.479 * \left(blue + red + \frac{red}{blue}\right) + 3.394 * \left(\frac{lngreen}{lnswir2}\right. \quad (2.3.2\text{-}4)$$

$$SQRT(Secchi)_{natural} = 1.172 - (1.003 * logchl - a) - (1.031 * logred) \quad (2.3.2\text{-}5)$$

$$SQRT(Secchi)_{artificial} = 3.927 - 1.365 * \left(\frac{green}{blue}\right) - 0.318 * \left(\frac{red}{swir1}\right) - 0.361 * logchla \quad (2.3.2\text{-}6)$$

$$LogTPgeneral = -1.425 + 0.452 * logChla - 0.573 * \left(\frac{lnred}{lnswir1}\right) \quad (2.3.2\text{-}7)$$

$$LogTPnatural = -0.633 - (0.704 * logSecchi) - 0.392 * \left(\frac{green}{red}\right) \quad (2.3.2\text{-}8)$$

All *in-situ* datasets (general, natural, artificial) concern lakes with mean depth higher than 5 meters to surely avoid the bottom reflectance noise

(McKinna and Werdell, 2018), but in case of TP, where *in situ* data are fewer than those concerning Secchi depths and Chl-*a* concentrations, the aforementioned depth criterion was set to 3.5 m. Additionally, in another effort to improve WQ models' quantification capability, we established a combined use of Landsat (7 ETM+, 8 OLI) and Sentinel 2 images. WQ models were once again applied to multi-sensor images while the selection of each image was based on the smallest time window between the satellite acquisition and *in-situ* date. In cases where this difference was the same for Landsat and Sentinel sensors, the Landsat image was selected as the WQ models are Landsat-developed and was hypothesized to be more effective compared to ones employing Sentinel images. A resampling procedure has not been performed since each satellite image was separately elaborated depending on the best matching date. Furthermore, the WQ models' performance- depending on the sensor used- was based on the Spearman's (r) correlation coefficient and the error metrics Mean Error (ME), Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE) and Normalized Root Mean Squared Error (NRMSE). MAPE metric is calculated based on the following equation (2.3.2-9):

$$\boldsymbol{MAPE} = \frac{1}{n}\sum_{t=1}^{t=n}\frac{|y'-y|}{y} * 100\% \qquad (2.3.2\text{-}9)$$

where y' is forecasted value, y is the true value and n is the total number of values in the dataset. Furthermore, Lewis (1982) created a table (Table 2.3.2-2) containing typical MAPE values and their interpretation concerning the forecasting potential. MAPE's greatest disadvantage is that the absolute percentage error distribution -characterised by having only positive values with no upper bound-usually has a right or positive skew brought about by the presence of outlier values to this side of the distribution (Moreno et al., 2013). Hence, if the denominator is extremely small or large, the MAPE value adopts the same behaviour.

Table 2.3.2-2. Typical MAPE values and interpretation (Lewis, 1982, p. 40).

MAPE	Interpretation
<10	Highly accurate forecasting
10-20	Good forecasting
20-50	Reasonable forecasting
>50	Inaccurate forecasting

In view of the aforementioned limitation of MAPE, NRMSE is used additionally and comparatively; NRMSE (Equation 2.3.2-10) is an extension of RMSE and often utilized to compare different datasets or predictive models of different scales (e.g. different units as in our case) while it has been calculated by using the range of the true values (difference of minimum and maximum values; Equation 2.3.2-10). Furthermore, it takes values 0-1.
Low values of all error metrics (ME, MAPE, RMSE and NRMSE) indicate the good performance of models.

$$\boldsymbol{NRMSE} = \frac{RMSE}{Ymax - Ymin} \qquad (2.3.2\text{-}10)$$

2.3.3 Results

2.3.3.1 Chlorophyll-a

Application of Chl-a_{general} model in Landsat-only and Sentinel-only images yielded similar results with a light superiority of Landsat-employing model based on ME (mean error) and RMSE values (Table 2.3.3-1). Median MAPE value of Sentinel-2 model touched the upper threshold for reasonable forecasting (50 %; Lewis, 1982) while the corresponding value of Landsat model slightly surpassed it. Combined utilization of Landsat and Sentinel 2 images has not indicated any further improvement of Chl-a_{general} model since all RMSE, NRMSE and MAPE values were higher than those accrued from one-sensor based applications (Figure 2.3.3-1; Figure 2.3.3-2).

As far as the Chl-a_{natural} model is concerned (Table 2.3.3-2), results are different. Sentinel-2 model is superior to the corresponding Landsat one based on RMSE, NRMSE and median MAPE values. Only the value of ME resulting from the utilization of Landsat images is significantly lower compared to Sentinel (-0.12 vs. 2.9 μg/l). Furthermore, application of Chl-a_{natural} model in joined Landsat and Sentinel 2 images has not managed to increase its performance since almost all values of error metrics are higher than those resulted from the independently employment of either Landsat or Sentinel 2 data.

Concerning the Chl-$a_{\text{artificial}}$ model (Table 2.3.3-3), values of ME and RMSE are lower with Landsat-employed data (Figure 2.3.3-1), indicating a better performance of this model compared to that employing Sentinel 2 or mixed satellite data.

Regarding general and natural models' applications, all correlations among *in-situ* and satellite-derived values were statistically significant with Sentinel 2 data presenting the highest values (general: r=0.71; natural: r=0.72), followed by Landsat (general: r=0.6; natural: r=0.697) and mixed satellites

(general: r=0.54; natural: r=0.64). Values of coefficient of correlation were similar for all applications in artificial lakes.

Table 2.3.3-1. Values of error metrics regarding the sensor-based applications of Chl-*a* general model (Underlined value indicates the lowest value among all three cases; ** Correlation is significant at the 0.01 level (2-tailed).

Chl-$a_{general}$ model	N	SPEARMAN r	Average In Situ Chl-*a* (μg/L)	Average Satellite Chl-*a* (μg/L)	ME (μg/L)	RMSE (Chl-*a*; μg/L)	NRMSE	MEDIAN MAPE (%)
Landsat 7 ETM+/8 OLI	74	0.601**	7.9	5.5	2.5	14.6	0.149	54.5
SENTINEL 2	67	0.711**	8.3	4.6	3.7	14.9	0.152	50
MIXED SATELLITES	67	0.54**	8.3	4.9	3.4	15.9	0.162	57.8

Table 2.3.3-2. Values of error metrics regarding the sensor-based applications of Chl-*a* natural model (** Correlation is significant at the 0.01 level (2-tailed).

Chl-$a_{natural}$ model	N	SPEARMAN r	Average In Situ Chl-*a* (μg/L)	Average Satellite Chl-*a* (μg/L)	ME (μg/L)	RMSE (Chl-*a*; μg/L)	NRMSE	MEDIAN MAPE (%)
Landsat 7 ETM+/8 OLI	28	0.697**	13.9	14	-0.12	21.5	0.22	57.9
SENTINEL 2	26	0.72**	14.7	11.8	2.9	16.4	0.17	42
MIXED SATELLITES	26	0.64**	14.7	12.1	2.6	24.1	0.25	52.4

Table 2.3.3-3. Values of error metrics regarding the sensor-based applications of Chl-*a* artificial model (Underlined value indicates the lowest value among all three cases; ** Correlation is significant at the 0.01 level (2-tailed).

Chl-$a_{artificial}$ model	N	SPEARMAN r	Average In Situ Chl-*a* (μg/L)	Average Satellite Chl-*a* (μg/L)	ME (μg/L)	RMSE (Chl-*a*; μg/L)	NRMSE	MEDIAN MAPE (%)
Landsat 7 ETM+/8 OLI	41	0.59**	4.2	2.5	0.95	3.7	0.18	49
SENTINEL 2	33	0.57**	4.6	2.03	2.6	7.74	0.17	41.8
MIXED SATELLITES	40	0.57**	4.3	2.3	2.1	7.3	0.16	53.5

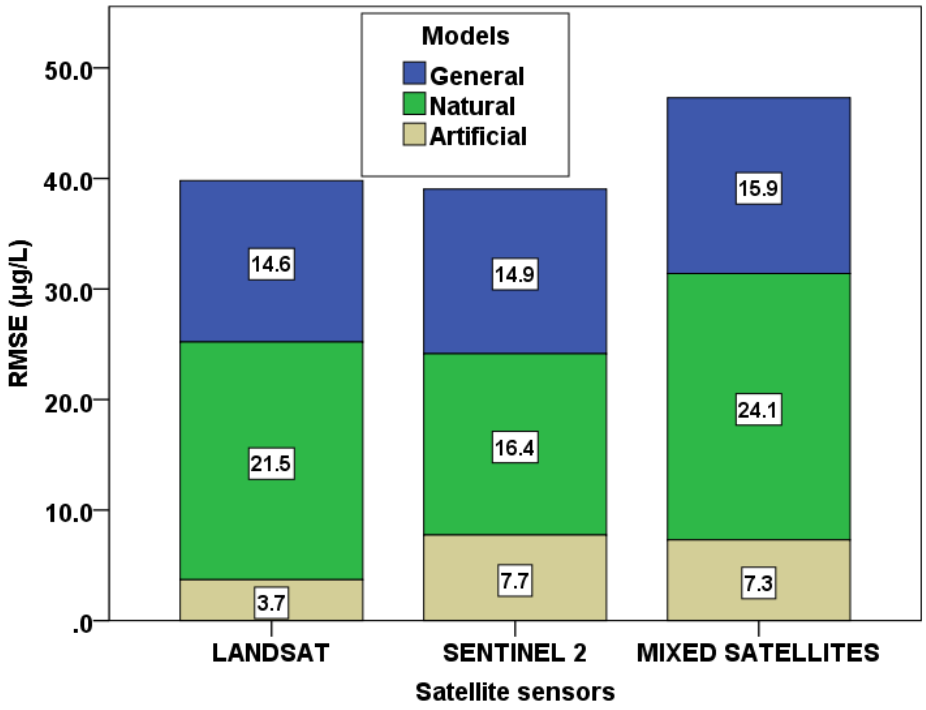


Figure 2.3.3-1. Stacked bars illustrating RMSE values of Chl-*a* concentrations (µg/l) per satellite sensor.

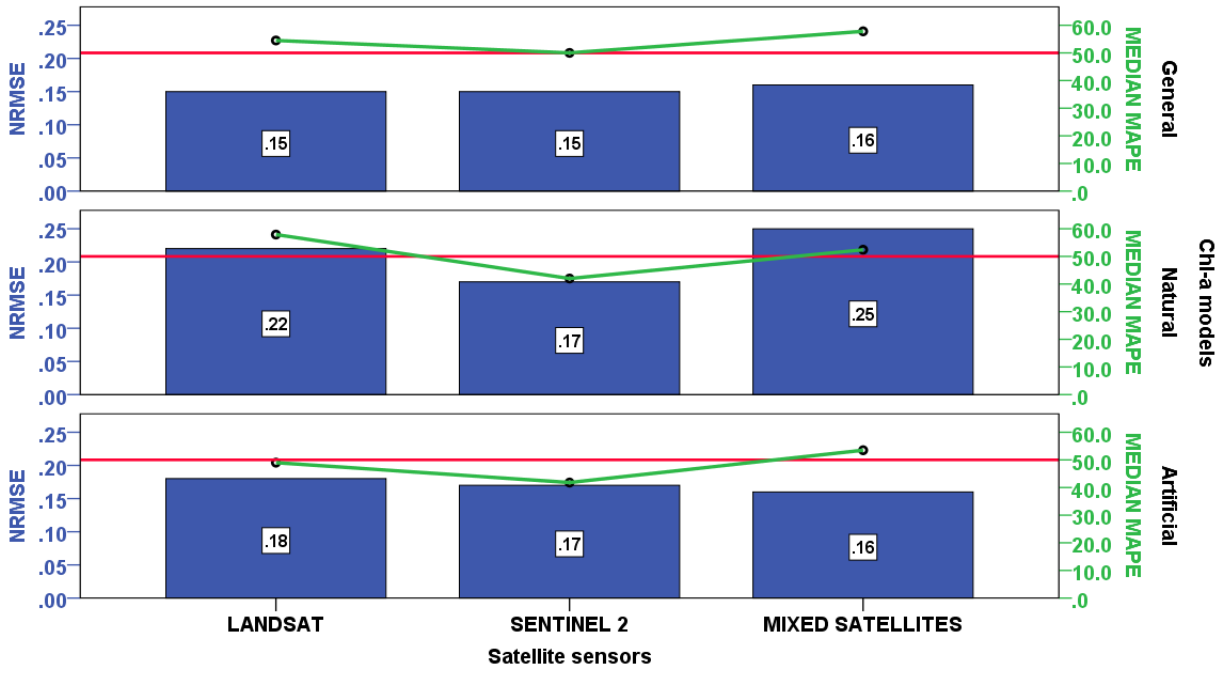


Figure 2.3.3-2. NRMSE and median MAPE (%) values per satellite sensor grouped by Chl-*a* models. (Red reference line to right Y axis is set to 50 %, upper threshold value for reasonable forecasting; Lewis, 1982).

2.3.3.2 Secchi depth

In general, application of $Secchi_{general}$ model on Landsat only images indicated better performance compared to Sentinel 2 data (Table 2.3.3-4). Moreover, the performance of general model is improved when mixed satellite data is used. In particular, the high decrease of ME value (0.25 vs 0.7 m) and the similar values of RMSE (Figure 2.3.3-3) to those of only-Landsat employed data prove that the combination of mixed satellite data yields reliable Secchi depth values.

$Secchi_{natural}$ model (Table 2.3.3-5) employing Landsat-only images was proven once again better based on RMSE value. On the other hand, the ME value is significantly lower when is accompanied by Sentinel 2 data, (0.1 vs 0.6 m). However, it should be noted that those differences are not significant. Results yielded from multi sensor images are similar to those of Landsat data except for the median MAPE value which seems to increase (38.7 vs. 30.2 %) with the combined satellite sensors.

Application of $Secchi_{artificial}$ model (Table 2.3.3-6) indicated a light superiority when is accompanied by Sentinel 2 data but no great differences are noticed. Similar but lower values of RMSE, NRMSE and median MAPE indicated that the Landsat-developed Secchi artificial model can perform satisfactorily and quantify reliable Secchi depth values in reservoirs based on Sentinel-2 reflectance data. $Secchi_{artificial}$ model performs equally well even when employing multi sensor data. Similar values of most of metrics but significantly lower value of ME (0.01 m) implied the suitability of this model even with combined satellite data.

Concerning the values of coefficient correlation, all correlations among satellite and *in-situ* data were statistically significant ranging from 0.54 (mixed satellites; general model) to 0.73 (Landsat; natural model).

Table 2.3.3-4. Values of error metrics regarding the sensor-based applications of $Secchi_{general}$ model (Underlined value indicates the lowest value among all three cases; ** Correlation is significant at the 0.01 level (2-tailed).

$Secchi_{general}$ model	**N**	**SPEARMAN r**	**Average In Situ SECCHI (m)**	**Average Satellite SECCHI (m)**	ME **(m)**	**RMSE (SECCHI; m)**	**NRMSE**	**MEDIAN MAPE (%)**
Landsat 7 ETM+/8 OLI	115	<u>0.57</u>**	4.5	3.8	0.702	2.512	0.165	<u>34.1</u>
SENTINEL 2	111	<u>0.57</u>**	4.64	5.24	-0.699	4.368	0.287	43.9
MIXED SATELLITES	111	0.54**	4.5	4.3	<u>0.252</u>	<u>2.375</u>	<u>0.156</u>	35.2

Table 2.3.3-5. Values of error metrics regarding the sensor-based applications of $Secchi_{natural}$ model (Underlined value indicates the lowest value among all three cases; ** Correlation is significant at the 0.01 level (2-tailed).

$Secchi_{natural}$ model	**N**	**SPEARMAN r**	**Average In Situ SECCHI (m)**	**Average Satellite SECCHI (m)**	ME **(m)**	**RMSE (SECCHI; m)**	**NRMSE**	**MEDIAN MAPE (%)**
Landsat 7 ETM+/8 OLI	28	0.73**	4.24	3.6	0.63	2.765	0.182	30.2
SENTINEL 2	28	0.56**	4.24	4.1	0.103	2.929	0.193	41.2
MIXED SATELLITES	28	0.66**	4.24	3.7	0.495	2.844	0.187	38.7

Table 2.3.3-6. Values of error metrics regarding the sensor-based applications of $Secchi_{artificial}$ model (Underlined value indicates the lowest value among all three cases; ** Correlation is significant at the 0.01 level (2-tailed).

$Secchi_{artificial}$ model	**N**	**SPEARMAN r**	**Average In Situ SECCHI (m)**	**Average Satellite SECCHI (m)**	ME **(m)**	**RMSE (SECCHI; m)**	**NRMSE**	**MEDIAN MAPE (%)**
Landsat 7 ETM+/8 OLI	40	0.56**	4.85	4.8	-0.071	1.612	0.183	24.7
SENTINEL 2	36	0.63**	4.47	4.32	0.149	1.458	0.159	24.22
MIXED SATELLITES	43	0.58**	4.6	4.6	0.008	1.528	0.166	25.1

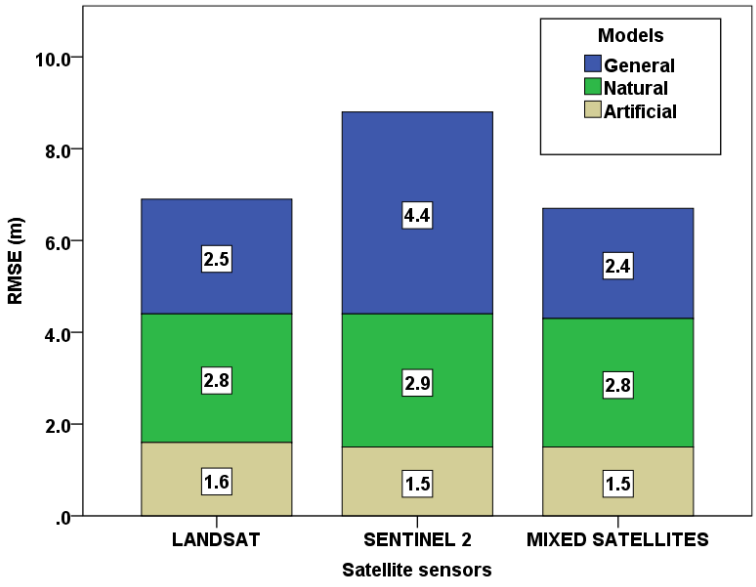


Figure 2.3.3-3. Stacked bars illustrating RMSE values of Secchi disk depths (m) per satellite sensor.

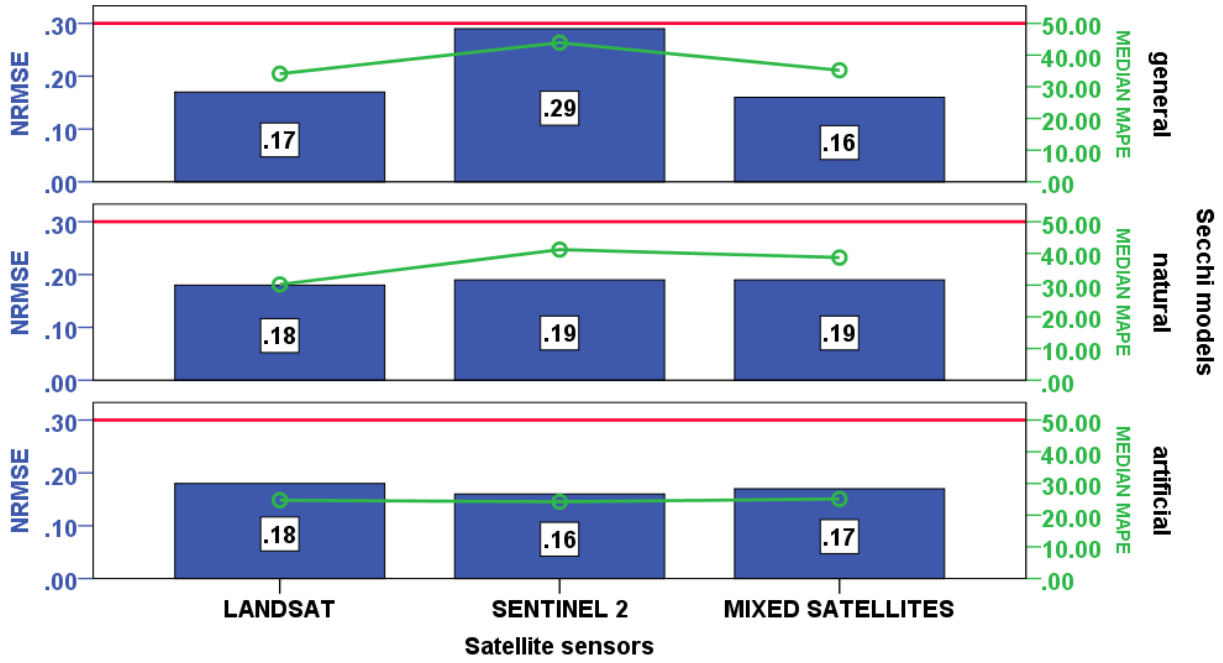


Figure 2.3.3-4. NRMSE and median MAPE (%) values per satellite sensor grouped by Secchi depth models. (Red reference line to right Y axis is set to 50 %, upper threshold value for reasonable forecasting; Lewis, 1982).

2.3.3.3 Total phosphorus

Application of $TP_{general}$ model in Landsat-only images demonstrated a clear superiority compared to the Sentinel 2-based employment (Table 2.3.3-7). All of the studied error metrics' values accrued from the Landsat used data are lower in comparison with those of the corresponding Sentinel 2. Furthermore, the combined adoption of multi sensor images has not revealed any further improvement while the majority of resulted statistical indices are similar to those resulted by the use of Sentinel 2 data.

Application of special TP model on only natural lakes (Table 2.3.3-8) yielded different results. Value of RMSE is lower when mixed satellite data are employed whereas ME value is significantly lower after the employment of Sentinel 2 images (Figure 2.3.3-5; green line). The employment of mixed satellite images indicated the natural model's suitability even when used in combination with Landsat and Sentinel 2 satellite data, based on RMSE metric.

Correlations among satellite derived and *in-situ* TP concentrations revealed in general strong relationships while Spearman values ranged from 0.67 (Landsat; natural model) to 0.75 (Sentinel 2; general model).

Table 2.3.3-7. Values of error metrics regarding the sensor-based applications of $TP_{general}$ model (Underlined value indicates the lowest value among all three cases; ** Correlation is significant at the 0.01 level (2-tailed).

$TP_{general}$ model	**N**	**SPEARMAN r**	**Average In Situ TP (mg/l)**	**Average Satellite TP (mg/l)**	ME **(mg/l)**	**RMSE (TP; mg/l)**	**NRMSE**	**MEDIAN MAPE (%)**
Landsat 7 ETM+/8 OLI	33	0.71**	0.075	0.077	-0.0014	0.028	0.139	21.7
SENTINEL 2	30	0.75**	0.08	0.07	0.0047	0.032	0.156	23.6
MIXED SATELLITES	30	0.73**	0.08	0.076	0.0016	0.0295	0.145	24.7

Table 2.3.3-8. Values of error metrics regarding the sensor-based applications of $TP_{natural}$ model (** Correlation is significant at the 0.01 level (2-tailed).

$TP_{natural}$ model	**N**	**SPEARMAN r**	**Average In Situ TP (mg/l)**	**Average Satellite TP (mg/l)**	ME **(mg/l)**	**RMSE (TP; mg/l)**	**NRMSE**	**MEDIAN MAPE (%)**
Landsat 7 ETM+/8 OLI	55	0.67**	0.07	0.076	-0.0051	0.0312	0.148	38.2
SENTINEL 2	49	0.7**	0.07	0.07	0.0005	0.0308	0.146	26.3
MIXED SATELLITES	49	0.69**	0.07	0.071	-0.0007	0.0300	0.142	26.3

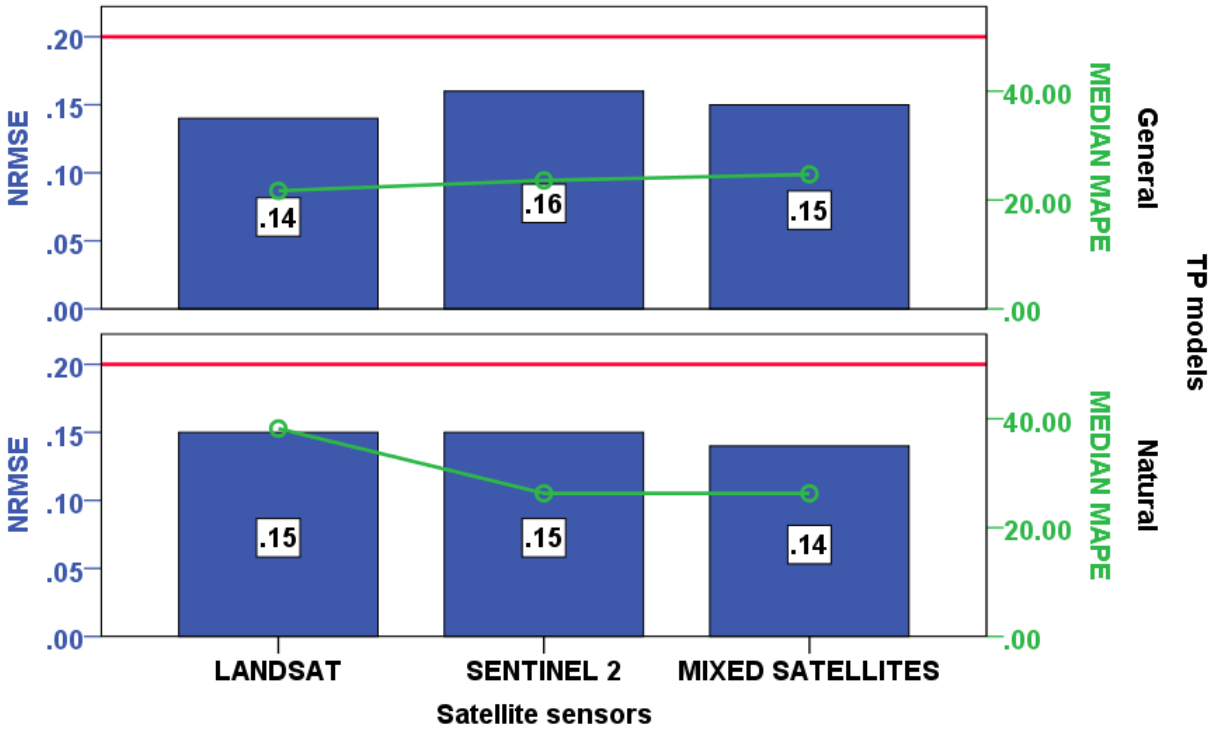


Figure 2.3.3-5. NRMSE and median MAPE (%) values per satellite sensor grouped by TP models. (Red reference line to right Y axis is set to 50 %, upper threshold value for reasonable forecasting; Lewis, 1982).

Concerning the overall efficiency of WQ models based on MAPE values, it is observed that Chl-*a* models are characterized by less quantification accuracy compared to respective Secchi and TP (Figure 2.3.3-6). Most of Chl-*a* models touch and even surpass the threshold MAPE value of 50% indicating marginally reasonable forecasting. Secchi and TP models achieved better results providing more superior forecasting.

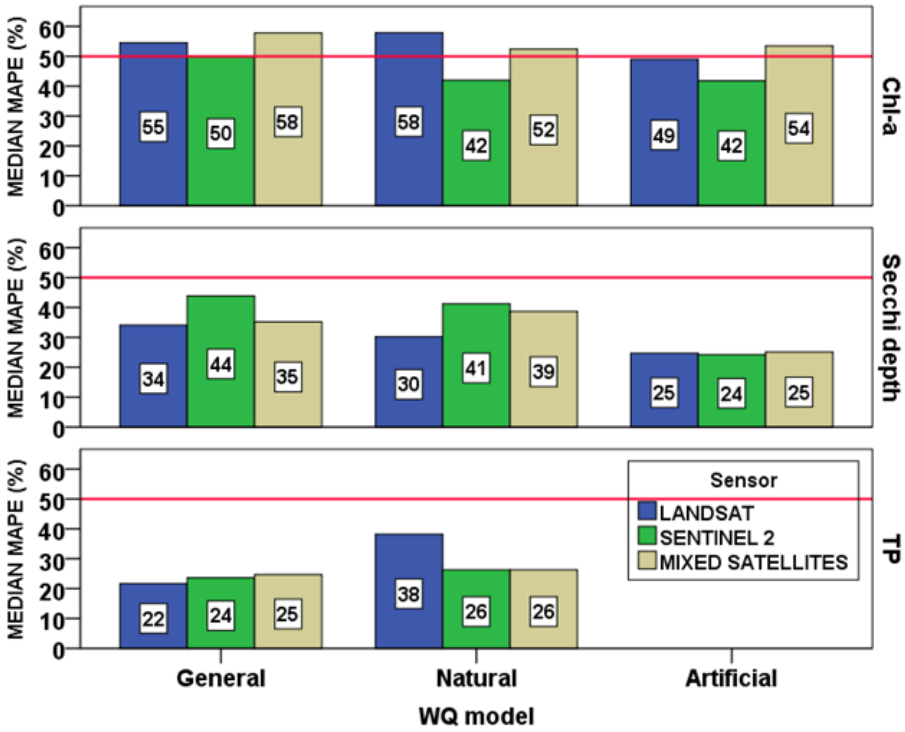


Figure 2.3.3-6. Median MAPE (%) values per WQ quantitative model grouped by WQ parameters. (Red reference line to Y axis is set to 50 %, upper threshold value for reasonable forecasting; Lewis, 1982).

2.3.4 Discussion

Main objective of this chapter section was to explore initially whether Landsat-based empirical WQ algorithms can be efficiently applied to Sentinel 2 images and then whether the combined use of multi-sensor data improves the algorithms' prediction accuracy. Furthermore, independently from whether there is some improvement or not, the ultimate goal of this combined approach was to decide whether multi-sensor images could be used with at least equally reliable results as those accrued from only-one sensor's utilization.

As far as the general models of all WQ elements (Chl-*a*, Secchi depth and TP) is concerned, all models were more efficient and accurate when were accompanied by Landsat images while no improvement was observed by using multi sensor images with the exception of $Secchi_{general}$ model. Natural models, though, demonstrated a different behavior. More particularly, Chl-*a* and TP natural models presented lower values of error metrics when employing Sentinel 2 images and only Secchi natural model performed better with Landsat data. Combined utilization of Landsat and Sentinel 2 images did not provide any improvement to corresponding Chl-*a* and Secchi models whereas the multi sensor images resulted in TP concentrations with equally

reliable outcomes as those employing Sentinel 2. Regarding the artificial algorithms, performance of Chl-*a* model was similar either by exploiting Landsat or Sentinel 2 data while Secchi model achieved slightly better efficiency with Sentinel 2 images.

The shortcoming of Sentinel 2 images to reach Landsat' performance in most of cases can probably be attributed to two reasons: first of all, hereby utilized WQ empirical models were developed based on Landsat-7 ETM+ and 8 OLI images; hence it is expected to be affected by the corresponding spectral configuration and perform better when employing Landsat rather than Sentinel 2 reflectance.

Furthermore, according to Mandanici and Bitelli (2016) who compared reflectance and index values of Landsat 8 OLI and Sentinel 2 imagery for a combined use, confirmed that MSI band 8A (vegetation red edge) is the optimal option from the radiometric point of view when Sentinel-2 images are associated with Landsat 8 ones. Instead, MSI band 8 (NIR) is highly recommended for a joint use with older Landsat series, such as Landsat 5. In the framework of this study though, the match of the Sentinel 2 B08 (NIR) band to bands B4 (L7 ETM+) and B5 (L8 OLI), may constituted an obstacle in achieving better and more accurate WQ quantifications when employing Sentinel 2 data. Further sources of different results between Landsat and Sentinel 2 images are the residual effects of water specular reflections, derived from the different azimuth and elevation of the sensors.

Similar works (Werff and Meer, 2016) having studied the potential combined usage of Landsat and Sentinel 2 images, indicated that by visual inspection satellite products of Landsat and Sentinel 2 sensors are similar; when however, reflectance values are compared there are differences which should in each case be evaluated (Mandanici and Bitelli, 2016).

Furthermore, the largely worse performance of Chl-*a* models compared to rest of WQ elements emphasizes once again the complexity that mapping of Chl-*a* in Case 2 waters (coastal and/or inland waters) hides. As optical properties are measured based on a compound of dissolved organic matter, dead organic and inorganic particulate matter, and phytoplankton (Chl-*a*), Chl-*a* determination is characterized by less accuracy since these constituents are not statistically correlated (Markogianni et al., 2020).

All in all, it is proven that hereby WQ models are proposed to employ principally Landsat images; however, the employment of Sentinel 2 data potentially produces reliable results with some (not significant) deviations, from reference *in-situ* data, regarding the assessment of lake WQ.

2.3.5 References

Deutsch, E.S.; Alameddine, I.; El-Fadel, M. Monitoring water quality in a hypereutrophic reservoir using Landsat ETM+ and OLI sensors: how transferable are the water quality algorithms? *Environ Monit Assess* **2018**, 190, 141. https://doi.org/10.1007/s10661-018-6506-9

Lewis, C. D. Industrial and business forecasting methods, London: Butterworths, 1982.

Li Peng; Chen Shenliang; Ji Hongyu; Ke Yinghai; Fu Yutao. Combining Landsat-8 and Sentinel-2 to investigate seasonal changes of suspended particulate matter off the abandoned distributary mouths of Yellow River Delta, Marine Geology, **2021**, 441, 106622. https://doi.org/10.1016/j.margeo.2021.106622.

Mandanici, E.; Bitelli, G. Multi-image and multi-sensor change detection for long-term monitoring of arid environments with Landsat series. Remote Sens. **2015**, 7, 14019–14038.

Mandanici, E.; Bitelli, G. Preliminary Comparison of Sentinel-2 and Landsat 8 Imagery for a Combined Use. *Remote Sens.* **2016**, *8*, 1014. https://doi.org/10.3390/rs8121014.

Markogianni, V.; Kalivas, D.; Petropoulos, G.P.; Dimitriou, E. Estimating Chlorophyll-*a* of Inland Water Bodies in Greece Based on Landsat Data. *Remote Sens.* **2020**, *12*, 2087. https://doi.org/10.3390/rs12132087.

Markogianni, V.; Kalivas, D.; Petropoulos, G.P.; Dimitriou, E. Modelling of Greek Lakes Water Quality Using Earth Observation in the Framework of the Water Framework Directive (WFD). Remote Sens. **2022**, 14, 739. https://doi.org/10.3390/rs14030739.

McKinna, L.; Werdell, P. Approach for identifying optically shallow pixels when processing ocean-color imagery. Opt. Express **2018**, 26, A915–A928.

Moreno Montaño, J.J.; Palmer Pol, A.; Sesé Abad, A.; Cajal Blasco, B. Using the R-MAPE index as a resistant measure of forecastaccuracy. Psicothema **2013**, 25, 500–506. https://doi.org/10.7334/psicothema2013.23.

Olmanson, L. G.; Brezonik, P. L.; Finlay, J. C.; Bauer, M. E. Comparison of Landsat 8 and Landsat 7 for regional measurements of CDOM and water clarity in lakes. Remote Sensing of Environment **2016**, 185, 119–128. https://doi.org/10.1016/j.rse.2016.01.007.

Pedreros-Guarda, M.; Abarca-del-Río, R.; Escalona, K.; García, I.; Parra, Ó. A Google Earth Engine Application to Retrieve Long-Term Surface Temperature for Small Lakes. Case: San Pedro Lagoons, Chile. Remote Sens. **2021**, 13, 4544. https://doi.org/10.3390/ rs13224544.

Van der Werff, H.; van der Meer, F. Sentinel-2A MSI and Landsat 8 OLI provide data continuity for geological remote sensing. Remote Sens. **2016**, 8, doi:10.3390/rs8110883.

Wang, L.; Xu, M.; Liu, Y.; Liu, H.; Beck, R.; Reif, M.; Emery, E.; Young, J.; Wu, Q. Mapping Freshwater Chlorophyll-*a* Concentrations at a Regional Scale Integrating Multi-Sensor Satellite Observations with Google Earth Engine. *Remote Sens.* **2020**, *12*, 3278. https://doi.org/10.3390/rs12203278.

3. Atmospheric correction analysis of lake WQ models by employing surface reflectance embedded in GEE platform

3.1 Introduction

Nowadays, the World's lakes' greatest threat is common; the so-called eutrophication, which is greatly connected to the increase of nutrients, mainly phosphorus and nitrogen (Pedreros-Guarda et al., 2021).

Water quality is the most significant indicator of a water body's state, while its assessment requires the continuous monitoring of mainly physico-chemical and biological elements (Fatima, 2018). Traditionally, WQ estimation is conducted based on *in-situ* sampling and laboratory analysis (Li et al., 2016). However, those methods are time- and labor-intensive particularly when large-scale investigations are the case (Zhang et al., 2014). Today, through the evolution of Remote Sensing (RS) techniques, satellite images offer valuable information facilitating the assessment of different WQ components, such as the total suspended matter (TSM) and colored dissolved organic matter (CDOM) content, the Secchi depth (SD), and the chlorophyll-*a* concentration (Brezonik et al., 2015; Sagan et al., 2020; Markogianni et al., 2020; Zhang et al., 2021; Song et al., 2022; Markogianni et al., 2022). Furthermore, due to their wide coverage, RS expedites the regional and large-scale WQ monitoring (Gholizadeh et al., 2016; Topp et al., 2020; Pizani and Maillard, 2022).

However, computing WQ properties from RS images may also become time-demanding and sophisticated because of the processing data chain particularly a great-scale WQ assessment and high-frequency time series demand (Kumar and Mutanga, 2018). Today, platforms for big EO Data Management and Analysis have emerged as computational solutions that provide functionalities for big EO data management, storage and access including processing without downloading big amounts of EO data sets and provision of images of certain pre- processing levels (Gomes et al., 2020). Gomes et al. (2020) overviewed and compared seven platforms among certain functionalities: Google Earth Engine (GEE), Sentinel Hub (SH), Open Data Cube (ODC), System for Earth Observation Data Access, Processing and Analysis for Land Monitoring (SEPAL), open EO, JEODPP and pipsCloud while the reviewed functionalities are the following: data abstraction, processing abstraction, physical infrastructure abstraction, open governance, reproducibility of science, infrastructure replicability, processing scalability, storage scalability, data access interoperability and extensibility. Based initial

on this survey and then on researches conducted by Maciel et al. (2021), Zhao et al. (2022) and Pizani and Maillard (2022), it was concluded that GEE outperforms among the available cloud computing systems as it proven the most significant cloud processing platform for the remote sensing community due to its ease of use and maturity.

Google Earth Engine (GEE) platform has changed the traditional RS data processing mode (Li et al., 2022) as it consists of a multi-petabyte analysis-ready data catalog while allowing users to compute massive-scale analysis and accomplish multiple RS and geospatial tasks at remarkable speeds and scales (Gorelick et al., 2017). The data repository of GEE includes publicly available geospatial datasets, along with observations from a variety of satellite and aerial imaging systems in both optical and non-optical wavelengths, environmental variables, weather forecasts, land cover and other datasets (Gorelick et al., 2017). In addition to this, several involved operators such as the United State Geological Survey (USGS), National Aeronautics and Space Administration (NASA), and European Space Agency (ESA) -among others- are collaborating with Google Inc. and have made satellite data available online through the Google Earth Engine (GEE) cloud platform (Wang et al., 2020). The satellite data catalog is updated on a daily basis with around 6000 new image scenes.

Recent studies have recorded applications of WQ monitoring based on Google Earth Engine (Jia et al., 2019; Zong et al., 2019; Maeda et al., 2019; Wang et al., 2020; Weber et al., 2020; Lobo et al., 2021; Somasundaram et al., 2021; Bioresita et al., 2021; Vaičiūtė et al., 2021; Kislik et al., 2022; Wen et al., 2022).

Based on all of the above, this effort will employ GEE-retrieved reflectance values to assess WQ elements in 50 Greek natural and artificial lakes, constituting the National Lake Monitoring Network in Greece for the WFD.

The hereby adopted methodological scheme includes the exploration of the performance of published empirically-developed WQ quantitative models of Chl-*a* (Markogianni et al., 2020), Secchi depth and Total phosphorus (Markogianni et al., 2022) when employing GEE-retrieved reflectance values subjected to different atmospheric correction (AC) methods. Precise AC is important for applications where small differences in surface reflectance (SR) are significant, such as retrieval of WQ parameters (Nazeer et al., 2014; Warren et al., 2019; Pahlevan et al., 2021). Furthermore, it enables direct comparison between different image dates and different sensors. AC methods fall into two types, namely physical (e.g. FLAASH, ATCOR, 6S) and image-based methods (e.g. DOS). Physical methods use a radiative transfer model to estimate SR while image-based methods obtain relevant parameters from the image (Nazeer et al., 2014).

In this study, *in-situ* measurements of Chl-*a*, Secchi depth and Total phosphorus of 2018 have been paired twice with concurrent satellite reflectance values derived from combined Landsat 7 +ETM/ 8 OLI (first dataset) and Sentinel 2 MSI (second dataset) images. The two-fold match concern initially the reflectance derived from manually downloaded and pre-processed images with the DOS1 method and then the GEE- derived reflectance (from the exact same satellite images) subjected to LaSRC (Landsat 8 OLI), LEDAPS (Landsat 7 ETM+) and Sen2Cor (Sentinel 2) correction algorithms. The aforementioned reflectance values were extracted from the points were WFD sampling sites are located. Linear regression analysis among the resulted WQ values was then conducted to highlight and potentially harmonize inherent differences primarily between the differently pre-processed reflectance values and afterwards among the different sensors used.

As it is already mentioned, used sensors in this research are the Landsat 7ETM+/8 OLI and Sentinel 2. Landsat (30 m spatial resolution) and Sentinel-2 (10–60 m spatial resolution) missions provide fine-scale spatial data and have been reported to be suitable for the quantification of multiple WQ indices in freshwater lakes and reservoirs (Bresciani et al., 2018; Markogianni et al., 2018; Pahlevan et al., 2020; Bramich et al., 2021; Zhou et al., 2021; Zhang et al., 2021; Song et al., 2022). Landsat sensors' temporal resolution is 16 days while Landsat 7 ETM+ and 8 OLI together, provide one (1) satellite image for every 8 days (Pedreros-Guarda et al., 2021). The Sentinel-2 mission carries two satellites—Sentinel-2A and Sentinel-2B—equipped with identical Multispectral Instruments (MSI)-, the revisit frequency of each satellite is 10 days while the combined revisit equals to 5 days.

However, those sensors differ in their orbital spatial, and spectral configuration, resulting in affecting the recorded radiometric values; hence sensor-based datasets have been processed independently to be less influenced (Mandanici and Bitelli, 2015). Moreover, available *in-situ* WQ data of years 2019 and 2020 have been paired with concurrent GEE-derived reflectance values to further validate and certify the strength and the suitability of the WQ universal models for estimating Chl-*a*, Secchi depth and Total phosphorus concentrations of optically-diverse inland waters at a national scale (Greece).

In purview of the above, present study's specific objectives are to: (1) test the spatiotemporal performance of empirically-developed WQ models when employing GEE-retrieved reflectance that is pre-processed with different correction methods; (2) highlight the differences and harmonize them by developing corrected WQ models; (3) develop sensor-specific corrected WQ monitoring algorithms individually for Landsat and Sentinel 2 sensors and (4) map WQ elements across Greek studied lakes, through GEE cloud-based platform.

3.2 Methodology

3.2.1 In-situ data

In-situ dataset used in this study concern the freely accessible data collected in the context of the Greek Water Monitoring Network for lakes (WFD) by the staff of the Goulandris Natural History Museum, Greek Biotope/Wetland Centre. Water samples have been collected along 50 lakes, natural and artificial, from 53 sites (27 surveillance; 26 operational) (Figure 3.2.1-1). In particular, herein used data include measured values of Chl-*a* and Total phosphorus concentrations and Secchi depth measurements on several dates during the years 2018, 2019 and 2020.

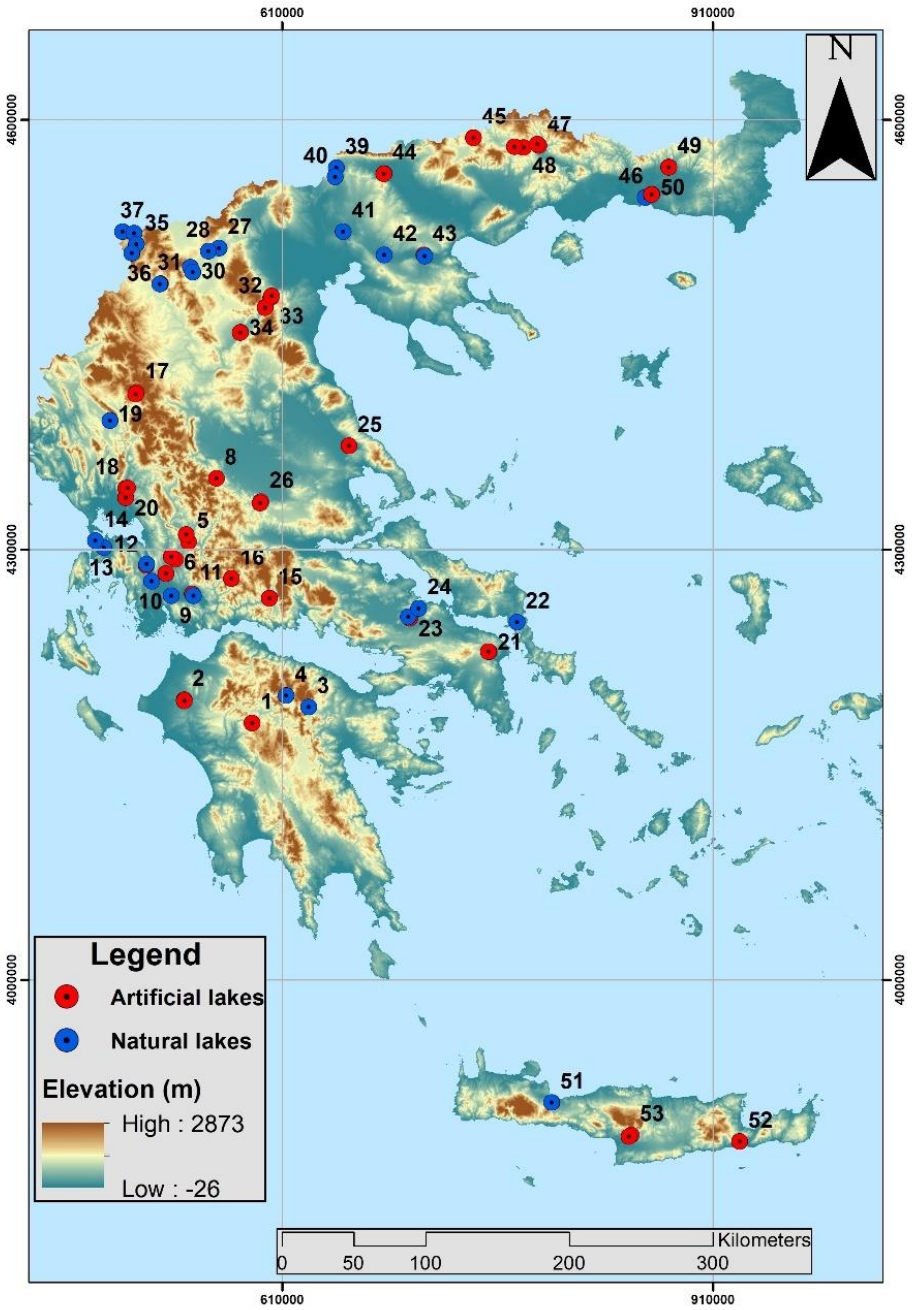


Figure 3.2.1-1. National Lake Monitoring Network in Greece (numbers of sampling stations coincide with the numbers presented in Table 2.2.2-1).

3.2.2 Satellite imagery selection and pre-processing

Acquired SR products from the GEE repository included Sentinel-2 MSI (Level-2A SR data) and Landsat 7 +ETM /8 OLI multispectral images (Collection 1 Level 1-precision and terrain correction- Tier 1; SR data). In this study, the time window for satellite acquisition from GEE, was set to ± 7 days in relation to the sampling date as in Wen et al. (2022). GEE-derived reflectance values have undergone different AC algorithms; more particularly

Landsat 8 OLI and Landsat 7 +ETM images are corrected using the Land Surface Reflectance Code (LaSRC) and the Ecosystem Disturbance Adaptive Processing System (LEDAPS) methods, respectively. LaSRC AC is performed using a radiative transfer model, auxiliary atmospheric data from MODIS and utilizes the coastal aerosol band for aerosol inversion tests. LEDAPS algorithm calculates the radiative transfer for atmospheric data from MODIS and NCEP (Ermida et al., 2020). Sentinel-2 products were processed with the Sentinel 2 Correction (Sen2Cor) algorithm. Sen2Cor algorithm is a combination of state-of-the-art techniques for performing AC together with a scene classification algorithm, which allows detection of clouds, snow and cloud shadows and generation of a classification map. This map consists of 3 different classes for clouds (including cirrus), 6 different classifications for shadows, cloud shadows, vegetation, not vegetated, water and snow (ESA, Sentinel Online, accessed on 10/07/2022). Then, SR values of several dates during the years 2018, 2019 and 2020 were extracted through the GEE platform at the points where the sampling stations are located, initially from Landsat (8 OLI, 7+ETM) and then from Sentinel 2 MSI images (Figure 3.2.2-1). Since the reflectance fraction in GEE is scaled by 10000, values were divided by 10000 to obtain 0-1 reflectance values from the respective cells. SR extraction from images in GEE platform was accomplished by the staff of GIS Research Unit of Agricultural University of Athens.

Concerning the manual pre-processing of satellite images, Landsat 7 ETM+ and 8 OLI images of 2018 were downloaded from the USGS (United States Geological Survey) Data Centre (https://earthexplorer.usgs.gov/) in the context of the study conducted by Markogianni et al. (2022). Moreover, Sentinel 2 images of 2018, were as well downloaded from the Copernicus open access hub (https://scihub.copernicus.eu/dhus/#/home), in the context of the previous chapter, studying the transferability/performance of Landsat-based WQ models to Sentinel 2 ones. Manually downloaded Sentinel 2 and Landsat images have been subjected to the same pre-processing procedure, as described in Markogianni et al. (2022) and more particularly, they were imported in the semi-automatic classification plugin (SCP) of the free and open-source cross-platform desktop Quantum Geographic Information System (Q-GIS), v. 3.6.3-Noosa to perform: (a) conversion of images from digital numbers (DN) to top-of-atmosphere reflectance (TOA), (b) AC by using the DOS1 method (Chavez J., 1988; applied to all bands except for thermal ones), and (c) the creation of a band stack set for each image. DOS1 method minimizes the additive effect of the atmosphere caused by haze. The main assumption is that dark objects represent 1% of reflectance while they are identified by an area with clear water in deep lakes or by the histogram method, which selects the DN of haze from the DN frequency histogram of an image (El Alem et al., 2021). The selection of DOS method was based on the

studies of Nazeer et al. (2014) and Doña et al. (2014), who evaluated 5 (6S, FLAASH, ATCOR, DOS and ELM) and 3 (DOS; ATCOR3; MODTRAN5) different AC methods, respectively. Nazeer et al. (2014) overviewed those methods over sand, artificial surface, grass and water while they concluded that DOS performed well over water, it showed higher differences than the physical methods and is proposed as a good choice for SR estimation of dark surfaces such as water. Doña et al. (2014), evaluated the aforementioned methods across certain Spanish lakes and ponds and concluded that DOS performed better than the others, reporting the lowest errors.

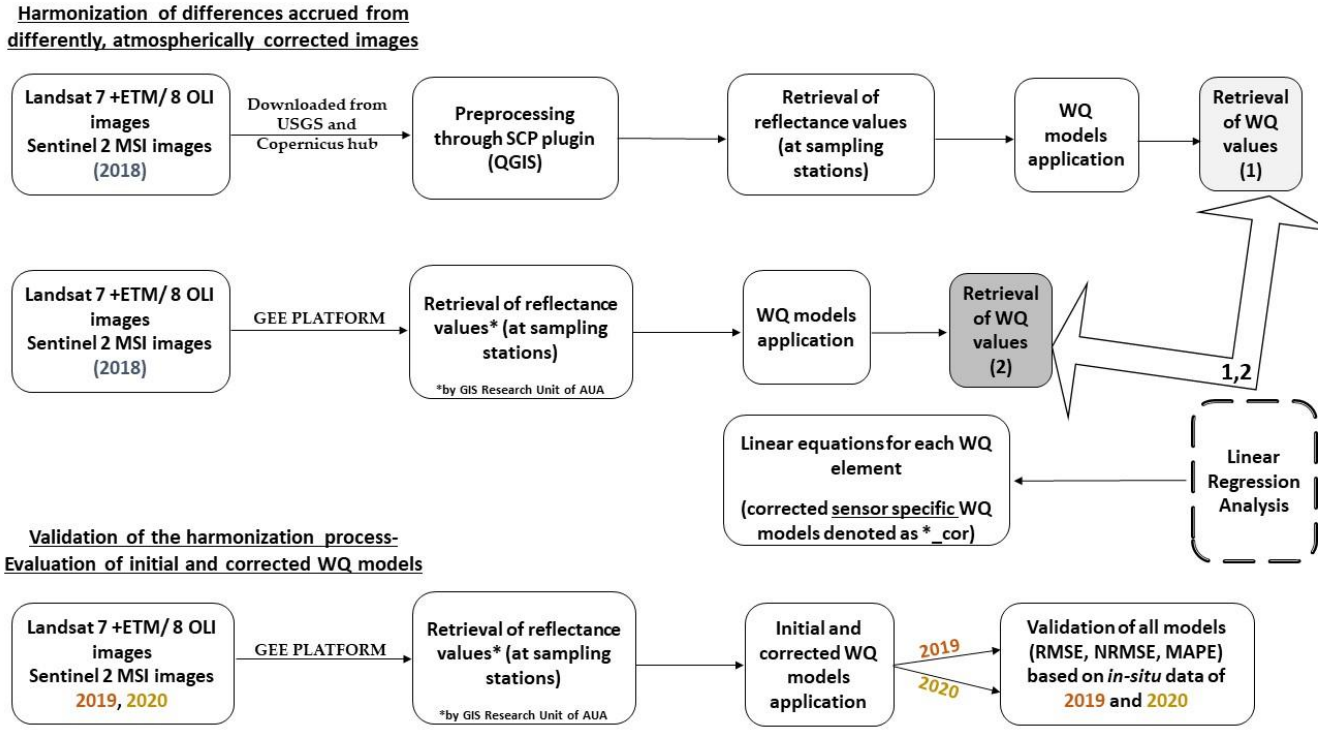


Figure 3.2.2-1. Flowchart of research methodology steps (harmonization and validation processes). Different shades of grey color [also distinguished with numbers (1) and (2)] represent the distinctive image datasets that are differently pre-processed (DOS1-manually and LaSRC; LEDAPS; Sen2Cor-GEE) and employed in WQ models. Different years of *in-situ* datasets are highlighted with distinct colors.

3.2.3 Harmonization among SR products subjected to different atmospheric correction methods

In the context of this study, the WQ models developed by Markogianni et al. (2020; Chl-*a*); and Markogianni et al. (2022; Secchi depth; TP) were initially applied to manually-downloaded and DOS1-pre-processed Landsat 7+ETM/ 8 OLI and Sentinel 2 images of 2018 (Figure 3.2.2-1). Then, WQ models employed GEE-derived SR values originating from the same images,

but undergone different AC methods (Figure 3.2.2-1). Application of WQ models includes Chl-*a* models (General- Equation 3.2.3-1; Natural- model-Equation 3.2.3-2; Artificial model -Equation 3.2.3-3), Secchi depth models (General-Equation 3.2.3-4; Natural model- Equation 3.2.3-5; Artificial model - Equation 3.2.3- 6) and TP models (General-Equation 3.2.3-7; Natural model - Equation 3.2.3-8). The basic goal of this elaboration is the exploration and establishment of relationships between same-located lake WQ values which originate from different AC methods and sensors. Linear regression analyses incorporate WQ values accrued by the employment of a) manually DOS1 corrected reflectance (dependent variable) and b) LaSRC, LEDAPS and Sen2Cor-corrected reflectance in GEE (independent variable) at corresponding locations. The development of linear equations should facilitate the normalization of the WQ models' results to acceptable and comparable values when employing different-from-DOS1 corrected reflectance values for each WQ model, its specifications (general model, natural-only, artificial-only; Figure 3.2.2-1) ans each sensor. This normalization was based on the dataset of 2018 (Figure 3.2.2-1) while the analysis yielded 192 and 210 match-up points of *in-situ* data and Landsat 7 +ETM/8 OLI and Sentinel 2 images embedded in GEE platform, respectively. Furthermore, the possible detection of strong relationships would indicate the suitability of WQ models to employ images not only subjected to aforementioned AC methods except for DOS but also of Sentinel 2 sensor. Regression analysis concern all cases presented in Table 3.2.3-1.

$$\log Chla = 3.599 - 0.63 * \left(\frac{blue}{red}\right) - 2.183 * \left(\frac{\ln red}{\ln swir2}\right) \quad (3.2.3\text{-}1)$$

$$\log Chla = 4.443 - 1.421 * \left(\frac{blue}{green}\right) - 3.454 * \left(\frac{\ln red}{\ln swir2}\right) + 1.304 * \left(\frac{red}{green}\right) \quad (3.2.3\text{-}2)$$

$$\log Chla = 2.919 - 2.011 * \left(\frac{\ln red}{\ln swir1}\right) + 1.449 * \left(\frac{red}{green}\right) - 1.441 * \left(\frac{\ln red}{\ln blue}\right) \quad (3.2.3\text{-}3)$$

$$SQRT(Secchi)_{general} = 1.215 - 2.479 * \left(blue + red + \frac{red}{blue}\right) + 3.394 * \left(\frac{lngreen}{lnswir2}\right) \quad (3.2.3\text{-}4)$$

$$SQRT(Secchi)_{natural} = 1.172 - (1.003 * logchl - a) - (1.031 * logred) \quad (3.2.3\text{-}5)$$

$$SQRT(Secchi)_{artificial} = 3.927 - 1.365 * \left(\frac{green}{blue}\right) - 0.318 * \left(\frac{red}{swir1}\right) - 0.361 * logchla \quad (3.2.3\text{-}6)$$

$$LogTPgeneral = -1.425 + 0.452 * logChla - 0.573 * (\frac{lnred}{lnswir1})$$ (3.2.3-7)

$$LogTPnatural = -0.633 - (0.704 * logSecchi) - 0.392 * (\frac{green}{red})$$ (3.2.3-8)

Table 3.2.3-1. WQ models involved in regression analysis (2018 dataset) for harmonization of different reflectance values emerged from different pre-processing methods and satellite sensors.

Equation	Models and preconditions	Sensor
3.2.3-1	Chl-*a*_general; mean depth> 5m	Landsat 7+ETM/8 OLI; Sentinel2
3.2.3-2	Chl-*a*_natural; mean depth> 5m	Landsat 7+ETM/8 OLI; Sentinel2
3.2.3-3	Chl-a_artificial; mean depth> 5m; date difference (sampling/satellite): +-5 days	Landsat 7+ETM/8 OLI; Sentinel2
3.2.3-4	Secchi_general; mean depth> 5m	Landsat 7+ETM/8 OLI; Sentinel2
3.2.3-5	Secchi_natural; mean depth> 5m	Landsat 7+ETM/8 OLI; Sentinel2
3.2.3-6	Secchi_artificial; mean depth> 5m; date difference (sampling/satellite): +-5 days	Landsat 7+ETM/8 OLI; Sentinel2
3.2.3-7	TP_general; mean depth> 3.5 m	Landsat 7+ETM/8 OLI; Sentinel2
3.2.3-8	TP_natural; mean depth> 3.5 m	Landsat 7+ETM/8 OLI; Sentinel2
3.2.3-7	TP_artificial; Application of TP general model on artificial lakes: mean depth> 3.5; date difference (sampling/satellite): +-5 days (where was possible)	Landsat 7+ETM/8 OLI; Sentinel2

Since dataset of 2018 was used for the development of the linear equations e.g. the corrected WQ models, *in-situ* datasets of 2019 and 2020 were used for their validation (Figure 3.2.2-1). In particular, as far as the year 2019 is concerned, 239 pairs of *in-situ* measurements and Landsat 7 +ETM/ 8 OLI values of reflectance (GEE) were created, accompanied by 242 pairs with Sentinel 2 images. Concerning the year of 2020, the paired reflectance of Landsat 7+ETM/ 8 OLI and Sentinel 2 (GEE) with *in-situ* data were 220 and 286, respectively. Further validation of initial and corrected WQ models' performance (Equations 3.2.3-1 to 3.2.3-8) for years 2019 and 2020 was based on the error metrics MAPE, RMSE and NRMSE. Additionally, MAPE values are interpreted according to Lewis (1982) concerning their forecasting potential (Table 2.3.2-2).

3.3 Results

3.3.1 Harmonization among SR values subjected to different AC processors

After the application of WQ models, initially by employing DOS1-corrected reflectance and then GEE-derived reflectance values connected to 2018 *in-situ* values, linear regression analyses were conducted among the resulted values, concerning all the three WQ elements (Chl-*a*, Secchi depth, TP) and utilized satellite sensors (integrated Landsat 7 +ETM/8 OLI images; Sentinel 2). The aforementioned statistical analyses yielded linear equations (Table 3.3.1-1) accompanied by high coefficient of determination (R^2) values except for Chl-$a_{artificial}$ model based on Landsat reflectance and Chl-$a_{general}$, Chl-$a_{artificial}$ and $Secchi_{general}$ models based on Sentinel 2 reflectance.

Application of Chl-$a_{general}$ model on Landsat images indicated a superiority concerning the ME value when the LaSRC and LEDAPS correction methods are used (compared to DOS1) even though both values of RMSE are similar. On the other hand, the Chl-$a_{natural}$ model employing LaSRC and LEDAPS corrected data performs worse than the respective DOS1 one, based on values of both ME and RMSE. Chl-$a_{artificial}$ model performs equally well with all AC methods.

Regarding the Secchi depths, the general model (DOS1; Table 3.3.1-1) presents a lower ME value and a similar RMSE to the respective employing the corrected reflectance retrieved from the GEE platform. The $Secchi_{natural}$ model, using LaSRC and LEDAPS correction methods, introduces a lower ME value but similar RMSE with the respective using DOS-1 corrected data. Moreover, $Secchi_{artificial}$ model applied on Landsat images performs better with DOS1-corrected reflectance.

Based on the given statistics, all models predicting total phosphorus by using Landsat images are achieving better results when they employ DOS1 corrected reflectance values.

As far as the application of WQ models on Sentinel 2 images is concerned, Chl-*a* models present a similar pattern to Landsat-based analysis. More particularly, only the Chl-$a_{general}$ model performs better with Sen2Cor-corrected reflectance whereas natural-only and artificial-only models are more successful when they employ DOS1 reflectance data. The same behavior is observed with Secchi models, where the general one presents better results with Sen2Cor methodwhereas natural and artificial ones perform better when employing the DOS1 corrected reflectance. Finally, the performance of $TP_{general}$ model is comparable with both types of corrected reflectance (Sen2Cor, DOS1) while the $TP_{natural}$ model offers better results when exploits the DOS1 correction method.

Table 3.3.1-1. Regression analysis basic statistics between the resulted values after the two-fold employment of SR (DOS1- and LaSRC; LEDAPS; Sen2Cor-corrected reflectance in GEE) in WQ models (developed by Markogianni et al., 2020; 2022) concerning datasets of year 2018. (Different reflectance products are referred as DOS1 and as GEE reflect. for LaSRC; LEDAPS and Sen2Cor correction methods. The units of ME and RMSE are μg/l, meters and mg/l for Chlorophyll-*a*, Secchi depth and Total phosphorus, respectively).

No.	WQ element	ME	RMSE	Equation (corrected WQ models)	N	R	R^2	Std. Error of the Estimate	Durbin-Watson	Sensor
1a	Chl-*a*_general (DOS1)	13.8	49.6	LogChl-a_general=0.221+0.61*(logChl-a_general_GEE)	115	0.88	0.78	0.19	1.99	Landsat
1b	Chl-*a*_general (GEE reflect.)	2.9	49.1							
2a	Chl-*a*_natural (DOS1)	1.8	20.7	LogChl-a_natural= -0.109+(0.747*logchl-a_natural_GEE)	26	0.87	0.76	0.23	2.03	
2b	Chl-*a*_natural (GEE reflect.)	-23.9	51.1							

3a	Chl-*a*_artificial (DOS1)	1.37	4.04	LogChl-a_artificial= 0.279+(0.209*logchl-a_artificial_GEE)	33	0.39	0.15	0.28	1.68	
3b	Chl-*a*_artificial (GEE reflect.)	1.9	4.3							
4a	Secchi_general (DOS1)	0.52	2.32	SQRTSecchi_general= 0.671+(0.737* SQRTSecchi_general_GEE)	95	0.91	0.83	0.25	1.99	Landsat
4b	Secchi_general (GEE reflect.)	1.15	2.5							
5a	Secchi_natural (DOS1)	0.7	2.87	SQRTSecchi_natural= 0.079+(0.875* SQRTSecchi_natural_GEE)	26	0.97	0.94	0.14	2.32	
5b	Secchi_natural (GEE reflect.)	-0.052	2.64							
6a	Secchi_artificial (DOS1)	-0.06	1.37	SQRTSecchi_artificial= 1.391+(0.475* SQRTSecchi_artificial_GEE)	33	0.88	0.78	0.15	1.72	
6b	Secchi_artificial (GEE reflect.)	1.75	2.41							
7a	TP_general (DOS1)	-0.003	0.029	LogTP_general= -0.127+(0.925* LOGTP_general_GEE)	28	0.99	0.98	0.03	2.25	
7b	TP_general (GEE reflect.)	-0.014	0.034							
8a	TP_natural (DOS1)	-0.003	0.028	LogTP_natural= -0.177+(0.796* LOGTP_natural_GEE)	40	0.96	0.93	0.08	1.83	
8b	TP_natural (GEE reflect.)	0.007	0.031							
9a	Application of TP general model on artificial	-0.015	0.026	LogTP_artificial = -0.143+(0.905* LOGTP_artificial	11	0.99	0.98	0.04	1.89	

	lakes (DOS1)			_GEE)						
9b	Application of TP general model on artificial lakes (GEE reflect.)	-0.023	0.039							
10a	Chl-*a*_general (DOS1)	2.77	11.5	LogChl-a_general= 0.338+(0.329*logchl-a_general_GEE)	64	0.6	0.36	0.43	2.18	Sentinel 2
10b	Chl-*a*_general (GEE reflect.)	0.56	7.6							
11a	Chl-*a*_natural (DOS1)	2.1	13.2	LogChl-a_natural= 0.125+(0.549*logchl-a_natural_GEE)	23	0.82	0.68	0.28	0.74	
11b	Chl-*a*_natural (GEE reflect.)	-28.9	73.5							
12a	Chl-*a*_artificial (DOS1)	2.6	7.7	LogChl-a_artificial= 0.06+(0.299*logchl-a_artificial_GEE)	33	0.38	0.15	0.62	1.9	
12b	Chl-*a*_artificial (GEE reflect.)	-10.2	75.9							
13a	Secchi_general (DOS1)	-0.78	4.5	SQRTSecchi_general= 1.646+(0.291*SQRTSecchi_general_GEE)	103	0.42	0.17	0.65	1.86	
13b	Secchi_general (GEE reflect.)	-0.2	2.9							
14a	Secchi_natural (DOS1)	-0.13	2.9	SQRTSecchi_natural= 0.233+(0.841*SQRTSecchi_natural_GEE)	24	0.97	0.95	0.13	1.44	
14b	Secchi_natural (GEE	-0.6	2.8							

	reflect.)	8								
15a	Secchi_artificial (DOS1)	0.15	1.46	SQRTSecchi_artificial= 1.490+(0.399* SQRTSecchi_artificial_GEE)	36	0.79	0.62	0.24	1.89	
15b	Secchi_artificial (GEE reflect.)	1.98	2.59							
16a	TP_general (DOS1)	0.006	0.033	LogTP_general= -0.086+(0.963* LOGTP_general_GEE)	27	0.99	0.98	0.04	1.84	Sentinel 2
16b	TP_general (GEE reflect.)	-0.003	0.033							
17a	TP_natural (DOS1)	-0.001	0.032	LogTP_natural= -0.178+(0.742* LOGTP_natural_GEE)	42	0.93	0.87	0.11	2.37	
17b	TP_natural (GEE reflect.)	0.02	0.04							
18a	Application of TP general model on artificial lakes (2 RECORDS) (DOS1)	-0.016		No linear regression analysis between TP values (accrued from the application of general model to artificial lakes) derived both from GEE and manual analysis (2018 dataset) due to few records						
18b	Application of TP general model on artificial lakes (2 RECORDS) (GEE Reflect.)	-0.011								

3.3.2 Validation of initial and corrected lake WQ models employing LaSRC, LEDAPS and Sen2Cor corrected reflectance values retrieved from the GEE platform

WQ models of Chl-*a*, Secchi depth and TP, developed by Markogianni et al. (2020); (2022), initially employed reflectance values from images in

GEEand were matched with corresponding *in-situ* datasets of 2019 and 2020 for both satellite sensors (mixed Landsat 7+ETM/8 OLI and Sentinel 2 images). Then, initial WQ models incorporated the developed corrected WQ models (Table 3.3.1-1) (noted as *_cor in Tables 3.3.1-2; 3.3.1-3; 3.3.1-4) and once again employed the same GEE-derived reflectance values of Landsat and Sentinel 2 images with basic aim the exploration of any further enhancement of each WQ element's quantification.

3.3.2.1 Chl-*a* models

Employment of GEE-retrieved reflectance values of Landsat and Sentinel images of 2019 in Chl-$a_{general}$ models (1a;4a; Table 3.3.1-2) yielded similar results based on ME and RMSE values while the employment of Sentinel reflectance resulted in lower MAPE values (122.3 vs 221.7).

Considering the corresponding corrected equations (1b; 4b), Landsat-employing model resulted in lower ME value (1.5 vs 3.7 μg/l) and slightly lower RMSE value (16.4 vs 16.8) compared to Sentinel-employing model. In general, the application of the corrected equations (1b; 4b) did not contribute to any further enhancement of Chl-*a* prediction (general model).

Regarding the Chl-$a_{natural}$ models, initial model employing Sentinel images (5a) performed better compared to Landsat (2a), while the corrected models (2b;5b) improved greatly the Chl-*a* prediction in natural-only lakes, especially regarding the Landsat-based model, and according to ME, RMSE and MAPE values (Table 3.3.1-2).

Chl-$a_{artificial}$ model achieved better results utilizing Landsat reflectance (3a vs. 12a), especially based on ME value, while no improvement was observed concerning the corrected models (3b; 6b), except for the MAPE value connected to Sentinel reflectance.

Table 3.3.1-2. Basic statistical error metrics evaluating the Chl-*a* models' performance in conjunction with *in-situ* WQ datasets of 2019. (The units of ME and RMSE are μg/l, NRMSE has no units while MAPE has percentage units).

No.	Model	ME	RMSE	NRMSE	MAPE	Sensor
1a	Chl-*a*_general	-1.34	16.4	0.11	221.7	**Landsat**
1b	Chl-*a*_general_cor	1.5	16.4	0.11	169.5	
2a	Chl-*a*_natural	-13.6	29.1	0.2	350.5	
2b	Chl-*a*_natural_cor	4.6	25.4	0.2	98.9	
3a	Chl-*a*_artificial	0.81	4.6	0.17	79.3	

3b	Chl-*a*_artificial_cor	1.71	4.9	0.18	87.5	
4a	Chl-*a*_general	1.65	15.4	0.11	122.3	**Sentinel 2**
4b	Chl-*a*_general_cor	3.7	16.8	0.12	116.8	
5a	Chl-*a*_natural	-3.6	14.98	0.11	177.7	
5b	Chl-*a*_natural_cor	7.4	25.4	0.18	84.9	
6a	Chl-*a*_artificial	1.9	4.5	0.17	82.7	
6b	Chl-*a*_artificial_cor	2.4	4.99	0.19	65.5	

3.3.2.2 Secchi Disk models

$Secchi_{general}$ models (Table 3.3.1-3; 1a; 4a) performed, in general, better than Chl-*a* ones and especially when employing Landsat data (1a) compared to Sentinel ones (4a). The Secchi corrected model enhanced in a great extent the assessment of Secchi depths by using Landsat images (1b) while the respective model employing Sentinel reflectance (4b) presented only a slight refinement (except for MAPE values in both cases).

$Secchi_{natural}$ models performed adequately regarding the prediction of Secchi depths while both the corrected models (2b; 5b) enhanced further their initial performance, especially the Sentinel-employing model (Table 3.3.1-3).

$Secchi_{artificial}$ models performed almost similarly regarding the sensor used (3a; 6a) yielding Secchi Depth values with adequate accuracy in relation to *in-situ* ones, while the corrected models (3b; 6b), further improved the Secchi prediction based on ME and RMSE values (Table 3.3.1-3).

Table 3.3.1-3. Basic statistical error metrics evaluating the Secchi Disk models' performance in conjunction with *in-situ* WQ datasets of 2019. (The units of ME and RMSE are meters, NRMSE has no units while MAPE has percentage units).

No.	Model	ME	RMSE	NRMSE	MAPE	Sensor
1a	Secchi_general	1.27	2.4	0.19	46.5	**Landsat**
1b	Secchi_general_cor	0.65	2.2	0.17	52.7	
2a	Secchi_natural	-0.53	1.83	0.15	56.2	
2b	Secchi_natural_cor	0.31	1.9	0.16	43.3	
3a	Secchi_artificial	1.55	2.5	0.3	51.7	
3b	Secchi_artificial_cor	-0.2	2.01	0.24	83.2	
4a	Secchi_general	-0.3	3.8	0.3	80.8	**Sentinel**

4b	Secchi_general_cor	-0.58	2.6	0.2	132.7	**2**
5a	Secchi_natural	-1.16	3.4	0.27	57.4	
5b	Secchi_natural_cor	-0.49	2.7	0.22	48.7	
6a	Secchi_artificial	1.31	2.4	0.29	60.2	
6b	Secchi_artificial_cor	-0.13	2.1	0.25	83.9	

3.3.2.3 TP models

$TP_{general}$ initial model performed slightly better with Sentinel images (Table 3.3.3-4; 4a) compared to Landsat (Table 3.3.3-4; 1a), while the corrected ones (Table 3.3.3-4; 1b; 4b) did not manage to improve the TP prediction with the exception of MAPE value.

$TP_{natural}$ model employing Sentinel data (Table 3.3.3-4; 5a) performed better compared to Landsat (Table 3.3.3-4; 2a) while the corrected models (Table 3.3.3-4; 2b; 5b) have not offered any significant differentiation.

Application of $TP_{general}$ model on artificial lakes sampled on 2019, presented better results when employing Landsat reflectance (3a; Table 3.3.3-4) rather than Sentinel. The corrected TP model (3b) improved slightly the performance of the initial one whereas no corrected TP model has been built for Sentinel data, due to existence of few records.

The application of all WQ models (including the corrected) on Landsat and Sentinel images of 2020 illustrated similar results with those accrued from the dataset of 2019; hence the corresponding statistical error metrics, are presented in the Appendix (Table 3).

Table 3.3.3-4. Basic statistical error metrics evaluating the TP models' performance in conjunction with *in-situ* WQ datasets of 2019. (The units of ME and RMSE are mg/l, NRMSE has no units while MAPE has percentage units).

No.	Model	ME	RMSE	NRMSE	MAPE	Sensor
1a	TP_general	0.07	0.34	0.15	53.8	**Landsat**
1b	TP_general_cor	0.08	0.36	0.16	47.5	
2a	TP_natural	0.05	0.26	0.11	58.3	
2b	TP_natural_cor	0.04	0.27	0.12	71.1	
3a	Application of TP general model on artificial lakes	-0.01	0.04	0.21	44.4	

3b	Application of TP general model on artificial lakes_cor	-0.002	0.035	0.19	37.5	
4a	TP_general	0.03	0.19	0.15	45.2	**Sentinel 2**
4b	TP_general_cor	0.039	0.19	0.15	41.2	
5a	TP_natural	0.03	0.14	0.12	54.7	
5b	TP_natural_cor	0.01	0.15	0.12	74.7	
6	Application of TP general model on artificial lakes	0.002	0.01	0.37	27.5	

3.3.2.4 All WQ models

Observing basic statistical indices and in particular values of ME per model, clustered by year and utilized satellite sensor, it can be concluded that Chl-*a* models resulted in higher divergences from *in-situ* values compared to Secchi and TP models (Figure 3.3.1-1a). Based on negative residual values, it seems that Chl-$a_{general}$ model employing Landsat reflectance overestimates Chl-*a* concentrations while the same applies for Chl-$a_{natural}$ models for both sensors but in greater extent for Landsat.

Secchi models have the same behavior based on the sampling year (except for $Secchi_{natural\_cor}$ model employing Sentinel data) but present differences based on the utilized sensor. $Secchi_{general}$ and $Secchi_{general\_cor}$ models using Landsat data seems to underestimate Secchi depths whilst respective models employing Sentinel 2 data overestimate those measurements. TP models in general indicated low residual values (Figure 3.3.1-1a).

The highest RMSE values are also accrued from the application of Chl-*a* models, followed by Secchi and TP models (Figure 3.3.1-1b). Distribution of RMSE values per WQ model is similar between the two years except for the value resulted after applying the Chl-$a_{natural}$ model on Landsat images of 2020. Additionally, Landsat-based Chl-*a* models suggest higher RMSE values compared to the respective Sentinel.

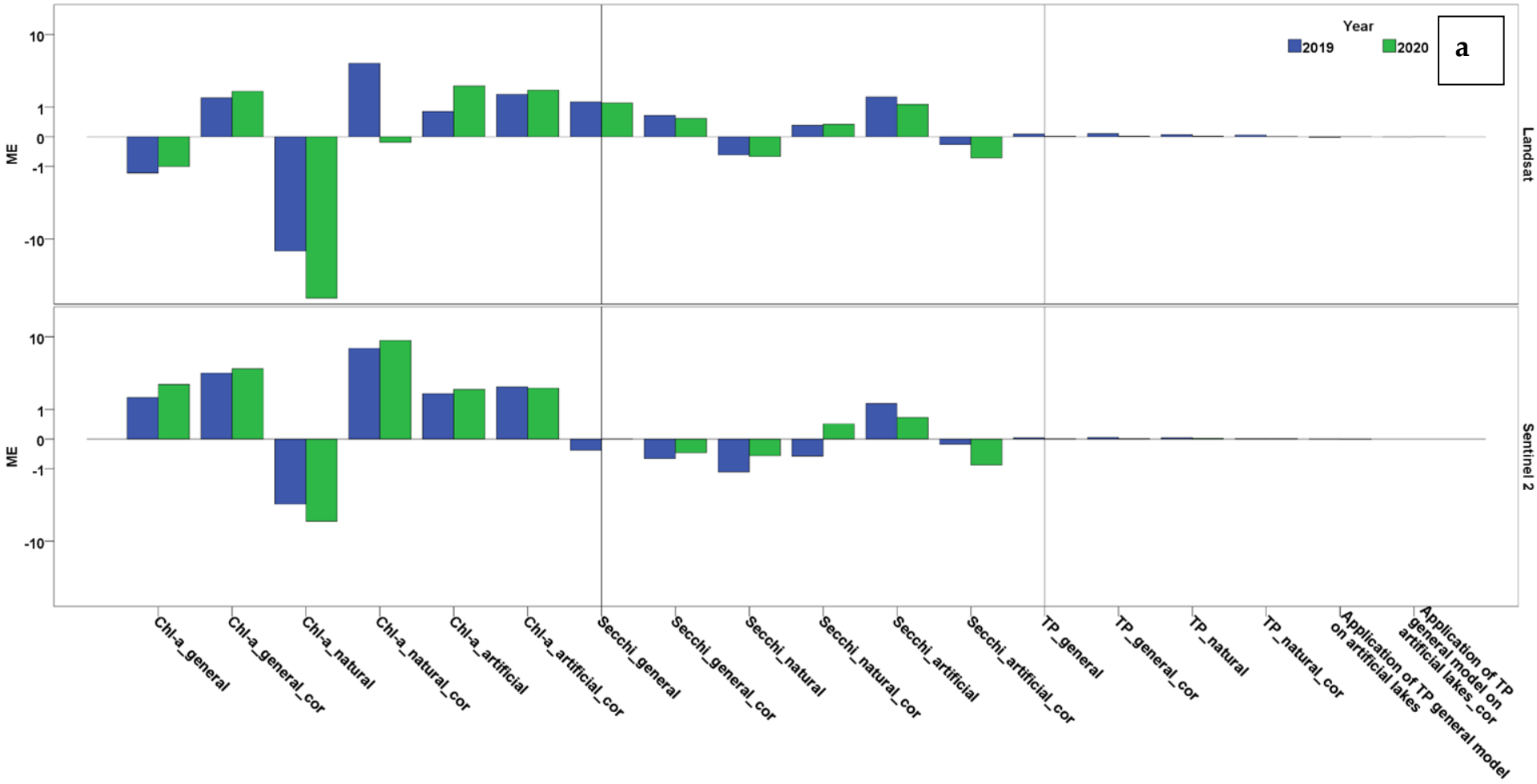
Examining the MAPE values derived from all WQ models and taking into consideration the threshold value of 50 (reasonable forecasting; Lewis 1982; Table 2.3.2-2), it can be concluded that mostly Secchi models, followed by the respective TP can be characterized as efficient enough to quantify each corresponding WQ element (Figure 3.3.1-1c). Concerning the application of Chl-$a_{general}$ models and the corresponding corrected ones, it can be declared that even though there has been an enhancement in models' performance,

MAPE values are still quite high independently from sensor or year (Figure 3.3.1-1c). Chl-$a_{artificial}$ models do not indicate any improvement concerning the year of 2020, but Sentinel-based model applied on 2019 dataset performs better than Landsat one (Figure 3.3.1-1c).

Even though $Secchi_{general}$ models have not been upgraded, the initial uncorrected ones presented highly acceptable MAPE values, indicating a good forecasting performance. In addition to Secchi models, $Secchi_{natural}$ models, after the fine tuning, have resulted in highly acceptable MAPE values. $Secchi_{artificial}$ models were not particularly enhanced but MAPE values accrued from the initial equations can guarantee a satisfactory Secchi quantification.

$TP_{general}$ models have been improved based on both years and utilized sensors and particularly the employment of 2019 dataset resulted in valuable outcomes (for both sensors). $TP_{natural}$ models were also not improved and concerning the initial models, only that employing Sentinel 2 reflectance of 2020, is considered reliable to use. As far as the application of $TP_{general}$ model on artificial lakes is considered, no safe conclusion can be drawn due to the existence of few available records. Despite this, Landsat-based models presented an improved performance for both studied years. Concerning the NRMSE metric, values' distribution is presented per WQ model clustered by satellite sensor and year (Figure 3.3.1-2). Thus, values range from 0.1 to 1.3, while both of them are observed in 2020 employing Landsat data in Chl-$a_{general}$ and Chl-$a_{natural}$ model, respectively. In general, low values close to 0 indicate the good performance of each respective WQ model while the median and average values of the whole dataset equal to 0.18 and 0.21, respectively.

Average NRMSE values per satellite sensor revealed a light superiority of models employing Landsat compared to Sentinel 2 reflectance (0.21 vs. 0.23) while the utilization of satellite and *in-situ* dataset of 2019 presented better performance than this of 2020 (0.18 vs. 0.24).


Year
2019
2020
a
ME
10
1
0
-1
-10
Landsat
Sentinel 2
Chl-a_general
Chl-a_general_cor
Chl-a_natural
Chl-a_natural_cor
Chl-a_artificial
Chl-a_artificial_cor
Secchi_general
Secchi_general_cor
Secchi_natural
Secchi_natural_cor
Secchi_artificial
Secchi_artificial_cor
TP_general
TP_general_cor
TP_natural
TP_natural_cor
Application of TP general model on artificial lakes
Application of TP general model on artificial lakes_cor

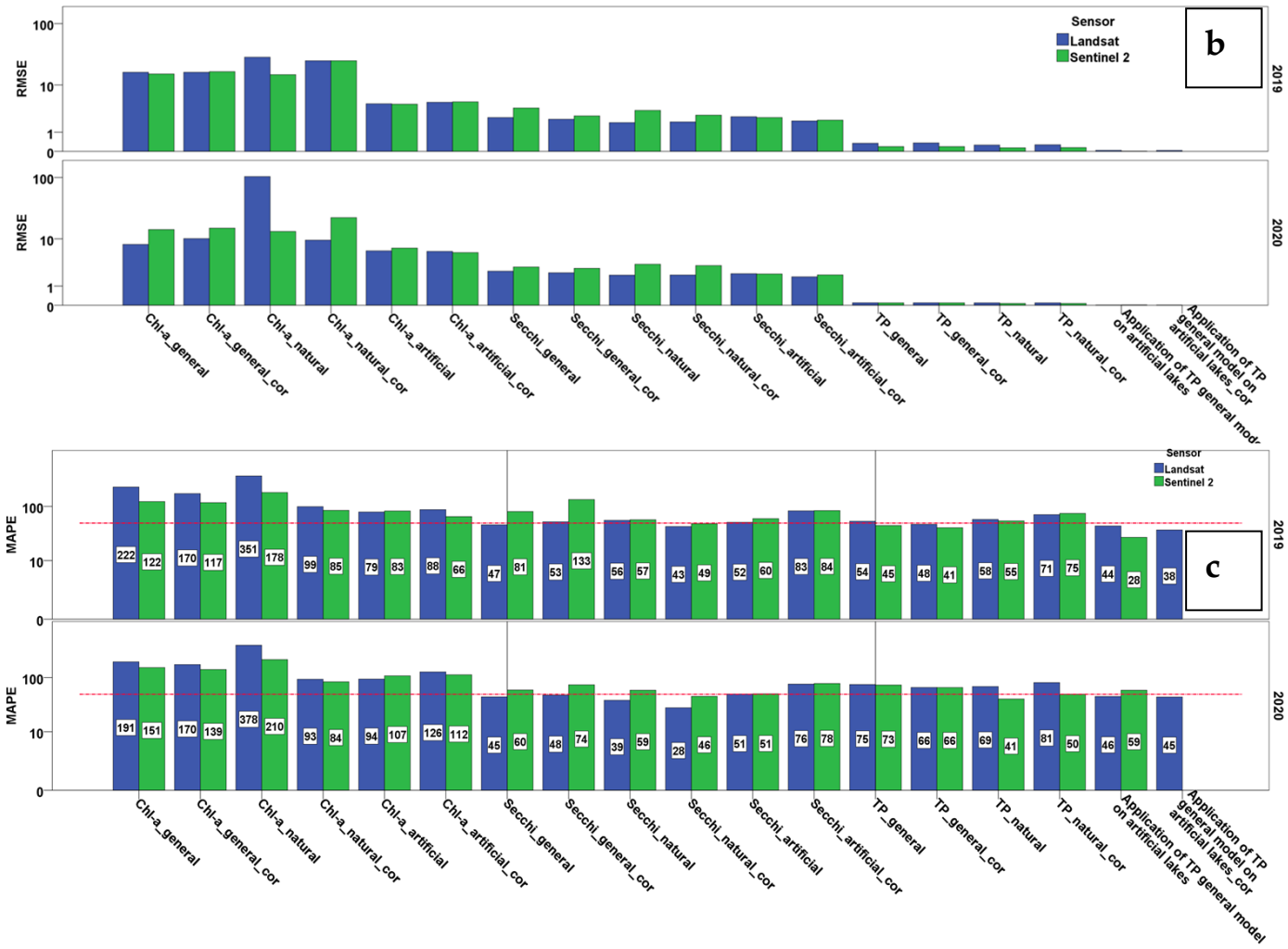


Figure 3.3.1-1. Distribution of (a) ME, (b) RMSE and (c) MAPE values per WQ model clustered by satellite sensor and year of sampling. (Red reference line to Y axis is set to 50, upper threshold value for reasonable forecasting.

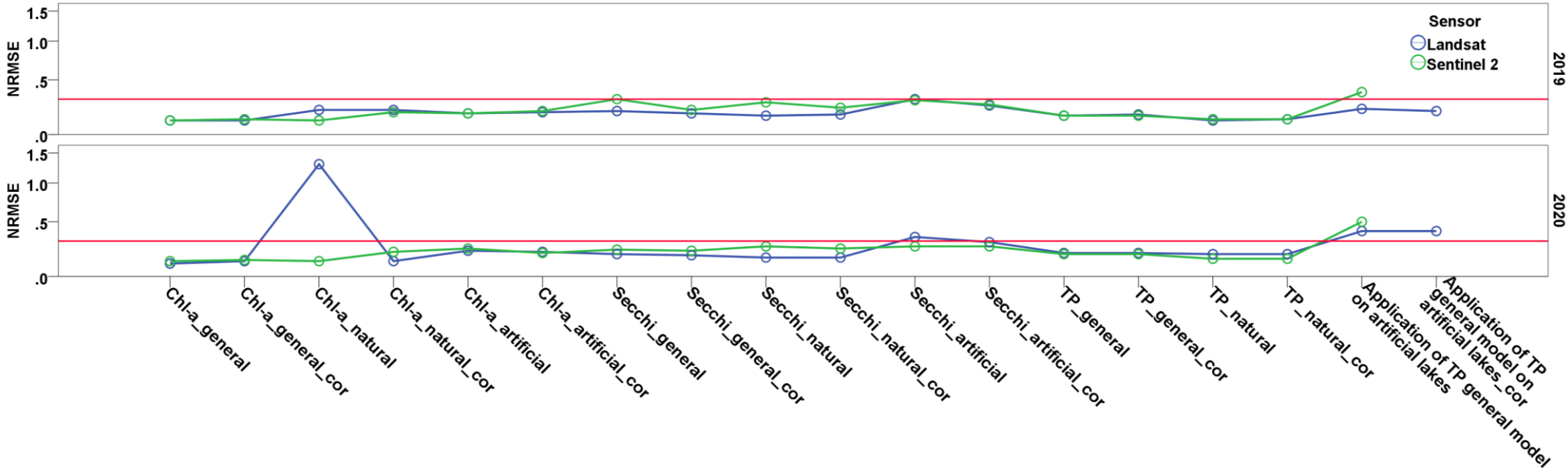


Figure 3.3.1-2. Distribution of NRMSE values per WQ model clustered by satellite sensor and grouped by year of sampling. (Red reference line to Y axis is set to 0.3 as an indicative low value).

3.4 Discussion

This study is perhaps the first attempt to facilitate the quantification of spatiotemporal lake WQ across the Greek Lake Monitoring Network of WFD, by using multi-sensor reflectance values retrieved from GEE platform.

Landsat (7 ETM+/8 OLI) and Sentinel 2 reflectance values in GEE were matched with concurrent WQ *in-situ* data of 2018 while the same pairs were created with reflectance derived from manually pre-processed respective images.

Published Landsat- based empirical WQ models of Chl-*a* (Markogianni et al., 2020), Secchi depth and Total Phosphorus (Markogianni et al., 2022) were applied twice employing two (2) different-atmospherically corrected reflectance values (DOS1 and other AC methods embedded in GEE) while linear regression analysis among resulted WQ values, separately for each sensor, yielded WQ-corrected linear equations accompanied by strong associations. Double employment (2018) of differently atmospheric corrected reflectance values in WQ models indicated the DOS1 as the most effective method for the quantification of lake WQ elements in almost all cases and for all sensors (Landsat/Sentinel 2); the only exceptions were the Chl-$a_{general}$ and $Secchi_{natural}$ models employing Landsat data, where LaSRC and LEDAPS methods were proved better and Chl-$a_{general}$ and $Secchi_{general}$ models employing Sentinel 2 data (Table 3.4-1; 2018 dataset), indicating their better performance after the application of Sen2Cor method. Results from several studies agree with the superiority of DOS method regarding the WQ monitoring of inland waters (Nazeer et al., 2014; Doña et al. 2014; El Alem et al., 2021; Abdelal et al. 2022). In particular, Abdelal et al. (2022) studied the extraction of WQ parameters in King Talal reservoir (Jordan) by testing several atmospheric correction methods, including DOS, in Landsat 8 and Sentinel 2 images. According to their atmospheric correction analysis, the DOS algorithm was the most successful in representing the Sentinel-2 satellite image while they recorded that it can be applied on images of both satellite sensors with not much accuracy loss which is not the case for the rest correction techniques examined (dark spectrum fitting -DSF-, atmospheric and topographic correction -ATCOR-, and exponential extrapolation -EXP). Furthermore, El Alem et al. (2021) compared image-based and physical correction models for retrieving suspended particulate matter (SPM) concentrations in lakes (United States and Canada) using Landsat imagery. Based on the results, image-based models, particularly the COST and DOS, are more appropriate than physical models for retrieving SPM concentrations in inland waters if the inputs of the physical atmospheric parameters are not well controlled. The basic assumption concerning the physical methods is that they usually use two or more NIR (or SWIR) wavebands, where the marine signal is assumed to be zero (open ocean waters). However, the signal in the NIR (or SWIR) is not

negligible in Case-2 waters, due to the concentrations of particulate matter in inland water bodies, and, consequently, maritime correction over inland water causes low or even negative water reflectance in the visible bands. As a consequence, reflectance over inland water bodies is assessed based on assumptions including 0 water-leaving radiance in the NIR (or SWIR) and aerosol origin/type models, resulting to the confusion of this natural optical–physical relationship (in terms of reflectance) of WQ parameters across the electromagnetic spectrum (El Alem et al., 2021).

Addittionally, cases of 2018 dataset where low values of coefficient of determination among same-located WQ values were observed, concern mainly Chl-*a* models (general and artificial) employing Sentinel 2 reflectance. On one hand it is well known that mapping Chl-*a* in Case 2 waters is a complicated task and characterized by less accuracy since the optical properties are measured based on a compound of dissolved organic matter, dead organic-inorganic particulate matter, and phytoplankton (Chl-*a*; Markogianni et al., 2020). On the other hand, hereby utilized WQ empirical models have been developed based on Landsat-7 ETM+ and 8 OLI images which were atmospherically corrected with the DOS1 method; hence it is expected to be affected not only by this factor but also by the corresponding spectral composition and eventually perform better when employing Landsat rather than Sentinel 2 reflectance, as it is hereby observed (Table 3.4-1; datasets of 2019 and 2020). Major exceptions constitute the Chl-$a_{natural}$, $TP_{general}$ and $TP_{artificial}$ models which seem to present more reliable results, for both validation years, with Sentinel 2 images.

Concerning the question whether the corrected WQ models contribute to the improvement of WQ elements' quantification, the answer is, in general, positive. In particularly, regarding Chl-*a* models, Chl-$a_{natural}$ model is presented widely enhanced for both satellite sensors and validation years (except for Sentinel 2 in 2020). Secchi models (general, natural, artificial) illustrated the greatest improvement, compared to Chl-*a* and TP models, with the exception of general and artificial-only models employing Sentinel 2 images in 2020. TP models also provided refined values based on *in-situ* datasets, except for the natural-only ones.

Table 3.4-1. Summarized results indicating the best performance of empirical WQ models employing a) different AC processors (2018 dataset) and b) GEE-derived reflectance values (2019, 2020) and exploration of the correction necessity via the application of the sensor-specific models. (x symbol denotes the best performance among the sensors used; NO* denotes that only reduced MAPE values were observed while YES* denotes increased MAPE values).

2018	**Chl-$a_{general}$**	**Chl-$a_{natural}$**	**Chl-$a_{artificial}$**	**Secchi$_{general}$**	**Secchi$_{natural}$**	**Secchi$_{artificial}$**	**TP$_{general}$**	**TP$_{natural}$**	**TP$_{artificial}$**
Landsat	LaSRC, LEDAPS	DOS1	ALL	DOS1	LaSRC, LEDAPS	DOS1	DOS1	DOS1	DOS1
Sentinel 2	Sen2Cor	DOS1	DOS1	Sen2Cor	DOS1	DOS1	ALL	DOS1	DOS1
2019	**Chl-$a_{general}$**	**Chl-$a_{natural}$**	**Chl-$a_{artificial}$**	**Secchi$_{general}$**	**Secchi$_{natural}$**	**Secchi$_{artificial}$**	**TP$_{general}$**	**TP$_{natural}$**	**TP$_{artificial}$**
Landsat	**X**		**X**	**X**	**X**	**X**			**X**
Sentinel 2	**X**	**X**					**X**	**X**	
ENHANCEMENT									
Landsat	**NO***	**YES**	**NO**	**YES**	**YES**	**YES***	**YES**	**NO**	**YES**
Sentinel 2	**NO**	**YES**	**NO***	**YES**	**YES**	**YES***	**YES**	**NO**	NO MODEL
2020	**Chl-$a_{general}$**	**Chl-$a_{natural}$**	**Chl-$a_{artificial}$**	**Secchi$_{general}$**	**Secchi$_{natural}$**	**Secchi$_{artificial}$**	**TP$_{general}$**	**TP$_{natural}$**	**TP$_{artificial}$**
Landsat	**X**		**X**	**X**	**X**	**X**	**X**		**X**
Sentinel 2		**X**					**X**	**X**	
ENHANCEMENT									
Landsat	**NO**	**YES**	**NO**	**YES***	**YES**	**YES***	**YES**	**NO**	**YES**
Sentinel 2	**NO***	**NO***	**NO***	**NO**	**YES**	**NO**	**YES**	**NO**	NO MODEL

Chl-$a_{general}$ model employing Landsat reflectance yielded for both validation years, an average RMSE of 12.21 μg/l whereas the corresponding value related to Sentinel 2 data equals to 14.9 μg/l. The Chl-$a_{natural}$ corrected model is proposed to be utilized in conjunction with Landsat data while the average RMSE value is 17.45 μg/l. The Chl-$a_{artificial}$ model (without correction) presented lower RMSE values, 5.4 and 5.7 μg/l for Landsat and Sentinel 2 reflectance, respectively.

To our knowledge, there are a few recent studies trying to estimate Chl-*a* concentrations at a large regional scale with GEE. Lin et al. (2018) combined *in*

situ Chl-*a* data from 1157 lakes (2007) with Landsat data and developed a well-validated lake national model (RMSE = 34.9 μg/L), by using machine learning algorithms built into the GEE. Wang et al. (2020) used GEE to automatically form match-up points from multi-sensor satellite observations with ground WQ samples and then an SVM was developed to map Chl-*a* concentrations across 12 lakes in the tri-state region of Kentucky, Indiana and Ohio (USA). Furthermore, RMSE of Chl-*a* of the SVM model trained by Landsat 8 OLI imagery was 4.42 μg/L. Kislik et al. (2022) analyzed four spectral indices - Normalized Difference Vegetation Index (NDVI), Normalized Difference Chlorophyll Index (NDCI), B8AB4, and B3B2 - to retrieve chlorophyll-*a* data for algal bloom identification in two highly dynamic freshwater reservoirs by using Sentinel 2-MSI in GEE. Among the results, NDCI most accurately identified chlorophyll-*a* across all study sites (highest adjusted R^2 = 0.84, lowest RMSE = 0.02 μg/l), followed by NDVI. A few studies have also been conducted to estuarine and marine environments. Li et al. (2022) extracted Chl-*a* concentrations of SeaWiFS and Terra/Aqua MODIS embedded in GEE across the Yellow Sea to examine their relationship to green tide while Bioresita et al. (2021) used Sentinel 2 images through the GEE platform to monitor Chl-*a* concentrations in the Kali Porong Estuary (Indonesia) employing certain estimation formulas.

Except for GEE utilization, other studies of remotely estimation of Chl-*a* concentrations have yielded comparable and even higher RMSE values than ours; Bonansea et al. (2018) presented an RMSE of 18.47 μg/l in the largest artificial reservoir in Córdoba province (Rio Tercero, Argentina) while Doña et al. (2014) showed an RMSE of 40 μg/l across certain Spanish lakes and ponds. Additionally, Zhang et al. (2020) developed an SVM model on Landsat 8 OLI images to estimate the Chl-*a* concentrations of multiple lakes in China, while they reported an RMSE of 22.64 μg/L.

As far as the Secchi models are concerned, all of corrected ones presented enhanced results; $Secchi_{general}$ yielded average 2-year RMSE values of 2.22 m (Landsat) and 2.7 m (Sentinel 2), $Secchi_{natural}$ model 1.95 m (Landsat) and 2.95 m (Sentinel 2) while the respective RMSE values resulted from the $Secchi_{artificial}$ model equal to 1.91 m (Landsat) and 2.05 m (Sentinel 2). One of the few studies that combined the derivation of Secchi depth in reservoirs and GEE, was conducted by Somasundaram et al. (2021) while the resulted RMSE value is particularly low and equals to 32.6 cm. Considering the high difference between this RMSE value and the hereby derived one, it should be noticed that Somasundaram et al. (2021) applied a Zsd (Secchi Disk depth) model consisted of a combination of the Normalized Difference Chlorophyll Index (NDCI) and a mechanistic model for the derivation of the absorption coefficient and backscattering. According to literature (IOCCG, 2006) those models depict more applicability and reliable results compared to those

utilizing the relationship between WQ parameters and *in-situ* measurements. Furthermore, Zhang et al. (2021) documented the most recent Secchi Disk estimation models used in previous studies since 1993, based on remote sensing techniques. Referring to the comparison of their performance, the average RMSE value among those studies is 1.13 m while the highest (1.7 m) has been recorded by Allan et al. (2011).

Considering the performance of TP models, it is evident that only the $TP_{general}$ model needs the corrected version while the specially developed model for natural-only lakes stands efficiently without correction. Corrected $TP_{general}$ model yielded RMSE values of 0.23 mg/l (Landsat) and 0.14 mg/l (Sentinel 2), $TP_{natural}$ model 0.18 mg/l (Landsat) and 0.11 mg/l (Sentinel 2) while the application of $TP_{general}$ model on artificial lakes- illustrated only in Landsat images- resulted in an RMSE value of 0.02 mg/l. During a thorough literature review, none recent study was detected utilizing GEE for the quantification of total phosphorus concentrations in lakes. Nevertheless, a survey was conducted to record RMSE values of remotely-sensed phosphorus concentrations to compare with hereby results. Zeng et al. (2022) developed a novel-semi-analytical algorithm in the eutrophic Lake Taihu, China and the validation showed satisfactory performance (RMSE=0.01 mg/l). Lim and Choi (2015), who also constructed multiple regression equations to retrieve total phosphorus concentrations in Nakdong river Korea, reported a TP regression model accompanied by an RMSE value of 0.01 mg/l. Lastly, Song et al. (2012) established a hybrid model combining genetic algorithms and partial least square (GA-PLS) to estimate remotely TP concentrations in 3 central Indiana reservoirs and RMSE values ranged from 0.009 to 0.03 mg/l, depending on *in-situ* datasets.

Taking into consideration the hereby-developed WQ models' evaluation, it is proven that GEE public data is sufficient for mapping Chl-*a*, Secchi depth and TP concentrations in a large geographical region and particularly at a national scale (Greece). Even though the WQ models were developed based on multiple linear regression analyses (MLRs) and Landsat 7 +ETM and 8 OLI images, their efficiency was indicated when were applied in GEE images, despite the pre-processing differentiation.

3.5 Conclusions

Estimation of important WQ elements in lakes across Greece employing satellite data embedded in the GEE platform, facilitates the monitoring and the estimation of their trophic status at a national scale. Hereby derived results indicated that WQ models, empirically developed, are applicable to both archived and future Landsat and Sentinel 2 image data despite the different pre-processing methodologies applied. Further, the aforementioned

models were trained based on a big dataset collected over different lakes with various optical properties and covering a long enough period of time.

Efficient application of empirical WQ models in GEE platform exempt users from the complicated AC of raw image products, key procedure for achieving stable performance. The hereby results confirm the spatio-temporal stability of the models while when combined with GEE-retrieved SR, offer scientists and Greek competent authorities the opportunity to exploit this massive warehouse of data for map long-term trend in WQ of lakes and identify the underlying factors and possible pollutant threats.

Acknowledgments: I wish to thank for the freely shared data used in this study, collected in the framework of the National Water Monitoring Network for lakes, according to the Joint Ministerial Decision 140384/2011, implemented by The Goulandris Natural History Museum, Greek Biotope/Wetland Centre (EKBY). I would also like to thank Mr. Rigas Giovos (Geographer at the GIS Research Unit of Agricultural University of Athens) for extracting the multi-sensor reflectance values from the GEE platform.

3.6 References

Abdelal Qasem, Assaf N. Mohammed, Al-Rawabdeh Abdulla, Arabasi Sameer, Nathir A. Rawashdeh. Assessment of Sentinel-2 and Landsat-8 OLI for Small-Scale Inland Water Quality Modeling and Monitoring Based on Handheld Hyperspectral Ground Truthing. *Journal of Sensors* **2022**, vol. 2022, Article ID 4643924, 19 pages, https://doi.org/10.1155/2022/4643924.

Allan M.G., Hamilton D.P., Hicks B.J., Brabyn L. Landsat remote sensing of chlorophyll a concentrations in central north island lakes of New Zealand. *Int. J. Remote Sens.* **2011**, 32, 2037–2055.

Bioresita, Filsa *et al*. *IOP Conf. Ser.: Earth Environ. Sci.* **2021**, 936 012011

Bonansea Matias, Bazán Raquel, Ferrero Susana, Rodríguez Claudia, Ledesma Claudia, Pinotti Lucio. Multivariate statistical analysis for estimating surface water quality in reservoirs. *Int. J. Hydr. Sci. and Technol.* **2018**, 8(**1),** 52-68, https://doi.org/10.1504/IJHST.2018.088675.

Bramich J., Bolch C.J.S. Fischer A. Improved red-edge chlorophyll-a detection for Sentinel 2. *Ecol. Indic.* **2021**, 120, 106876.

Bresciani M., Cazzaniga I., Austoni M., Sforzi T., Buzzi F., Morabito G., Giardino C. Mapping phytoplankton blooms in deep subalpine lakes from Sentinel-2A and Landsat-8. *Hydrobiologia* **2018**, 824, 197–214.

Brezonik, P. L; Olmanson, G. L.; Finlay, J.C.; Bauer, M. E. Factors affecting the measurement of CDOM by remote sensing of optically complex inland waters. *Remote Sensing of Environment*, **2015**, 157, 199-215, ISSN 0034-4257, http://dx.doi.org/10.1016/j.rse.2014.04.033.

Chavez, P.S., Jr. An Improved Dark-Object Subtraction Technique for Atmospheric Scattering Correction of Multispectral data. *Remote Sens. Environ.* **1988**, 24, 459–479.

Doña, C., Sánchez M. J., Caselles V., Jose Antonio Domínguez, and Antonio

Camacho, Empirical Relationships for Monitoring Water Quality of Lakes and Reservoirs Through Multispectral Images, *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **2014**, 7(5).

El Alem, A.; Lhissou, R.; Chokmani, K.; Oubennaceur, K. Remote Retrieval of Suspended Particulate Matter in Inland Waters: Image-Based or Physical Atmospheric Correction Models? *Water* **2021**, 13, 2149. https://doi.org/10.3390/w13162149.

Ermida, S.L.; Soares, P.; Mantas, V.; Göttsche, F.-M.; Trigo, I.F. Google Earth Engine Open-Source Code for Land Surface Temperature Estimation from the Landsat Series. *Remote Sens.* **2020**, *12*, 1471. https://doi.org/10.3390/rs12091471.

Fatima, R. Studies on physical, chemical and bacteriological characteristics on quality of spring water in Hajigak iron ore mine, Bamyan province, central Afghanistan, *Open J. Geol.* **2018**, vol. 8, pp. 313–332.

Gholizadeh, M.; Melesse, A.; Reddi, L. A comprehensive review on water quality parameters estimation using remote sensing techniques. *Sensors* **2016**, 16 (8), 1298. https://doi.org/10.3390/s16081298.

Gomes, V.C.F.; Queiroz, G.R.; Ferreira, K.R. An Overview of Platforms for Big Earth Observation Data Management and Analysis. *Remote Sens.* **2020**, *12*, 1253. https://doi.org/10.3390/rs12081253.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., **2017**. Google Earth Engine: planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment* **2017**, 202, 18-27, https://doi.org/10.1016/j.rse.2017.06.031.

IOCCG. Remote Sensing of Inherent Optical Properties: Fundamentals, Tests of Algorithms, and Applications; Lee, Z., Ed.; Reports of the International Ocean-Colour Coordinating Group, No. 5; IOCCG: Dartmouth, NS, Canada, **2006**; ISBN 9781896246567.

Jia, T.; Zhang, X.; Dong, R. Long-Term Spatial and Temporal Monitoring of Cyanobacteria Blooms Using MODIS on Google Earth Engine: A Case Study in Taihu Lake. *Remote Sensing* **2019**, 11, 2269.

Kislik, Chippie; Dronova, Iryna; Grantham, E. Theodore; Kelly, Maggi. Mapping algal bloom dynamics in small reservoirs using Sentinel-2 imagery in Google Earth Engine. *Ecological Indicators* **2022**, Volume 140, 109041, ISSN 1470-160X, https://doi.org/10.1016/j.ecolind.2022.109041.

Kumar, L.; Mutanga, O. Google Earth Engine applications since inception: Usage, trends, and potential. *Remote Sens.* **2018**, 10, 1509.

Lewis, C. D. Industrial and business forecasting methods, London: Butterworths, 1982.

Li, J. *et al.* MODIS observations of water color of the largest 10 lakes in china between 2000 and 2012. *Int. J. Digit. Earth* **2016**, vol. 9, no. 8, pp. 788–805.

Li, W.; Yang, Q.; Ma, Y.; Yang, Y.; Song, K.; Zhang, J.; Wen, Z.; Liu, G. Remote Sensing Estimation of Long-Term Total Suspended Matter Concentration from Landsat across Lake Qinghai. *Water* **2022**, *14*, 2498. https://doi.org/10.3390/w14162498.

Lim, J.; Choi, M. Assessment of water quality based on Landsat 8 operational land imager associated with human activities in Korea. *Environ. Monit. Assess.* **2015**, 187, 384.

Lin, S.; Novitski, L.N.; Qi, J.; Stevenson, R.J. Landsat TM/ETM+ and machine-learning algorithms for limnological studies and algal bloom management of

inland lakes. *J. Appl. Remote Sens.* **2018**, 12, 1.

Lobo, F. de L.; Nagel, G.W.; Maciel, D.A.; Carvalho, L.A.S. de; Martins, V.S.; Barbosa, C.C. F.; Novo, E.M.L. de M. AlgaeMAp: Algae Bloom Monitoring Application for Inland Waters in Latin America. *Remote Sensing* **2021**, 13, 2874.

Maciel DA, Barbosa CCF, de Moraes Novo EML, Junior RF and Begliomini FN. Water clarity in Brazilian water assessed using Sentinel-2 and machine learning methods. *ISPRS Journal of Photogrammetry and Remote Sensing* **2021**, 182: 134–152.

Maeda, E.E.; Lisboa, F.; Kaikkonen, L.; Kallio, K.; Koponen, S.; Brotas, V.; Kuikka, S. Temporal patterns of phytoplankton phenology across high latitude lakes unveiled by long-term time series of satellite data. *Remote Sens. Environ.* **2019**, 221, 609–620.

Mandanici, E.; Bitelli, G. Multi-image and multi-sensor change detection for long-term monitoring of arid environments with Landsat series. *Remote Sens.* **2015**, 7, 14019–14038.

Markogianni, V.; Kalivas, D.; Petropoulos, G.; Dimitriou, E. An appraisal of the potential of Landsat 8 in estimating chlorophyll-a, ammonium concentrations and other water quality indicators. *Remote Sens.* **2018**, 10, 1018.

Markogianni, V.; Kalivas, D.; Petropoulos, G.P.; Dimitriou, E. Estimating Chlorophyll-a of Inland Water Bodies in Greece Based on Landsat Data. *Remote Sens.* **2020**, 12, 2087. https://doi.org/10.3390/rs12132087.

Markogianni, V.; Kalivas, D.; Petropoulos, G.P.; Dimitriou, E. Modelling of Greek Lakes Water Quality Using Earth Observation in the Framework of the Water Framework Directive (WFD). *Remote Sens.* **2022**, 14, 739. https://doi.org/10.3390/rs14030739.

Nazeer M., J. E. Nichol, Y.-K. Yung. Evaluation of Atmospheric Correction Models and Landsat Surface Reflectance Product in an Urban Coastal Environment. *International Journal of Remote Sensing,* **2014**, 35 (16): 6271–6291. doi:https://doi.org/10.1080/01431161.2014.951742.

Pahlevan Nima, Antoine Mangin, Sundarabalan V. Balasubramanian, Brandon Smith, Krista Alikas, Kohei Arai, Claudio Barbosa, Simon Bélanger, Caren Binding, Mariano Bresciani, Claudia Giardino, Daniela Gurlin, Yongzhen Fan, Tristan Harmel, Peter Hunter, Joji Ishikaza, Susanne Kratzer, Moritz K. Lehmann, Martin Ligi, Ronghua Ma, François-Régis Martin-Lauzer, Leif Olmanson, Natascha Oppelt, Yanqun Pan, Steef Peters, Nathalie Reynaud, Lino A. Sander de Carvalho, Stefan Simis, Evangelos Spyrakos, François Steinmetz, Kerstin Stelzer, Sindy Sterckx, Thierry Tormos, Andrew Tyler, Quinten Vanhellemont, Mark Warren. ACIX-Aqua: A global assessment of atmospheric correction methods for Landsat-8 and Sentinel-2 over lakes, rivers, and coastal waters, *Remote Sensing of Environment*, **2021**, 258,112366, https://doi.org/10.1016/j.rse.2021.112366.

Pahlevan, N., B. Smith, J. Schalles, C. Binding, Z. Cao, M. Ronghua, K. Alikas, et al. Seamless Retrievals of Chlorophyll-A from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in Inland and Coastal Waters: A Machine-learning Approach. *Remote Sensing of Environment,* ***2020***, 240 :111604. doi: https://doi.org/10.1016/j.rse.2019.111604.

Pedreros-Guarda, M.; Abarca-del-Río, R.; Escalona, K.; García, I.; Parra, Ó. A Google Earth Engine Application to Retrieve Long-Term Surface Temperature for Small Lakes. Case: San Pedro Lagoons, Chile.

Remote Sens. **2021**, 13, 4544. https://doi.org/10.3390/ rs13224544.

Pizani, F.; Maillard, P. The determination of water quality parameters by remote sensing technologies: 2000 -2020. Universidade Federal de Minas Gerais, [S. l.] **2022** p. 1–30. DOI: 10.13140/RG.2.2.32203.87849.

Sagan, V.; Peterson, T.K.; Maimaitijiang, M.; Sidike, P.; Sloan, J.; Greeling, A.B.; Samar, M.; Adams, C. Monitoring inland water quality using remote sensing: Potential and limitations of spectral indices, bio-optical simulations, machine learning, and cloud computing. *Earth-Sci. Rev.* **2020**, 205, 103187.

Somasundaram, D.; Zhang, F.; Ediriweera, S.; Wang, S.; Yin, Z.; Li, J.; Zhang, B. Patterns, Trends and Drivers of Water Transparency in Sri Lanka Using Landsat 8 Observations and Google Earth Engine. *Remote Sens.* **2021**, 13, 2193. https://doi.org/10.3390/rs13112193.

Song, K., Wang, Q., Liu, G., Jacinthe, P.-A., Li, S., Tao, H., Du, Y., Wen, Z., Wang, X., Guo, W., Wang, Z., Shi, K., Du, J., Shang, Y., Lyu, L., Hou, J., Zhang, B., Cheng, S., Lyu, Y., Fei, L., 2022. A unified model for high resolution mapping of global lake (>1 ha) clarity using landsat imagery data. *Sci. Total Environ.* **2022**, 810, 15118, https://doi.org/10.1016/j.scitotenv.2021.151188.

Song, K.; Li, L.; Li, S.; Tedesco, L.; Hall, B.; Li, L. Hyperspectral Remote Sensing of Total Phosphorus (TP) in Three Central Indiana Water Supply Reservoirs. *Water Air Soil Pollut.* **2012**, 223, 1481–1502.

Topp, S.; Pavelsky, T.; Jensen, D.; Simard, M.; Ross, M. Research trends in the use of remote sensing for inland water quality science: Moving towards multidisciplinary applications. *Water* **2020**, 12, 169.

Vaičiūtė, D.; Bučas, M.; Bresciani, M.; Dabulevičienė, T.; Gintauskas, J.; Mėžinė, J.; Tiškus, E.; Umgiesser, G.; Morkūnas, J.; De Santi, F.; Bartoli, M. Hot moments and hotspots of cyanobacteria hyperblooms in the Curonian Lagoon (SE Baltic Sea) revealed via remote sensing-based retrospective analysis. *Sci. Total Environ.* **2021**, 769, 145053.

Wang, L.; Xu, M.; Liu, Y.; Liu, H.; Beck, R.; Reif, M.; Emery, E.; Young, J.; Wu, Q. Mapping Freshwater Chlorophyll-*a* Concentrations at a Regional Scale Integrating Multi-Sensor Satellite Observations with Google Earth Engine. *Remote Sens.* **2020**, *12*, 3278. https://doi.org/10.3390/rs12203278.

Warren, M.A., Simis, S.G.H., Martinez-Vicente, V., Poser, K., Bresciani, M., Alikas, K., et al. Assessment of atmospheric correction algorithms for the Sentinel-2A MultiSpectral Imager over coastal and inland waters. *Rem. Sens. Environ*. **2019**, 225, 267–289, https://doi.org/10.1016/j.rse.2019.03.018.

Weber, S.J.; Mishra, D.R.; Wilde, S.B.; Kramer, E. Risks for cyanobacterial harmful algal blooms due to land management and climate interactions. *Sci. Total Environ.* **2020**, 703, 134608.

Wen Zhidan; Qiang, Wang; Ge, Liu; Pierre-Andre, Jacinthe; Xiang, Wang; Lili, Lyu; Hui, Tao; Yue, Ma; Hongtao, Duan; Yingxin, Shang; Baohua, Zhang; Yunxia, Du; Jia, Du; Sijia, Li; Shuai, Cheng; Kaishan, Song. Remote sensing of total suspended matter concentration in lakes across China using Landsat images and Google Earth Engine, *ISPRS Journal of Photogrammetry and Remote Sensing* **2022**, Volume 187, Pages 61-78, ISSN 0924-2716, https://doi.org/10.1016/j.isprsjprs.2022.02.018.

Zeng Shuai; Shaohua, Lei; Yunmei, Li; Heng, Lyu; Xianzhang, Dong; Junda, Li;

Xiaolan, Cai. Remote monitoring of total dissolved phosphorus in eutrophic Lake Taihu based on a novel algorithm: Implications for contributing factors and lake management. *Environmental Pollution* **2022**, Volume 296, 118740, ISSN 0269-7491, https://doi.org/10.1016/j.envpol.2021.118740.

Zhang F. *et al.* Validation of a synthetic chlorophyll index for remote estimates of chlorophyll-a in a turbid hypereutrophic lake. *Int. J. Remote Sens.* **2014**, vol. 35, no. 1, pp. 289–305.

Zhao, Q.; Yu, L.; Du, Z.; Peng, D.; Hao, P.; Zhang, Y.; Gong, P. An Overview of the Applications of Earth Observation Satellite Data: Impacts and Future Trends. *Remote Sens.* **2022**, 14, 1863. https://doi.org/10.3390/rs14081863.

Zhou Yadong, Liu Hui, He Baoyin, Yang Xiaoqing, Feng Qi, Kutser Tiit, Chen Feng, Zhou Xinmeng, Xiao Fei, Kou Jiefeng. Secchi Depth estimation for optically-complex waters based on spectral angle mapping - derived water classification

https://sentinels.copernicus.eu/web/sentinel/user-guides/sentinel-2-msi/processing-levels/level-2

Zhang, T.; Huang, M.; Wang, Z. Estimation of chlorophyll-*a* Concentration of lakes based on SVM algorithm and Landsat 8 OLI images. *Environ Sci Pollut Res* **2020**, 27, 14977–14990. https://doi.org/10.1007/s11356-020-07706-7.

Zhang, Y.; Zhang, Y.; Shi, K.; Zhou, Y.; Li, N. Remote sensing estimation of water clarity for various lakes in China. *Water Res.* **2021**, 192, 116844.

using Sentinel-2 data, *Int. J. Remote Sens.* ***2021***, 42:8, 3123-3145, DOI: 10.1080/01431161.2020.1868606.

Zong, J.-M.; Wang, X.-X.; Zhong, Q.-Y.; Xiao, X.-M.; Ma, J., Zhao; B. Increasing Outbreak of Cyanobacterial Blooms in Large Lakes and Reservoirs under Pressures from Climate Change and Anthropogenic Interferences in the Middle-Lower Yangtze River Basin. *Remote Sensing* **2019**, 11, 1754.

4. Operational Development of Techniques for Characterizing Water Quality of oligotrophic Case-2 waters

4.1 Analysis on the WQ models' Performance in Oligotrophic Case-2 waters

4.1.1 Introduction

The classification of waters in Case 1 (oceanic) and Case 2 (coastal regions, rivers, and lakes), refined by Gordon and Morel (1983), is characterized by great importance when RS techniques are utilized to monitor their WQ and/or trophic status. The distinction between the two cases has some significant effects on the interpretation and modelling of optical data. In particular, according to this classification scheme, the optical properties of Case 1 waters are determined by phytoplankton and co-varying substances, while Chl-*a* is considered a proxy of phytoplankton concentration. This assumption has facilitated the implementation of large-scale optical models and the development of Chl-*a* predicting algorithms for Case 1 waters (Markogianni et al., 2022).

It is, however, acknowledged that Case 2 waters are more complex than Case 1 concerning their composition and optical properties. Hence, satellite ocean color algorithms, primarily developed for ocean, cannot be always applied to lakes due to, except for optical complexity, atmospheric conditions, altitudes and land proximity (IOCCG, 2018; Seegers et al., 2021). One of the main factors hindering accurate WQ monitoring in Case 2 waters is the fact that suspended material, yellow substances, and perhaps bottom reflectance vary independently of each other. Moreover, alterations in optical signal and the concentrations of the dissolved constituents are often so small that they hinder the ability to extract reliable information (Gholizadeh et al., 2016).

Hence, given the difficulty that WQ monitoring of Case 2 waters constitutes a multi-variable, non-linear problem, it is more realistic to establish a series of algorithms rather than a single all-purpose one. In this way, more than one algorithm contributes to capturing and solving the problem for all variables and over several and different ranges of concentrations (IOCCG, 2000). Those different ranges of concentrations correspond to classes of trophic status. Carlson (1977) developed a method of trophic status classification considering Chl-*a* and phosphorus concentrations and Secchi disk depths (*ZSD*). Ranges of those WQ elements were associated with three (3) main trophic classes: oligotrophic, mesotrophic and eutrophic

(McCullough, 2012) including also transitional categories (e.g. ultra-oligotrophic, hypertrophic; Watanabe et al., 2020).

Based on this rationale, very clear lakes are classified as oligotrophic Case-2 rather than Case-1 (Gons et al., 2008) since they typically receive significant levels of terrigenous input (Gons and Auer, 2004), while dissolved organic carbon (DOC), in this type of lakes, has an exceptionally powerful influence on water clarity (Gunn et al., 2001). Hence, there is a need for further algorithm development, especially for oligotrophic water bodies, while, of principal value is the choice of the appropriate wavelengths.

On one hand, red/NIR bands are usually utilized for the assessment of Chl-*a* concentrations in Case-2 waters (O'Reilly and Werdell, 2019; Seegers et al. 2021); however, AC algorithms need to further improve to replicate the spectral shape in the NIR bands so that the NIR-red band ratio algorithms can be used in such turbid waters (Warren et al., 2019). On the other hand, escpecially for clear waters, the use of blue-green ratio has been reported as the most effective for the monitoring of WQ elements (Binding et al. 2019; O'Reilly and Werdell, 2019, Warren et al., 2019) since their turbidities are non-algal and inorganic (Warren et al., 2019)

In purview of the above, hereby-developed lake WQ quantitative models (Chl-*a*, Secchi depth and Total phosphorus), based on wide concentration ranges, were applied to Landsat 8 OLI images illustrating two (2) Greek oligotrophic lakes. Basic aims of this effort are to explore the efficiency of the aforementioned WQ models in monitoring Trichonis and Amvrakia lakes' trophic status and reach final conclusions concerning whether there is indeed a need for the development of special algorithms exclusively oriented to oligotrophic waterbodies.

4.1.2 Methodology

4.1.2.1 Study areas

Trichonis Lake (Figure 4.1.2-1) is the largest natural freshwater body in Greece and it receives pollutants from numerous anthropogenic activities, especially from intensive agricultural practices, urban sewages, stock grazing land and small industries. Even though large quantities of fertilizers are applied in the lake's catchment, the trophic status of the lake is oligotrophic to oligomesotrophic (Koussouris 1993; Zacharias et al., 2002; Bertahas et al., 2006). Trichonis Lake is a deep freshwater body which has a surface area of 97 km^2, a maximum depth of 58 m and a potential water volume of approximately $2.8x10^9$ m^3 (Figure 4.1.2-1; Dimitriou et al., 2001).

Lake Amvrakia (Figure 4.1.2-1) belongs to the European Ecological Network Natura 2000, has a surface area of 14 km^2 and a maximum water depth of 50 m. Lake Amvrakia is characterized by strong water level

fluctuations due to high evaporation rates, especially during the summer, and the irrigation of the surrounding agricultural area. These alterations usually lead to the drainage of the shallower northern part (Figure 4.1.2-1) of the basin in certain periods and, consequently, to the fluctuation of the surface area of the lake (Zotos et al., 2021).

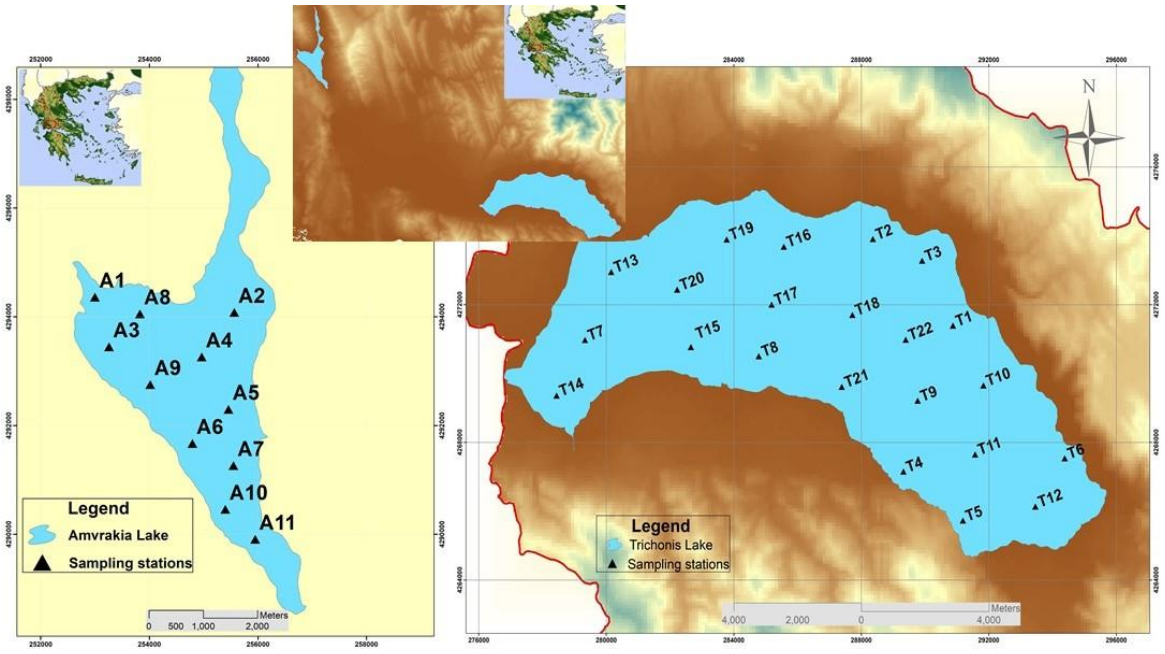


Figure 4.1.2-1. Sampling stations in Trichonis and Amvrakia lakes.

4.1.2.2 Water sampling, in-situ data and Chemical analyses

Water samplings were conducted by HCMR staff in the framework of a research project studying water quality of lakes and rivers located in the western part of central Greece. A total of twenty-two (22) and eleven (11) water samples were collected across the surfaces (5-10 cm) of lake Trichonis and Amvrakia, respectively while a GPS (Global Position System) was utilized to record the coordinate data of each station. Water was collected with NIO samplers of 1.5-l capacity in 29-30/10/2013 (Trichonis), 31/10/2013 (Amvrakia), 30/08/2014 (Trichonis) and 31/08/2014 (Amvrakia). Following collection, the water samples for nutrient analysis were preserved by the addition of $HgCl_2$ and on return to the HCMR laboratories were filtered and analyzed for total phosphorus concentrations. Samples were filtered through 0.45 μm cellulose acetate filters that had been precleaned with 10% hydrochloric acid (pH = 2) followed by rinsing with Milli-Q water.

A specific quantity of water samples for chlorophyll-*a* (usually 1 L) was filtered through Whatman GF/F filters immediately after collection. These filters were maintained in a dry and dark environment at −15 °C and then transferred to HCMR laboratories for further analysis.

In-situ data utilized in the framework of this study includes Chl-*a* and TP concentrations. However, since Secchi disk measurements have not been conducted in none of the studied lakes, an effort has been made to estimate Secchi depths from turbidity data based on the following regression model (Rasmussen et al., 2009; Equation 4.1.2-1):

$$Secchi\ Depth = 11.123 * TBDY^{-0.637} \quad (4.1.2\text{-}1)$$

where Secchi Depth is in feet and TBDY is turbidity in FNU. This equation includes a bias correction factor of 1.01 while more information can be found in http://pubs.usgs.gov/tm/tm3c4/. Furthermore, it should ne noted that turbidity measurements were conducted only in water sampling of 2013, thus Secchi depth values concern only this year. Turbidity was measured with the HACH 2100Q IS Portable Turbimeter.

Concentrations of TP were determined in the soluble fraction using the photometer Merck Nova 400. The Chl-*a* concentrations were determined with a TURNER 00-AU- 10U fluorometer according to the method of Holm-Hansen et al. (1965), modified by Welschmeyer (1994).

Trophic status classification

In order to classify the water quality of Trichonis and Amvrakia lakes, the EPA (Environmental Protection Agency) classification system was used (EPA, 2000). According to this scheme, the classification of lakes into seven quality classes (Table 4.1.2-1) is based on the total phosphorus concentration, water transparency and trophic index (Trophic State Index—TSI). Trophic index TSI is calculated for each classification quality parameter as follows (Carlson and Simpson, 1996):

$$TSI\ (SD) = 60 - 14.41 * LN(SD) \quad (4.1.2\text{-}2)$$

$$TSI\ (Chla) = 9.81 * LN(Chla) + 30.6 \quad (4.1.2\text{-}3)$$

$$TSI(TP) = 14.42 * LN(TP) + 4.15 \quad (4.1.2\text{-}4)$$

where SD is the Secchi disk (m) and Chl-*a* and TP (μg/l) are the concentrations of chlorophyll-*a* and total phosphorus, respectively. In the context of this study, there are no available data of Secchi disk therefore, this water quality

classification effort is based only on TP and Chl-*a* values, aiming to better understanding of the prevailing conditions during the sampling periods and afterwards to ascertain lakes' oligotrophic nature.

Table 4.1.2-1. Proposed lake WQ classification system by United States EPA (Carlson and Simpson, 1996).

TSI average	SD (m)	TP (μg/l)	Chl-*a* (μg/l)	Trophic status-Attributes
< 30	> 8	< 6	< 0.94	Oligotrophic-Clear water, oxygen throughout the year in the hypolimnion
30 - 40	8 - 4	6 - 12	0.94 – 2.6	Oligotrophic -A lake will still exhibit oligotrophy, but some shallower lakes will become anoxic during the summer
40 - 50	4 - 2	12 - 24	2.6 – 6.4	Mesotrophic-Water moderately clear, but increasing probability of anoxia during the summer
50 - 60	2 - 1	24 - 48	6.4 - 20	Eutrophic-Lower boundary of classical eutrophy: Decreased transparency, warm-water fisheries only
60 - 70	0.5 - 1	48 - 96	20 - 56	Eutrophic-Dominance of blue-green algae, algal scum probable, extensive macrophyte problems
> 70	< 0.25	> 96	> 56	Hypereutrophic, Heavy algal blooms possible throughout the summer, often hypereutrophic

4.1.2.3 Satellite data and pre-processing

Two Landsat 8 OLI images (Path 184, Row 33) illustrating Trichonis and Amvrakia lakes of 30 October 2013 and 30 August 2014 were used for this study. The satellite images were acquired from the USGS (United States Geological Survey) Data Centre (http://glovis.usgs.gov/) while image processing was completed in ENVI software (EXELIS Visual Information Solutions, Version 5.1).

Each band for both Landsat 8 OLI images was radiometrically and geometrically corrected (using GCP). After assessing geometric accuracy based on Global Position System measurements (coordinate data) taken in the study areas, the geometrical accuracy was determined to be less than one half pixel (<15 m). Finally, each band was converted to top-of-atmosphere (TOA) reflectance with sun angle correction using radiometric calibration coefficients provided in the metadata file to normalize the images for comparison

between different days. For atmospheric correction, dark object subtraction (DOS) technique was used, which takes the minimum value in each band and removes it from each pixel (Lathrop et al. 1992; Keiner and Yan 1998; Vincent et al., 2004).

4.1.2.4 Application of WQ models in Landsat 8 OLI images and performance evaluation

Hereby used WQ models concern quantitative Chl-*a* models developed by Markogianni et al. (2020) and Secchi depth and TP models developed by Markogianni et al. (2022). Application of WQ models includes Chl-*a* models (General- Equation 4.1.2-5; Natural- model-Equation 4.1.2-6), Secchi depth models (General-Equation 4.1.2-7; Natural model- Equation 4.1.2-8) and TP models (General-Equation 4.1.2-9; Natural model -Equation 4.1.2-10). $Secchi_{general}$ and $Secchi_{natural}$ models have been applied only in Landsat 8 OLI image of 2013 while the same applies for $TP_{natural}$ model (Equation 4.1.2-10) since it employs Secchi depth data.

$$\log Chla = 3.599 - 0.63 * \left(\frac{blue}{red}\right) - 2.183 * \left(\frac{\ln red}{\ln swir2}\right) \quad (4.1.2\text{-}5)$$

$$\log Chla = 4.443 - 1.421 * \left(\frac{blue}{green}\right) - 3.454 * \left(\frac{\ln red}{\ln swir2}\right) + 1.304 * \left(\frac{red}{green}\right) \quad (4.1.2\text{-}6)$$

$$SQRT(Secchi)_{general} = 1.215 - 2.479 * \left(blue + red + \frac{red}{blue}\right) + 3.394 * \left(\frac{lngreen}{lnswir2}\right) \quad (4.1.2\text{-}7)$$

$$SQRT(Secchi)_{natural} = 1.172 - (1.003 * logchl - a) - (1.031 * logred) \quad (4.1.2\text{-}8)$$

$$LogTPgeneral = -1.425 + 0.452 * logChla - 0.573 * \left(\frac{lnred}{lnswir1}\right) \quad (4.1.2\text{-}9)$$

$$LogTPnatural = -0.633 - (0.704 * logSecchi) - 0.392 * \left(\frac{green}{red}\right) \quad (4.1.2\text{-}10)$$

After the models' application, satellite-derived values of Chl-*a* and TP (general) concentrations of 2013 and 2014 and Secchi depth and TP (natural) of 2013, acquired from both Trichonis and Amvrakia lakes, were compared

with the corresponding *in-situ* values. WQ models' performance was based on the Spearman's (r) correlation coefficient and the error metrics Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE) and Normalized Root Mean Squared Error (NRMSE).

4.1.3 Results

4.1.3.1 Trophic status classification

Considering the concentrations of total phosphorus and Chl-*a* and the estimated average Trophic Index (TSI) of both the sampling campaigns, Amvrakia Lake is characterized as oligotrophic to oligomesotrophic for both years (Table 4.1.3-1) due to increased TP concentrations. Furthermore, according to Markogianni et al. (2018), Trichonis lake was also classified as oligotrophic to oligomesotrophic in 2013 and oligotrophic in 2014.

Table 4.1.3-1. EPA classification system and estimated TSI for Amvrakia Lake.

	avg Chl-*a* (μg/l)	TSI (Chl-*a*)	TP (μg/l)	TSI (TP)	TSI average	Classification
2013	0.77	27.99	21.00	48.05	38	Oligotrophic to oligomesotrophic
2014	0.43	22.4	32.45	54.33	38.4	Oligotrophic to oligomesotrophic

4.1.3.2 WQ models' application in Trichonis and Amvrakia lakes

Application of WQ models in Trichonis and Amvrakia lakes indicated their bad performance concerning all WQ elements. Even though the correlations between *in-situ* and satellite Chl-*a* (general and natural) data were statistically significant (Table 4.1.3-2), the values of error metrics NRMSE and median MAPE are particularly high (indicating inaccurate forecasting). Median MAPE was selected as it is considered more resilient to outliers than MAPE. Relatively better performance was presented by TP models based on RMSE, NRMSE and MAPE metrics; however, no correlation was detected among satellite and *in-situ* data. Secchi models' performance was also poor but, in this point, it should be noted that *in-situ* Secchi values have emerged via the transformation of turbidity values; factor that has surely affected their effectiveness and contributed to statistical insignificant correlations.

Table 4.1.3-2. Error metrics' values after the WQ models' application to Landsat 8 OLI images (** Correlation is significant at the 0.01 level (2-tailed).

WQ model	Units	avg *in-situ*	avg satellite	ME	RMSE	NRMSE	median MAPE	Spearman r
Chl-$a_{general}$	μg/l	0.66	2.25	-1.6	1.9	1.6	256.8	0.513**
Chl-$a_{natural}$	μg/l	0.66	1.84	-1.19	1.8	1.5	176.6	0.44**
$TP_{general}$	mg/l	0.031	0.02	0.012	0.022	0.32	54.7	-0.16
$TP_{natural}$ (2013)	mg/l	0.035	0.033	0.002	0.024	0.34	37.9	-0.17
$Secchi_{general}$ (2013)	m	3.31	6.43	-3.12	3.35	0.79	93.6	0.099
$Secchi_{natural}$ (2013)	m	3.31	8.26	-4.95	5.1	1.2	143.4	0.08

4.1.4 Discussion

Landsat 7 ETM+/8 OLI-developed WQ models have been established on the basis of *in-situ* data, sampled from 50 different lake water bodies across Greece. Chl-*a*, Secchi and TP quantitative models have been built based on a wide range of concentrations (Markogianni et al., 2020; 2022) representing almost all trophic status classes. In the framework of this work, an effort has been made to apply those WQ models and explore their performance in a distinct category of Case-2 waters, e.g. oligotrophic lakes. Trichonis and Amvrakia lakes are classified as oligotrophic based on both the recorded literature and the hereby trophic status classification, which relied on two (2) field trips in 2013 and 2014. Furthermore, trophic status of both lakes has been assessed by utilizing *in-situ* data from numerous sampling stations across each studied lake.

Application of WQ models in oligotrophic Trichonis and Amvrakia lakes was ineffective while particularly low and homogeneous measured Chl-*a* concentrations indicated lakes where the greatest optical contribution originates from non-algae particles. Considering the relevant literature there are a plethora of studies with similar to hereby results. Seegers et al. (2021) evaluated the Cyanobacteria Index (CI)-based Chl-*a* algorithm (Chl_{T16}) by using MERIS radiometric time series (2002-2012) for over 2300 waterbodies (United States) and more than 5000 in Alaska while they tried to derive a new CI-to-Chl-*a* relationship (Chl_{BS}). According to their results, the Chl_{BS} algorithm performed best in the >7 μg/l range, while underachieved at the lowest chlorophyll concentrations (oligotrophic-mesotrophic). Seegers et al. (2021) attributed the bad performance mainly to the need of $\varrho s(\lambda)$ signal in the NIR to overcome the absorption of pure water in that range of low concentrations and afterwards to the fact that satellite instrument performance in the NIR

may confound meaningful retrievals. Moreover, Gilerson et al. (2010) reported that *Chl* algorithms using the red/NIR surpassed blue-green band ratio algorithms at concentrations higher than 5 μg/l while Binding et al. (2019) also found that the MCI (Maximum Chlorophyll Index) and CI performed better than band-ratio approaches for *Chl-a* concentrations higher than 10 μg/l.

Another study was conducted by Gons and Auer (2004) who attempted to use spectral reflectance R (0, λ) for Chl-*a* retrieval in the Keweenaw Bay (Lake Superior). Measured spectra were typical of oligotrophic lacustrine waters while strong absorption by water was observed in the red region which hindered the accurate detection of Chl-*a* absorption. Ultimately, they also indicated the need of algorithm development for oligotrophic waterbodies.

Considering the utilized bands for Chl-*a* retrieval in oligotrophic lakes, Gons et al. (2008) implied that Chl-*a* mapping in oligotrophic areas of the Great Lakes (north America) remains problematic for the current generation of satellite sensors, in particular MERIS and MODIS. More specifically, they proposed the existence of more and narrower bands in the red-NIR spectral region in the case where the adequate performance of their empirically developed algorithm- employing blue-to-green bands- was coincidental. O'Reilly and Werdell (2019) also suggested an approach using a blue-green band ratio algorithm in oligotrophic systems and a red/NIR method in lakes with concentrations ranging 3–155 μg/L.

Considering the contribution of machine learning methods in trophic status classification of diverse water bodies, Watanabe et al. (2020) tested Artificial Neural Network (ANN), Random Forest (RF) and Support Vector Machine (SVM) algorithms in four (4) reservoirs in Brazil based on *in-situ* reflectance measurements while all of them exhibited the poorest modelling for oligotrophic samples.

As far as the calculation of Secchi depths in Trichonis and Amvrakia lakes is concerned, no safe conclusions can be drawn in the context of this study, due to absence of available data. However, according to literature, the Secchi depth monitoring seems also problematic in lakes with fewer particles dissolved in water. More particularly, established relationships between trophic status and Secchi depths of Maine lakes indicated more accurate estimates -in relation to observed conditions- for eutrophic and mesotrophic (on average within 1 m) than for oligotrophic (on average deviated higher than 1 m) lakes (Maine Pearl, 2011). According to Lathrop (1992), increased turbidity and phytoplankton connected with higher chlorophyll-*a* concentrations, result in escalating energy received by the satellite; thus, red band is a less accurate predictor of Secchi depth in clear waters (McCullough, 2012). McCullough (2012) also highlights that the longer red band may reach the bottom before the deepest SDD is attained, yielding ambiguous results.

Therefore, based on the hereby utilized $Secchi_{natural}$ model which employs Chl-*a* concentrations and the red band, inaccurate results are anticipated concerning Secchi depths in Trichonis and Amvrakia lakes.

Concerning TP estimation in hereby studied lakes, results were proven better compared to Chl-*a* and Secchi depth except for the absence of a statistically significant correlation between predicted and observed values. Despite the oligotrophic character of Trichonis and Amvrakia lakes, measured TP concentrations for both years characterize them as marginally eutrophic (EPA, 2000). Regardless of the high measured concentrations which theoretically would contribute to a better prediction, TP cannot be assessed remotely because is characterized by weak optical characteristics and a low signal noise ratio (Markogianni et al., 2022). However, TP is highly correlated with optically active constituents while the hereby applied TP models also employ Chl-*a* concentrations (general model) and Secchi depths (natural model). The involvement initially of Chl-*a* values (which are exceptionally low) and afterwards of Secchi depths which originated from the Turbidity transformations hindered the achievement of a higher accuracy due to reasons extensively described above.

Poor performance of WQ models could further be attributed to other sources of inaccuracy such as the employment of multiple laboratory techniques and the lack of knowledge about regional phytoplankton community composition. Some algorithms were proven to present sensitivity to community composition (diatom- or cyanobacteria dominated sampling stations; Binding et al., 2019); hence more research is required involving known Chl-*a* distributions, particles and CDOM before the establishment of special algorithms, exclusively oriented to oligotrophic lakes.

4.1.5 References

Bertahas, I.; Dimitriou, E.; Karaouzas, I.; Laschou, S.; Zacharias, I. Climate change and agricultural pollution effects on the trophic status of a Mediterranean lake. Acta hydrochim. hydrobiol. **2006**, vol.34, pp. 349–359.

Binding, C.E.; Zastepa, A.; Zeng, C. The impact of phytoplankton community composition on optical properties and satellite observations of the 2017 western Lake Erie algal bloom. J. Great Lakes Res. **2019**, 45, 573–586. https://doi.org/10.1016/j.jglr.2018.11.015.

Carlson, R. E.; Simpson, J. A coordinator's guide to volunteer lake monitoring methods. North American Lake Management Society **1996**, 96.

Carlson, R.E. A trophic state index for lakes. *Limnol. Oceanogr.* **1977**, 22, 361–369.

Dimitriou, E.; Zacharias, I.; Koussouris, Th. Water resources management plan for Trichonis Lake catchment. Final report. In: Zacharias, I., Koussouris, Th. (Eds.), Actions for the Protection of Calcareous Bogs/Fens in Trichonis Lake. Technical Report, NCMR/IIW, pp.70, **2001**.

EPA - US Environmental Protection Agency, *Nutrient Criteria, Technical Guidance Manual, Lakes and Reservoirs*, First Edition, EPA-822-B00-001, 2000.

Gholizadeh, M.; Melesse, A.; Reddi, L. A comprehensive review on water quality parameters estimation using remote sensingtechniques. Sensors **2016**, 16, 1298.

Gilerson, A.A.; Gitelson, A.A.; Zhou, J.; Gurlin, D.; Moses, W.; Ioannou, I.; Ahmed, S.A. Algorithms for remote estimation of chlorophyll-a in coastal and inland waters using red and near infrared bands. Opt. Express **2010**, 18 (23), 24109–24125.

Gons, J. Herman; Auer, T. Martin. Some Notes on Water Color in Keweenaw Bay (Lake Superior), Journal of Great Lakes Research **2004**, Volume 30, Supplement 1, Pages 481-489, ISSN 0380-1330, https://doi.org/10.1016/S0380-1330(04)70408-6.

Gons, J. Herman; Auer, T. Martin; Effler, W. Steven. MERIS satellite chlorophyll mapping of oligotrophic and eutrophic waters in the Laurentian Great Lakes, Remote Sensing of Environment **2008**, Volume 112, Issue 11, Pages 4098-4106, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2007.06.029.

Gordon, H.R.; Morel, A. Remote Assessment of Ocean Color for Interpretation of Satellite Visible Imagery. A Review, Lecture Notes on Coastal and Estuarine Studies; Barber, R.T., Mooers, N.K., Bowman, M.J., Zeitzschel, B., Eds.; Springer: New York, NY, USA, **1983**; p. 114.

Gunn, J. M.; Snucins, E.; Yan, N. D.; Arts, M. T. Use of water clarity to monitor the effects of climate change and other stressors on oligotrophic lakes. Environmental Monitoring and Assessment **2001**, 67, 69–88.

Holm-Hansen, O.; Lorenzen, C. J.; Hormes, R. W.; Strickland, J. D. H. Fluorometric determination of chlorophyll. *Journal du Conseil/Conseil Permanent International pour l'Exploration de la Mer* **1965**, vol.30, pp. 3–15, 1965, doi:10.1093/icesjms/30.1.3.

IOCCG, **2000**. Remote Sensing of Ocean Colour in Coastal, and Other Optically-Complex, Waters; Reports of the International Ocean-Colour Coordinating Group, No. 3; Sathyendranath, S., Ed.; IOCCG: Dartmouth, NH, Canada.

IOCCG, **2018**. Earth observations in support of global water quality monitoring. In: Greb, S., Dekker, A., Binding, C. (Eds.), IOCCG Report Series. No. 17, International Ocean Colour Coordinating Group, Dartmouth, Canada.

Keiner, L.E.; Yan, X. A neural network model for estimating sea surface chlorophyll and sediments form Thematic Mapper imagery. Remote Sens. Environ. **1998**, 66, 153–165.

Koussouris, Th. Contribution of the impact study of agricultural and cattle-raising activities to the trophic status of Trichonis Lake and the possibilities of fishery development, PhD thesis, University of Thessaloniki, **1993**.

Lathrop, R. Landsat Thematic Mapper monitoring of turbid inland water quality. *Photogramm. Eng. & Rem. S.* **1992**., 58(4), 465–470.

Maine PEARL, **2011**. Lakes Guide. University of Maine, Orono: Senator George J. Mitchell Center for Environmental Research http://www.pearl.maine.edu/windows/community/default.htm.

Markogianni, V.; Kalivas, D.; Petropoulos, G.; Dimitriou, E. An appraisal of the potential of Landsat 8 in estimating chlorophyll-a, ammonium concentrations and other water quality indicators. Remote Sens. **2018**, 10, 1018.

Markogianni, V.; Kalivas, D.; Petropoulos, G.P.; Dimitriou, E. Estimating Chlorophyll-*a* of Inland Water Bodies in Greece Based on Landsat Data. *Remote Sens.* **2020**, *12*, 2087. https://doi.org/10.3390/rs12132087.

Markogianni, V.; Kalivas, D.; Petropoulos, G.P.; Dimitriou, E. Modelling of Greek Lakes Water Quality Using Earth Observation in the Framework of the Water Framework Directive (WFD). *Remote Sens.* **2022**, 14, 739. https://doi.org/10.3390/rs14030739.

McCullough, Ian M. Remote Estimation of Regional Lake Clarity with Landsat TM and MODIS Satellite Imagery. *Electronic Theses and Dissertations*. **2012**, 1744.

O'Reilly, J.E.; Werdell, P.J. Chlorophyll algorithms for ocean color sensors-OC4, OC5 & OC6. Remote Sens. Environ. **2019**, 229, 32–47.

Rasmussen, P.P.; Gray, J.R.; Glysson, G.D.; Ziegler, A.C., **2009**. Guidelines and procedures for computing time-series suspended-sediment concentrations and loads from in-stream turbidity-sensor and streamflow data: U.S. Geological Survey Techniques and Methods book 3, chap. C4, 53 p.

Seegers, N. Bridget; Werdell, P. Jeremy; Vandermeulen, A. Ryan, Salls, Wilson, Stumpf, P. Richard; Schaeffer, A. Blake; Owens, J. Tommy; Bailey, W. Sean; Scott P. Joel; Loftin A. Keith. Satellites for long-term monitoring of inland U.S. lakes: The MERIS time series and application for chlorophyll-a, *Remote Sensing of Environment* **2021**, Volume 266, 112685, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2021.112685.

Vincent, R.K.; Qin, X.; McKay, R.M.; Miner, J.; Czajkowski, K.; Savino J. et al. Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie. *Remote Sensing of Environment* **2004**, vol. 89, no. 3, pp. 381–392.

Warren, M.A.; Simis, S.G.H.; Martinez-Vicente, V.; Poser, K.; Bresciani, M.; Alikas, K.; Spyrakos, E.; Giardino, C.; Ansper, A. Assessment of atmospheric correction algorithms for the Sentinel-2A MultiSpectral Imager over coastal and inland waters, *Remote Sensing of Environment* **2019**, 225, 267-289.

Watanabe, Fernanda S.Y.; Miyoshi, Gabriela T.; Rodrigues, Thanan W.P.; Bernardo, Nariane M.R.; Rotta, Luiz H.S.; Alcântara, Enner; Imai, Nilton N. Inland water's trophic status classification based on machine learning and remote sensing data, *Remote Sensing Applications: Society and Environment* **2020**, Volume 19, 100326, ISSN 2352-9385, https://doi.org/10.1016/j.rsase.2020.100326.

Welschmeyer, N. A. Fluorometric analysis of chlorophyll a in the presence of chlorophyll b and pheopigments. *Limnol. Oceanogr.* **1994**, vol. 39, 1994, http://dx.doi.org/ 10.4319/lo.1994.39.8.1985.

Zacharias, I.; Bertachas, I.; Skoulikidis, N.; Koussouris, T. Greek Lakes: Limnological overview. *Lakes & Reservoirs, Research & Management* **2002**, vol.7, no.1, pp. 55–62.

Zotos, Anastasios; Kosma, Chariklia; Triantafyllidis, Vassilios; Kakabouki, Ioanna; Kehayias, George; Roussis, Ioannis; Mavroeidis, Antonios; Tataridas, Alexandros; Bilalis, Dimitrios. Plant Species Diversity of the Wet Meadows under Natural and Anthropogenic Interventions: The Case of the Lakes Amvrakia and Ozeros (W. Greece)". *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* ***2021***, 49 (3):12435. https://doi.org/10.15835/nbha49312435.

4.2 WQ Modelling of Oligotrophic Trichonis Lake

Based on:

V. Markogianni, D. Kalivas, G. Petropoulos, E. Dimitriou. Analysis on the Feasibility of Landsat 8 Imagery for Water Quality parameters assessment in an Oligotrophic Mediterranean Lake. World Academy of Science, Engineering and Technology, *International Journal of Geological and Environmental Engineering* **2017**, Vol: 11, No: 9, 906-914**.**

and

Markogianni, V.; Kalivas, D.; Petropoulos, G.P.; Dimitriou, E. An Appraisal of the Potential of Landsat 8 in Estimating Chlorophyll-*a*, Ammonium Concentrations and Other Water Quality Indicators. *Remote Sens.* **2018**, *10*, 1018. https://doi.org/10.3390/rs10071018

4.2.1 Introduction

Water resources are essential for the survival of all living organisms. Part of those resources is stored in lakes and reservoirs, and is used to satisfy environmental and human requirements. Unfortunately, in many cases WQ is chemically deteriorated, and water managers/ scientists need new means for efficient monitoring (Seegers et al., 2021).

The continuous monitoring of large water bodies is a complex task, since it demands frequent and detailed data collection and interpretation efforts. Only frequent fieldworks can fully attain the spatial and temporal variance of WQ key indicators. This requires a compromise concerning the number of sampling stations while keeping maintainance costs reasonable (Lyu et al., 2022).

Satellite remote sensing (RS) is a powerful supportive tool for assessing of spatial and temporal variations in WQ (Lyu et al., 2022). RS technologies enable researchers to acquire a unique, holistic perspective of the ecosystems. From the vantage point of space, satellite data become an invaluable tool in support of wetland management. This is of especial importance in the context of the increasingly strict environmental regulations approved by governments worldwide (e.g. Water Framework Directive and the European Marine Strategy Framework Directive) (Pizani and Maillard, 2022).

Since the European Commission Water Framework Directive (EC, 2000) was promulgated, Member States have started to develop lake ecological status assessment systems, and finished setting TP and Chl-*a* as reference conditions for European lakes in different lake types and ecoregions (Nikolaidis et al., 2022). In particular, the use of multi-spectral sensors makes possible to measure many of the parameters required by law (Gholizadeh et al., 2016). Apart from the law-required components, the major factors which can influence the quality of inland water bodies are the suspended sediments (turbidity), phytoplankton and cyanobacteria (i.e., chlorophylls, carotenoids), dissolved organic matter (DOM), organic and inorganic nutrients, pesticides, metals, thermal releases, macrophytic algae, pathogens and oils. The above-mentioned factors affect the optical properties of waters (except for nutrients) thus they directly change the signal acquired by optical sensors over water bodies (Gholizadeh et al., 2016, Pizani and Maillard, 2022). The parameters which can be directly quantified using RS techniques are the suspended particulate matter (SPM), which is placed in suspension by wind-wave stirring of shallow waters and can be a tracer for inflowing pollutants (Eleveld, 2012), the phytoplankton mainly as Chlorophyll-*a* (chl-*a*) or phycocyanin (PC), that can be used to indicate the trophic level, to evaluate the presence of potentially toxic algal blooms and as a proxy of phytoplankton biomass (Randolph et al., 2008; Ruiz-Verdu et al., 2008), the coloured DOM (CDOM), which is investigated because of its role in

protecting aquatic biota from ultraviolet solar radiation and its influence on specifically heterotrophic bacterial productivity in the water column, indicative of the shift from net autotrophy to net heterotrophy (Kutser et al., 2005; Giardino et al., 2014).

A number of satellite sensors have been used for the study of surface WQ (Kutser et al., 2009; Yacobi et al., 2011; Matthews, 2011; Odermatt et al., 2012a). Matthews (2011) and Kutser (2012) have provided a detailed review of RS instruments which can be used to assess WQ in inland and near-coastal waters. Medium spatial resolution multi-spectral sensor such as Advanced Land Imager (ALI) (30 m), Advanced Land Observation Satellite (ALOS) (10 m), SPOT-5 (10 m) and Landsat provide images in the visible and near-infrared wavelengths; compared to the higher spatial resolution sensors, these sensors are characterized by a higher radiometric performance which contributes to a more accurate assessment of the concentrations of quality parameters over water. On May 30, 2013, data from the Landsat-8 satellite (launched on 11 February, 2013) became available allowing the continuance of studies on WQ of lakes (Giardino et al., 2014).

Although Landsat sensors were not designed for aquatic applications (Kutser, 2012; McCullough et al., 2012a), we find numerous examples of applications of Landsat images for estimating and/or monitoring lake WQ. Several studies have proposed reliable algorithms between Landsat data and WQ parameters, including chlorophyll; phytoplankton and PC concentrations (Brezonik et al., 2005; Karakaya et al., 2011; Tebbs et al., 2013), water clarity (Hadjimitsis et al., 2006; Guan et al., 2011; Zhao et al., 2011), CDOM (Brezonik et al., 2005; Zhu et al., 2014; Brezonik et al., 2015), blooms of cyanobacteria (Vincent et al., 2004), macrophyte (Albright and Ode, 2011) and total suspended sediments (TSS; Zhou et al., 2006; Guang et al., 2006; Onderka and Pekarova, 2008; Bonansea et al., 2013). Few studies, though, have attempted to monitor and model nutrient data, since those data do not have optical properties and the regression models usually yield statistically insignificant results (Gholizadeh et al., 2016). In particular, Chen and Quan (2012) used Landsat TM imagery to predict nitrogen and phosphorus concentrations in Tiahu Lake, China with some successful results for phosphorus and less successful results for nitrogen. In general, the aforementioned studies considerably increase knowledge of WQ and most of their developed algorithms are commonly based on empirical relationships using classical simple linear regression models between remotely sensed reflectance values and measurements collected simultaneously in the field.

In contrast to the clear oceanic waters (Case-1 waters), retrieval problems of some WQ parameters have arisen for coastal and inland waters (Case-2 waters) (Gons et al., 2008). Monitoring of WQ parameters in Case 2 waters is not an easy task due to runoff and discharges from rivers/streams, which add to the complexity of the water constituent retrieval process.

Inflows from streams introduce different organic/ inorganic particles (e.g. TSS). As opposed to particles, Chl-*a* and particularly CDOM are absorbing components of water with CDOM absorbing the highest in short wavelengths (350–440 nm) and Chl-*a* representing two absorption peaks in the blue and the red regions of the spectrum (Pahlevan and Schott, 2013). Whereas Chl-*a* in Case-1 waters can be accurately estimated on the basis of the pigment's absorption peak in the blue, in oligotrophic Case-2 waters, estimation on the basis of the Chl-*a* absorption peak in the red can be no alternative due to the overwhelming absorption by water of the red and near-infrared (NIR) wavelength bands (Gons et al., 2008).

Moreover, findings from numerous published studies have indicated biological and chemical WQ parameters such as Chl-*a* can be measured using spectral indices. However, these indices appear to be less reliable in diverse water bodies e.g Case 2 waters (lakes, ponds, rivers and streams in coastal regions) (Yang et al., 2017). A variety of spectral indices derived from RS data based on empirical or semi-empirical relationships have been developed for transforming spectral data into WQ parameters. Water indices' usefulness has been demonstrated in different studies for drought monitoring and early warning assessment (Memon et al., 2015; Bohn et al., 2017). Nevertheless, the vegetation indices and reflectance (individual bands and band ratios) values' application is highly encouraged for the estimation of WQ parameters (i.e. chlorophyll-*a*, transparency) in lakes (Bonansea et al., 2015; Doña et al., 2015; Bohn et al., 2017). These indices may involve three (Yang et al., 2010; Song et al., 2013; Sun et al., 2014; Huang et al., 2014) and four spectral bands (Le et al., 2009). The majority, though, of spectral indices are based on reflectance ratios of two spectral bands (near infrared and red) for operational purpose. A band ratio between the near infrared (NIR, ~0.7 μm) and Red (~0.6 μm) has frequently been used to estimate Chl-*a* in waters due to a positive reflectivity of Chl-*a* in the NIR and an inverse behavior in the red while NIR and red bands are involved in most indices (Yang et al., 2017).

As well, lake water clarity can be estimated more accurately in eutrophic and mesotrophic than oligotrophic lakes, due to the absence of suspended particles in oligotrophic lakes that are evident by satellite sensors (McCullough, 2012). In oligotrophic lakes, water clarity is primarily controlled by the concentration of coloured dissolved organic matter (CDOM) (Gunn et al., 2001; Giardino et al., 2014), which, in turn, affects a wide range of chemical, physical and biological processes. These include thermal structure, light transmission for photosynthesis, attenuation of damaging levels of ultraviolet light, vertical distribution of plants and animals, as well as the form and availability of toxic metals (Gunn et al., 2001).

This study presents the analysis of L8 OLI imagery in combination with simultaneous field data to conduct basic spatial assessment of various WQ

parameters in a natural lake, characterized by particularly low concentration values and the absence of strong spatial and temporal variability. The main objective is to develop quantification algorithms and determine Chl-*a* concentration, CDOM absorption at 420 and 440 nm ($a_{cdom}(420)$; $a_{cdom}(440)$) and nutrient concentrations in the deep oligotrophic Lake Trichonida (Greece), using MLR analysis. Selected optimal algorithms were applied to another L8 image of different date but with available *in-situ* Chl-*a*, nutrient and CDOM absorption data, to validate the results while satellite derived values were compared to *in-situ* ones.

4.2.2 Methodology

4.2.2.1 Study area

Trichonis Lake (Figure 4.2.2-1) is the largest natural freshwater body in Greece while a more detailed description can be found at 4.1.2 Chapter (Study areas). A significant hydrogeologic aspect of Trichonis lake's catchment is that groundwater inflows to the lake during the dry periods are considerably high, which enhances the water abstraction potential for anthropogenic activities (Zacharias et al., 2003). Trichonis Lake's catchment is a 399 km^2 semi-mountainous area in Western Greece (Figure 4.2.2-1). The regional climate is characterized as semi-arid to arid Mediterranean with an average annual rainfall of 936 mm and an average annual temperature of 17 ^{0}C which fluctuates by 19 ^{0}C annually (Zacharias et al., 2005).

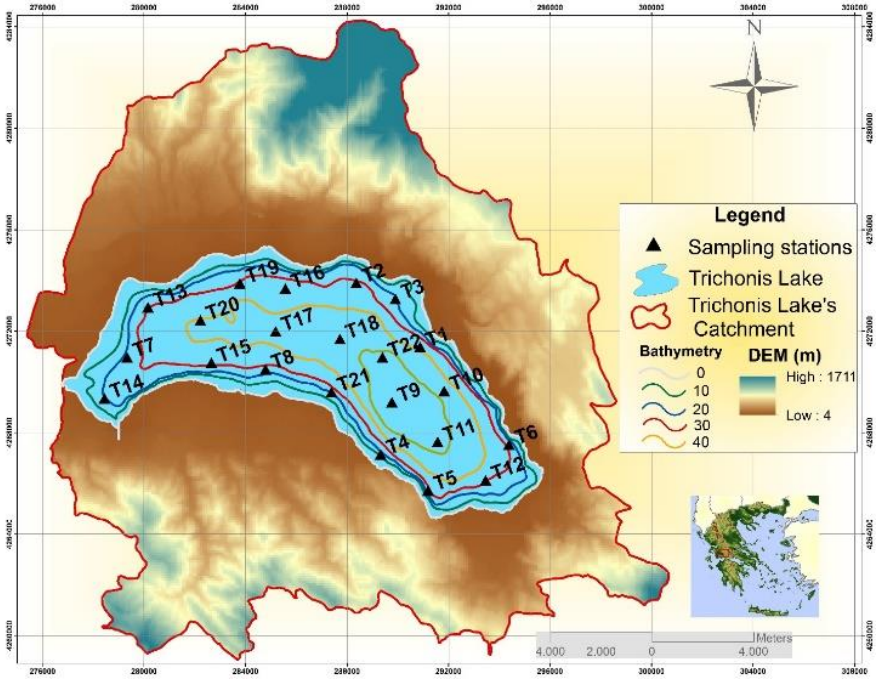


Figure 4.2.2-1. Trichonis Lake's catchment and bathymetry and Chl-*a, CDOM* and nutrients' sampling stations of 30-31/10/2013 and 30/08/2014.

4.2.2.2 Water sampling

Regarding the date difference between satellite image and fieldwork, water samplings were conducted at the same date as the satellite overpass. A total of 22 water samples were collected across the lake Trichonis's surface (5-10 cm) with NIO samplers of 1.5 l capacity in 29-30/10/2013 and 30/08/2014. Following collection, the preservation of the water samples for nutrient and Chl-*a* analysis was conducted as previously described in 4.1.2 Chapter (Water sampling, *in-situ* data and trophic status classification subchapter).

Water samples for CDOM absorption were filtered through 0.22 μm polycarbonate filter immediately after sampling. Filtered water was transferred into acid-cleaned (HCL 10%, 12 h) glass bottles and stored in the dark at ~-20 °C. Before measurement, the samples were allowed to stand until reaching room temperature.

4.2.2.3 Chemical Analyses and EPA quality classification system

Concentrations of nutrients (NO_3^-, NO_2^-, NH_4^+, and PO_4^{3-}) were determined in the soluble fraction using an ion analyser Metrohm, the automatic analyzer Radiometer and the photometer Merck Nova 400. The Chl-*a* concentrations were determined with a TURNER 00-AU- 10U fluorometer according to the method of Holm-Hansen et al. (1965), modified by Welschmeyer (1994). CDOM absorption spectra were obtained between 250 and 700 nm at 1 nm increments using a dual beam UV-visible spectrophotometer (Perkin Elmer, Lambda 25) equipped with 5 cm quartz cells and referenced to Milli-Q water. A baseline correction was applied by subtracting the average sample absorbance between 690 and 700 nm from the entire spectrum. In addition, a blank scan containing Milli-Q water was subtracted from each spectrum. Absorption units were converted to absorption coefficients using the relationship (Eq. 4.2.2-1):

$$a(\lambda) = 2.303 * A(\lambda)/l \qquad (4.2.2\text{-}1)$$

where $\alpha(\lambda)$ = absorption coefficient (m^{-1}), $A(\lambda)$ = absorbance, l = cell's light pathlength (m).

The EPA classification system was used for the WQ classification of Trichonis lake (EPA, 2000). According to this scheme, total phosphorus (TP) concentration, water transparency and trophic index (Trophic State Index—TSI) determine the classification of lakes into six quality classes (Table 4.2.2-1). Trophic index TSI is calculated for each quality parameter as follows (Carlson and Simpson, 1996):

$$TSI\ (SD) = 60 - 14.41 * LN(SD) \qquad (4.2.2\text{-}2)$$

$$TSI\ (Chla) = 9.81 * LN(Chla) + 30.6 \quad (4.2.2\text{-}3)$$

$$TSI(TP) = 14.42 * LN(TP) + 4.15 \quad (4.2.2\text{-}4)$$

where SD is the Secchi disk (m) and Chl-*a* and TP (μg/l) are the concentrations of Chl-*a* and TP, respectively.

Table 4.2.2-1. Proposed lake WQ classification system by United States EPA (Environmental Protection Agency) (Carlson and Simpson, 1996).

TSI average	SD (m)	TP (μg/l)	Chl-*a* (μg/l)	Trophic status-Attributes
< 30	> 8	< 6	< 0.94	Oligotrophic-Clear water, oxygen throughout the year in the hypolimnion
30 - 40	8 - 4	6 - 12	0.94 – 2.6	Oligotrophic -A lake will still exhibit oligotrophy, but some shallower lakes will become anoxic during the summer
40 - 50	4 - 2	12 - 24	2.6 – 6.4	Mesotrophic-Water moderately clear, but increasing probability of anoxia during the summer
50 - 60	2 - 1	24 - 48	6.4 - 20	Eutrophic-Lower boundary of classical eutrophy: Decreased transparency, warm-water fisheries only
60 - 70	0.5 - 1	48 - 96	20 - 56	Eutrophic-Dominance of blue-green algae, algal scum probable, extensive macrophyte problems
> 70	< 0.25	> 96	> 56	Hypereutrophic, Heavy algal blooms possible throughout the summer, often hypereutrophic

4.2.2.4 Satellite Data and Pre-Processing

L8 Operational Land Imager (OLI) images consist of 9 spectral bands with a medium spatial resolution (30 meters) for Bands 1 to 7 and 9. The ultra-blue Band 1 is advantageous for coastal and aerosol research. Furthermore, Band 9 is expendient for cirrus cloud observation. The resolution for Band 8 (panchromatic) is 15 meters (Table 4.2.2-2) (Barsi et al., 2014). Two L8 OLI images of Lake Trichonis (Path 184, Row 33) of 30 October 2013 (17:22:09Z) and 30 August 2014 (14:50:07Z) were used for this study. According to the large size of the Trichonis Lake, the number of sampling stations (22) were considered to be adequate for monitoring variability of CDOM, Chl-*a* and nutrient concentrations. The satellite images were acquired from the USGS (United States Geological Survey) Data Centre (http://glovis.usgs.gov/). The

image processing was completed in ENVI software (EXELIS Visual Information Solutions, Version 5.1) while further data elaboration and analysis were conducted in ESRI's software (ArcGIS v. 10.1).

The geometric accuracy for the two images was determined to be less than one half pixel (<15 m) based on Global Position System measurements (coordinate data) taken in the study area. Finally, each band was converted to top-of-atmosphere (TOA) reflectance with sun angle correction using radiometric calibration coefficients provided in the metadata file to normalize the images and facilitate the comparison between different days. For atmospheric correction, dark object subtraction (DOS) technique was used. The basic principle of this method is that within the image there are some pixels completely shadowed and their radiances received at the satellite originate entirely from atmospheric scattering (path radiance). This radiance value is then being subtracted from each pixel value in the image. The largest sources of errors for water constituents' retrieval is usually attributed to the bio-optical model that relates water leaving radiance (or reflectance) to the constituents' concentrations and to treatment of aerosol reflectance in the atmospheric correction procedure (Ruddick et al., 2000).

Furthermore, based on the demonstrated water indices' usefulness for the estimation of WQ parameters (i.e. chlorophyll-a, transparency) in lakes, several spectral (vegetation and water) indices were calculated (Table 4.2.2-3) to assess Chl-*a* concentrations.

Table 4.2.2-2. Landsat 8 spectral bands, wavelengths and spatial resolution.

Bands	Wavelength (micrometers)	Resolution (meters)
Band 1 - Ultra Blue (coastal/aerosol)	0.435 - 0.451	30
Band 2 - Blue	0.452 - 0.512	30
Band 3 - Green	0.533 - 0.590	30
Band 4 - Red	0.636 - 0.673	30
Band 5 - NIR	0.851 - 0.879	30
Band 6 - Shortwave Infrared (SWIR) 1	1.566 - 1.651	30
Band 7 - Shortwave Infrared (SWIR) 2	2.107 - 2.294	30
Band 8 - Panchromatic	0.503 - 0.676	15
Band 9 - Cirrus	1.363 - 1.384	30
Band 10 - Thermal Infrared (TIRS) 1	10.60 - 11.19	100 * (30)
Band 11- Thermal Infrared (TIRS) 2	11.50 - 12.51	100 * (30)

Table 4.2.2-3. Selected spectral indices calculated, according to literature.

INDEX	EQUATION	Source
Enhanced Vegetation Index (EVI)	EVI = *G * ((nir - red)/(nir + C1 * red - C2 * blue + L_evi))*	Liu and Huete, 1995
Normalised Ratio Vegetation Index (NRVI)	NRVI= *(red/nir - 1)/(red/nir + 1)*	Baret and Guyot, 1991
Normalised Difference Water Index (NDWI)	NDWI= *(green - nir)/(green + nir)*	McFeeters, 1996
Normalised Difference Water Index (NDWI2)	NDWI2= *(nir - swir2)/(nir + swir2)*	Gao, 1996
Modified Normalised Difference Water Index (MNDWI)	MNDWI= *(green - swir2)/(green + swir2)*	Xu, 2006
Green Normalised Difference Vegetation Index (GNDVI)	GNDVI= *(nir - green)/(nir + green)*	Gitelson et al., 1996
Normalised Difference Vegetation Index (NDVI)	NDVI= *(nir - red)/(nir + red)*	Rouse et al., 1974

4.2.2.5 Development of Models Relating L8 and WQ Data

MLR analysis was used in this study to develop relationships between remotely sensed reflectance data (independent) and Chl-*a*, log(Chl-*a*), spectral indices, CDOM and nutrient values (dependant). Initially, attempts were made to find combinations, transformations, or logarithmic transformations of L8 OLI bands which would provide more information about the under-study parameters in the lake than only one band. Subsequently, pixel values of each transformed image were retrieved from those regions where the 22 sampling stations are located. The transformed variables were denoted as log(Chl-*a*), Chl-*a*, ln($a_{CDOM(420)}$) and ln(acdom440). In addition to the above, the calculated vegetation and water indices (Table 4.2.2-3) were added to the analysis.

The first criterion considered in order to select the best quantitative model was the predictor importance chart conducted in IBM SPSS software Statistics Base v. 23.0. The predictor importance chart contributes to indicating the relative importance of each predictor in estimating the model; it does not relate to model accuracy but to the importance of each predictor in making a prediction. Subsequently, after having selected the predictors with the highest importance for each WQ parameter, they were further imported in a series of stepwise and backward linear regressions. Criteria of multicollinearity and values of tolerance factor, variance inflation factor (VIF) and condition indices (CI) were applied to a subset of strategic models to further help compare them. Ultimate goal was to select more straightforward models versus models with higher accuracy (higher R) but more complexity to pick an optimal one

to assess WQ attributes across Trichonis Lake. Then, the optimal quantitative models developed based on field sampling of 30 August 2014 and satellite image L8 of the same date, were applied to the Landsat image of 30/10/2013 to assess and validate their efficiency by comparing the resulting estimates with the respective available *in-situ* measurements.

4.2.3 Results

4.2.3.1 Statistical Summary of Trichonis lake's In-Situ Measurements and WQ Classification

In-situ dataset of both sampling campaigns covered wide ranges of WQ key indicators: Chl-*a*, a $a_{CDOM(420)}$, $a_{CDOM(440)}$, TP, total nitrogen (TN), nitrate, nitrite, phosphate and ammonium concentrations. *In-situ* nitrate, nitrite, phosphate and TN concentrations of 2014 were measured as lower than the detection limit of the instrument used (photometer Merck Nova 400), hence no statistical elaboration was conducted. Data distributions for the rest parameters were skewed with mostly low values and without extremely high values or outliers (Table 4.2.3-1). In general, most values of all parameters of 2013 were measured slightly higher than the values of 2014, without indicating great differences or existence of WQ deterioration in 2013.

Chl-*a* concentrations ranged from 0.5 to 1.4 μg/l with mean value 1.07 μg/l during the sampling campaign of 2013 and between 0.2 and 0.9 μg/l with average value 0.39 μg/l in 2014. Mean values of TP indicate the presence of similar conditions into the lake since those values for both years are equal to 0.04 and 0.02 mg/l for 2013 and 2014, respectively (Figure 4.2.3-1, Figure 4.2.3-2b).

Ammonium concentrations demonstrated even more resembling values, which ranged from 0.02 to 0.06 mg/l in 2013 and from 0.01 to 0.09 mg/l in 2014, with identical mean value equal to 0.03 mg/l. In general, concentrations of Chl-*a* and TP were measured slightly higher in 2013 than the values of 2014 compared to ammonium concentrations (Figure 4.2.3-1, Figure 4.2.3-2a; 4.2.3-2b). Those values though are slightly increased; thus, no WQ deterioration is indicated in 2013.

Table 4.2.3-1. Descriptive statistics-Summary table of *in-situ* Chl-*a*, TP and ammonium concentrations and $a_{CDOM\ (420)}$, $a_{CDOM\ (440)}$ of 2013 and 2014.

	N	Minimum		Maximum		Mean		Std. Deviation		Skewness			
										Statistic		Std. Error	
		2013	2014	2013	2014	2013	2014	2013	2014	2013	2014	2013	2014
Chl*a* (µg/l)	22	.5	0.2	1.4	.88	1.07	.39	.22	.14	-.51	2.15	.49	.49
a_{cdom} (420)	22	.1	.08	.4	.4	.19	.22	.09	.09	1.35	.46	.49	.49
a_{cdom} (440)	22	.07	.06	.33	.38	.16	.18	.07	.09	1.34	.97	.49	.49
TP (mg/l)	22	.03	.01	.08	.06	.04	.02	.013	.012	1.2	1.9	.49	.49
NH_4^+ (mg/l)	22	.02	.01	.06	.09	.03	.03	.01	.02	2.1	1.99	.49	.49

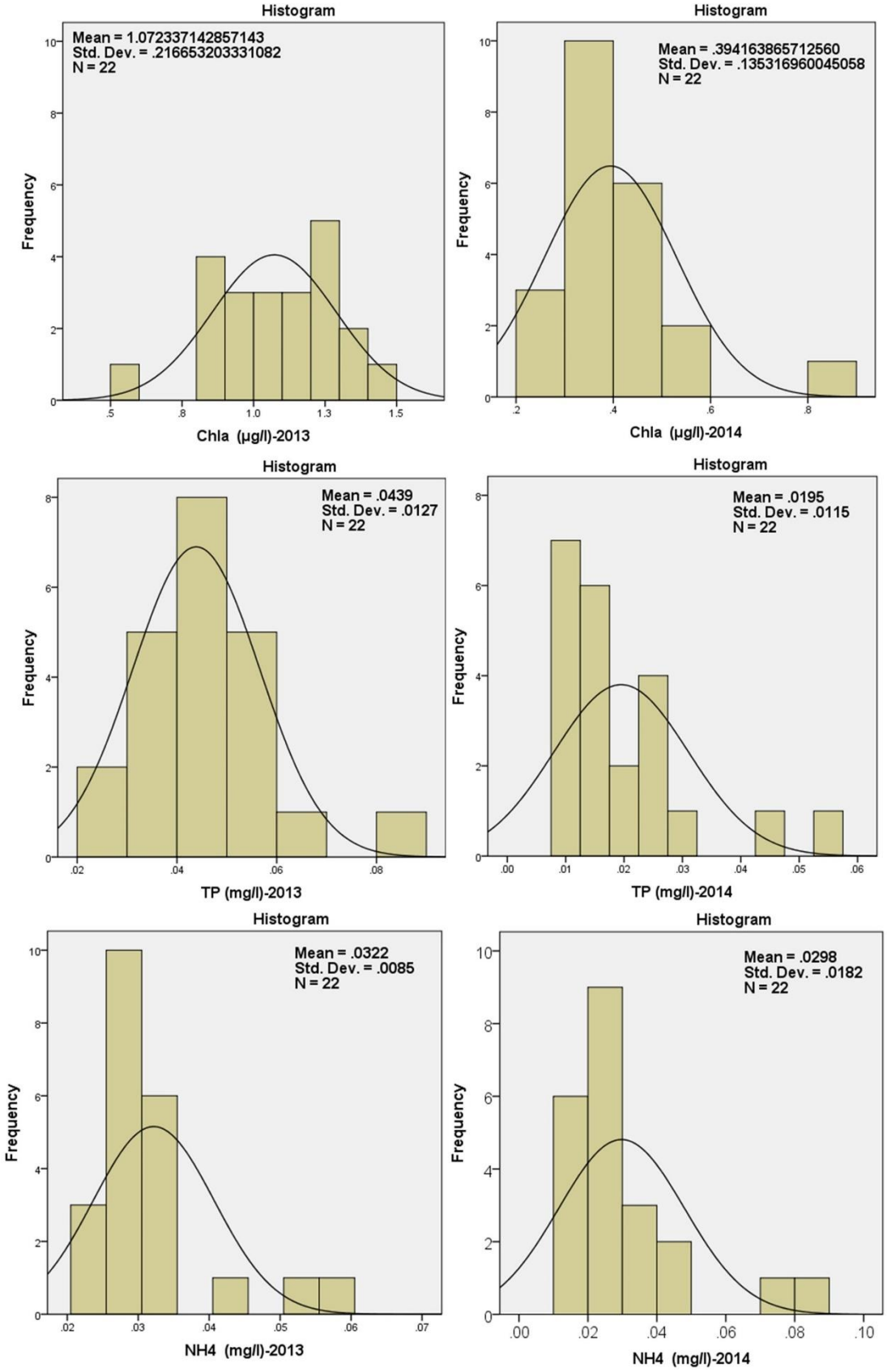


Figure 4.2.3-1. Frequency graphs presenting the distribution of the in-water constituents Chl-*a*, TP and NH_4^+ for both years.

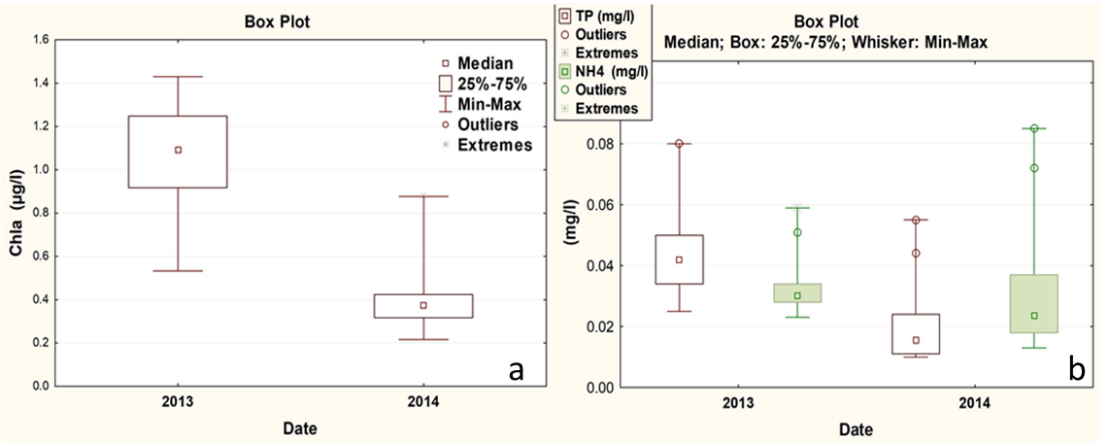


Figure 4.2.3-2. Temporal boxplots presenting basic descriptive statistics (median, percentiles, min-max, outliers and extremes) over sampling season of (a) Chl-*a* and (b) TP concentrations.

In order to classify the WQ of Trichonis lake, the EPA classification system (EPA, 2000) was used; however, in the context of this study, there are no available data of Secchi disk. Therefore, it should be noted that this WQ classification effort is developed to better understand the prevailing conditions during the sampling periods and not to definitely classify the WQ of the Trichonis Lake. Considering the concentrations of TP and Chl-*a* and the estimated average TSI of both the sampling campaigns, Trichonis Lake is characterized as oligotrophic to oligomesotrophic in 2013 and oligotrophic in 2014 (Table 4.2.3-2).

Table 4.2.3-2. EPA lake WQ classification system and estimated TSI for Trichonis Lake.

Date	TSI (TP)	TSI (Chl-*a*)	TSI average	Classification
2013	57.4	31.3	44.3	oligotrophic to oligomesotrophic
2014	47.35	21.4	34.4	oligotrophic

4.2.3.2 MLR analysis and regression models

MLR analysis concerning the *in-situ* data and L8 band combinations of 2013 returned statistical insignificant results. Regarding the correlations accompanied by the highest values of correlation coefficient, correlation analysis was subsequently attempted between the *in-situ* data of 2013 and mean remote sensed values. Those values were retrieved from 90 m buffer zones that were created around each sampling station and were transformed

into surface reflectance. The retrieval of a mean reflectance value around each *in-situ* sampling site was considered more appropriate in order to reduce sensor and algorithm noise (Hu et al., 2001). Those results were equally statistically insignificant. Subsequently, *in-situ* data and band combinations of satellite image of 2014 were correlated. This correlation analysis, after having tested more than 45 band combinations, yielded more statistically acceptable results compared to data of 2013.

Low and statistically insignificant relationships were detected particularly among reflectance values and TP concentrations while the most remarkable (but still statistical insignificant) results, are presented below (Table 4.2.3-3). TP quantitative model (1; Table 4.2.3-3) yielded R and R^2 values equal to 0.27 and 0.07, respectively while the predictors included are the subtraction between bands Green and Red and the natural logarithm of Red and Green ratio (Table 4.2.3-3). Then, using the backward linear regression, ln (Red/Green) was removed (Table 4.2.3-4) and quantitative model (2; Table 4.2.3-3) resulted in R and R^2 values equal to 0.24 and 0.06, respectively while Durbin-Watson's statistic indicates an absence of autocorrelation in the residuals (Table 4.2.3-3). Considering certain statistical indices (especially the value of R^2), all predictive models of TP were rejected due to their low performance.

Concerning the Chl-*a* regression model, coefficients Blue/ (Ultra Blue+Blue+Green) and (Ultra Blue+Blue)/2 were indicated by the predictor importance chart (Figure 4.2.3-3a) and were used presenting acceptable multicollinearity statistics with values of tolerance and VIF 0.96 and 1.04, respectively.

Concerning the spatial distribution of spectral indices, measured from satellite image of 2014, slight differences and variance were also indicated (Table 4.2.3-5). Moreover, the highest value range is apparent in NDVI values (0.0227) while the lowest is in the EVI index. Regarding all indices, no great difference is detected in maximum and minimum values, indicating once again the high spatial homogeneity and the lack of variability that characterizes Trichonis Lake. For the logChl-*a* model, four (4) equations were evaluated, with the following independent variables: the EVI, NDWI, MNDWI and NDVI vegetation and water indices (Table 4.2.3-6). Chl-*a* can be measured initially by using vegetation indices and by extension based on the Green and SWIR bands of water indices (NDWI, MNDWI) due to Chl-*a* absorbance in violet-blue and orange-red wavelengths and its reflection in green/yellow light. Concerning the selected logChl-*a* quantitative model, vegetation and water spectral indices EVI, NDWI, MNDWI and NDVI were employed presenting marginally acceptable statistics (Table 4.2.3-7).

MLR model involving Landsat 8 bands 2 (Blue), 3 (Green) and 4 (Red) proved to be the most suitable for predicting CDOM absorption at 420 nm in

Trichonis Lake (Table 4.2.3-7). Correlation coefficient equals to 0.48 (training data) while Durbin-Watson value indicates independence of residuals. The optimal estimating model of ammonium concentration includes the bands 1 (Ultra-Blue), 3 (Green), and 4 (Red) (Table 4.2.3-7; Figure 4.2.3-3b), while the value of the correlation coefficient is equal to 0.26 (training data). Collinearity statistics (Tolerance and VIF) of the coefficients are 1, excluding the possibility of multicollinearity.

Table 4.2.3-3. Regression analysis **s**tatistics and models' summary among reflectance values and total phosphorus concentrations (dependent variable).

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin - Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.268[a]	.072	-.026	.0117	.072	.734	2	19	.49	1.611
2	.239[b]	.057	.010	.0115	-.015	.301	1	19	.59	
Dependent Variable: TP (mg/l)										
a. Predictors: (Constant), B3-B4, ln (B4/B3)										
b. Predictors: (Constant), B3-B4										

Table 4.2.3-4. Variables entered/removed from TP predictive models depending on the regression method used.

Model	Variables Entered	Variables Removed	Method
1	B3-B4, ln (B4/B3)	.	Enter
2	.	ln (B4/B3)	Backward (criterion: Probability of F-to-remove >= .100).
3	.	B3-B4	Backward (criterion: Probability of F-to-remove >= .100).
a. Dependent Variable: TP (mg/l)			

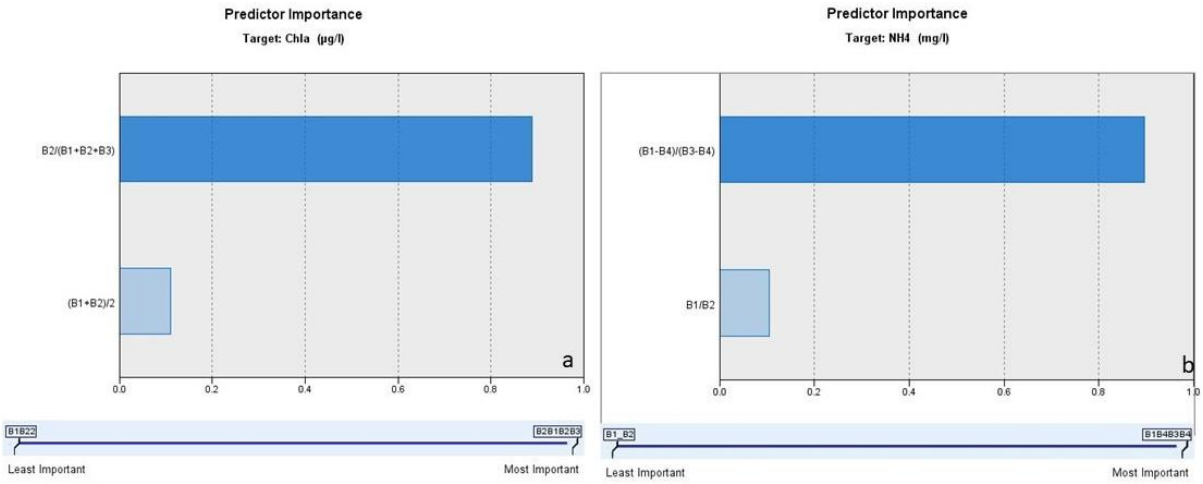


Figure 4.2.3-3. Predictor importance charts indicating the optimal factors for the assessment of Chlorophyll-*a* (a) and Ammonium concentrations (b).

Table 4.2.3-5. Descriptive statistics-Summary tables of selected spectral indices calculated from satellite image of 2014.

2014	N	Range	Min	Max	Mean	Std. Deviation	Skew-ness	Std. Error	Kurto-sis	Std. Error
EVI	22	0.0001	-0.002	-0.002	-0.002	0.0	0.17	0.49	1.32	0.95
NRVI	22	0.0002	-1.002	-1.002	-1.002	0.0	-0.58	0.49	-0.3	0.95
NDWI	22	0.0064	0.86	0.87	0.87	0.002	0.46	0.49	-0.6	0.95
MNDWI	22	0.0021	0.94	0.94	0.94	0.0005	0.02	0.49	0.08	0.95
GNDVI	22	0.0218	-0.424	-0.4	-0.42	0.006	0.72	0.49	-0.08	0.95
NDVI	22	0.023	-0.29	-0.26	-0.28	0.006	0.797	0.491	0.46	0.953

Table 4.2.3-6. Regression analysis statistics and models' summary among multiple spectral indices and log-chlorophyll-*a* concentrations (dependent variable).

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.578[a]	.334	.126	.12	.334	1.608	5	16	.214	
2	.576[b]	.332	.175	.12	-.002	.054	1	16	.819	
3	.493[c]	.243	.117	.12	-.089	2.275	1	17	.150	
4	.473[d]	.224	.142	.12	-.019	.449	1	18	.512	2.235
a. Predictors: (Constant), NDVI, MNDWI, EVI, NDWI2, NRVI										
b. Predictors: (Constant), NDVI, MNDWI, EVI, NDWI2										

c. Predictors: (Constant), NDVI, EVI, NDWI2
d. Predictors: (Constant), NDVI, NDWI2
e. Dependent Variable: LOGCHL-A

Table 4.2.3-7. Statistical summary and description of final water quality parameters' models.

Model	R	R^2	Std. Error of the Estimate	R^2 Change	Durbin-Watson
$acdom420 = -2.195 - (859.4 * Green) + (3426.1 * Red) - 497.51 * [(Blue + Red)/2]$	0.48	0.23	0.08	0.23	1.75
$NH4^{+} = -0.32 + 0.14 * [(Ultra\ Blue - Red)/(Green - Red)]$	0.26	0.07	0.02	0.7	2.33
$CHL - a = -38.62 + 92.05 * [(Blue/(Ultra\ blue + Blue + Green)] + 2239.7 * [(Ultra\ blue + Blue)/2]$	0.44	0.19	0.13	0.19	2.5
$\log Chl - a = -117.64 - (4894.002 * \mathrm{EVI}) - (313.07 * \mathrm{NDWI}) + (433.46 * \mathrm{MNDWI}) + (103.14 * \mathrm{NDVI})$	0.58	0.33	0.12	-0.002	2.24

4.2.3.3 Algorithm validation

In order to explore the reliability of the final regression WQ models, regressions between L8 estimates of Chl-*a*, logchl-*a*, $a_{CDOM(420)}$ and ammonium concentrations in Trichonis Lake versus respective *in-situ* measurements of 2013 were conducted. Several models (linear, logarithmic, quadratic, cubic, power and exponential) have been applied in order to detect the best potential agreement between the observed and satellite-estimated values with the cubic model presenting the highest correlation coefficients for all parameters, except for logChl-*a* where quadratic model was proven to yield slightly better results than the cubic one (Table 4.2.3-8). Nevertheless, the moderate fit between *in-situ* and predicted WQ parameters by each selected MLR indicated the moderate and low quantification capacity of these models.

Besides, the highest correlation coefficient among all validation models, is associated with the ammonium concentration assessment model and it is equal to 0.7 (standard error of estimates 0.004 mg/l), then follows Chl-*a* cubic model with R equal to 0.5 and finally logchl-*a* predictive model (employing spectral indices) with similar values between cubic and quadratic models, 0.4 and 0.41, respectively (Table 4.2.3-8). Following, correlation coefficient of $a_{CDOM(420)}$ was calculated 0.3 with standard error of estimates 0.17 m^{-1} (Table 4.2.3-8).

Table 4.2.3-8. Regression analysis statistics and summary of WQ parameters' regression models, used in validation process. (R^2: the proportion of variance in the dependent variable which can be predicted from the independent variables, R: Correlation coefficient, Std. Error of the Estimate: measure of the accuracy of predictions).

Chl-*a* (μg/l)	**R**	**R^2**	**Std. Error of the Estimate**
linear	0.200	0.040	0.220
logarithmic	0.223	0.05	0.218
quadratic	0.440	0.194	0.207
cubic	0.447	0.199	0.207
power	0.226	0.051	0.094
exponential	0.202	0.041	0.095
Chl-*a* (μg/l) (spectral indices)	**R**	**R^2**	**Std. Error of the Estimate**
Linear	0.34	0.11	0.02
Logarithmic	0.30	0.09	0.02
Quadratic	0.41	0.17	0.02
Cubic	0.40	0.17	0.02
Power	0.30	0.09	0.01
Exponential	0.34	0.11	0.01
$a_{cdom}420$ (m^{-1})	**R**	**R^2**	**Std. Error of the Estimate**
linear	0.11	0.012	0.162
logarithmic	0.131	0.017	0.162
quadratic	0.196	0.038	0.164
cubic	0.258	0.067	0.166
power	0.136	0.018	0.106
exponential	0.118	0.014	0.106
NH_4^+ (mg/l)	**R**	**R^2**	**Std. Error of the Estimate**
linear	0.325	0.106	0.005
logarithmic	0.252	0.064	0.006
quadratic	0.611	0.374	0.005

cubic	0.689	0.474	0.004
power	0.421	0.177	0.379
exponential	0.505	0.255	0.360

Spatial distribution of *in-situ* measurements of NH_4^+ (Figure 4.2.3-4a) and Chl-*a* concentration (Figure 4.2.3-5a, Figure 4.2.3-5a) was mapped through their spatial interpolation using the Spline method. Other interpolation methods, e.g., IDW (Inverse Distance Weighted) and natural neighbour were also tested, but Spline method generated the smoothest surfaces and representative values that were closer to the *in-situ* measured concentrations. The ammonium regression model was applied on the satellite image of 2013 and yielded concentrations ranging from 0 to 0.11 mg/l (Figure 4.2.3-4b) in relation to *in-situ* ammonium distribution which ranged from 0 to 0.08 mg/l (Figure 4.2.3-4a). Pixels having negative values were deleted and the few remained are illustrated with black colour.

Furthermore, the application of the regression models on the satellite data of 2013 indicated some increasing or decreasing assessment trends compared to the respective *in-situ* data. In particular, Chl-*a* regression models overestimated the actual Chl-*a* concentrations with the main difference that the model retrieved from singe band combinations present a more fluctuated value distribution (Figures 4.2.3-5b) than the Chl-*a* retrieved from specific spectral indices (Figures 4.2.3-6b).

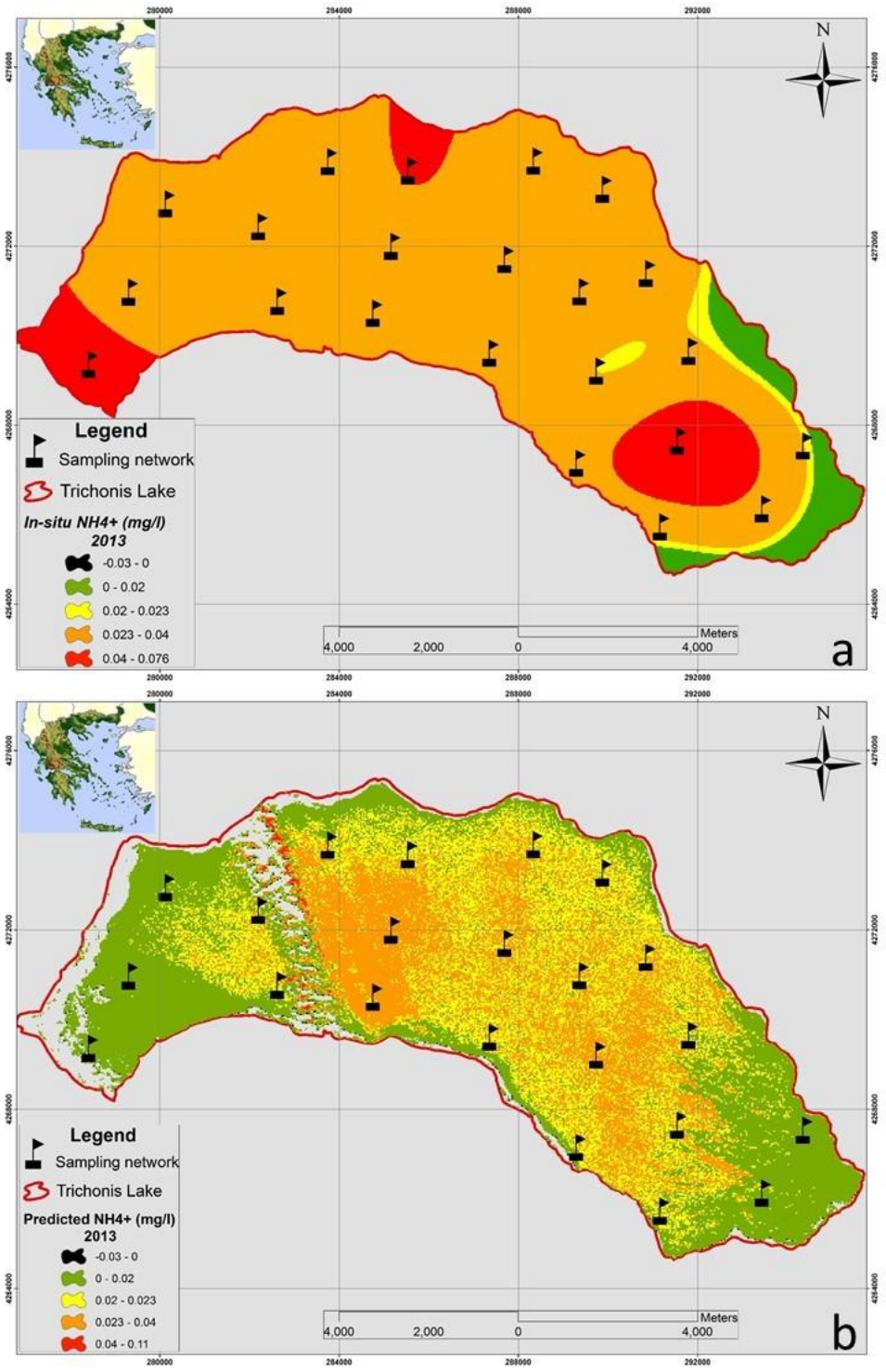


Figure 4.2.3-4. *In-situ* (a) and satellite derived (b) NH_4^+ (mg/l) of 2013 along the Trichonis Lake, after applying the satellite-regression algorithm.

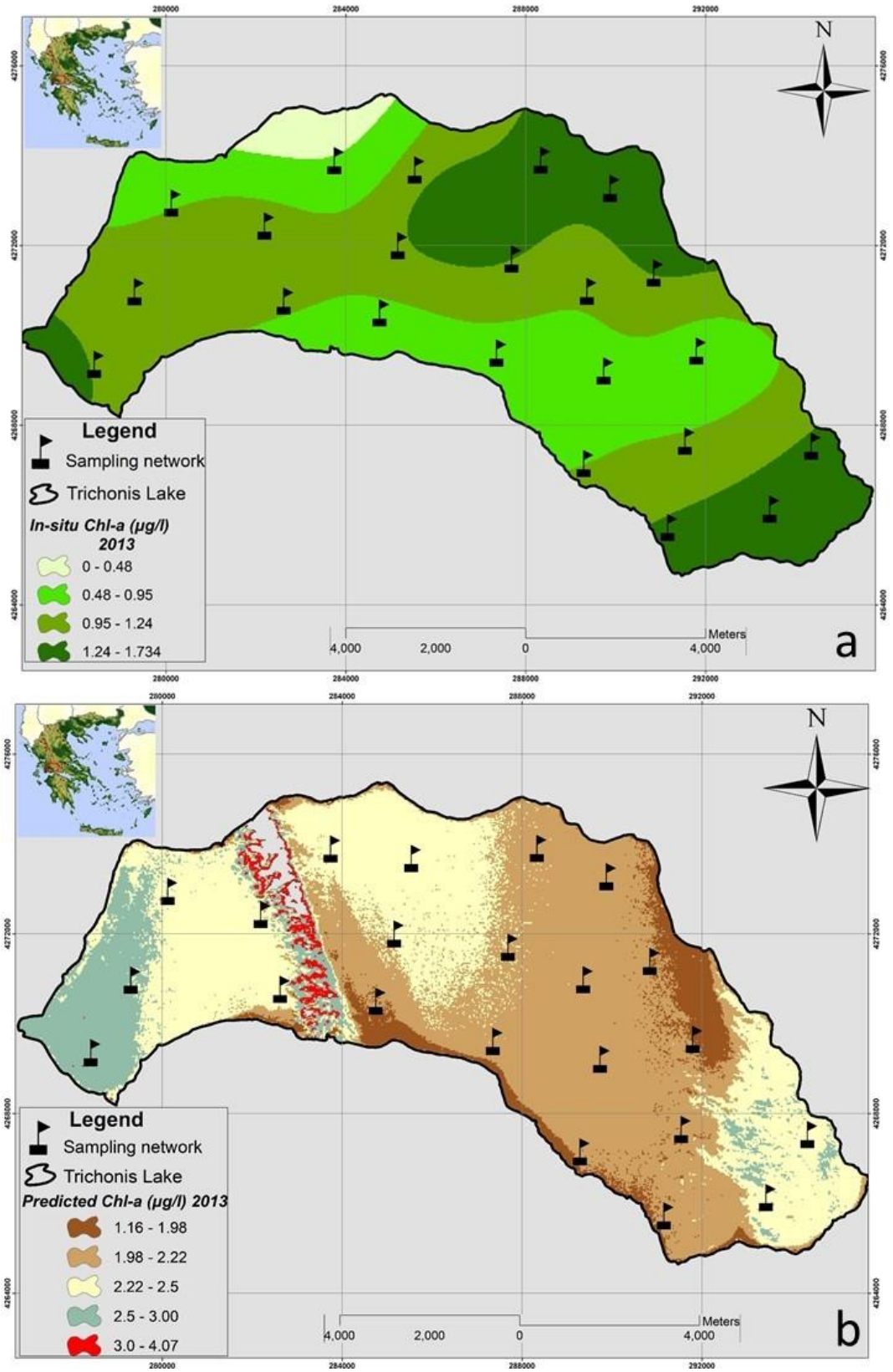


Figure 4.2.3-5. *In-situ* (a) and satellite derived (b) Chl-*a* (μg/l) spatial distribution of 2013 along the Trichonis Lake, after applying the satellite-regression algorithm using L8 bands.

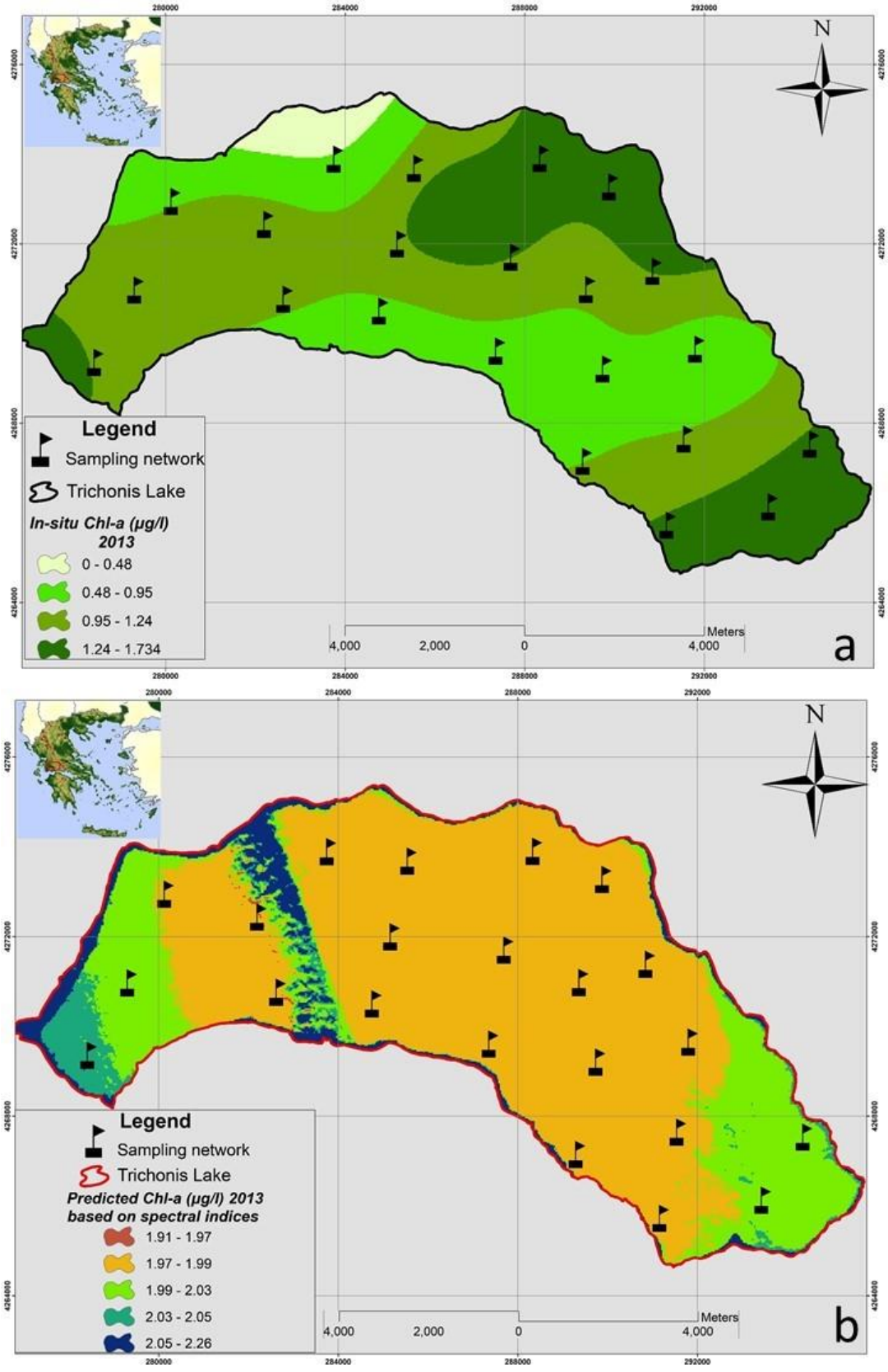


Figure 4.2.3-6. *In-situ* (a) and satellite derived (b) Chl-*a* (μg/l) spatial distribution of 2013 along the Trichonis Lake, after applying the satellite-regression algorithm using spectral indices.

4.2.4 Discussion

Remote sensing provides suitable information concerning WQ and aquatic systems management. In this study, the feasibility of Landsat 8 OLI imagery in combination with *in-situ* WQ parameters' concentrations to identify relevant algorithms for WQ assessment in an oligotrophic waterbody (Trichonis lake) was demonstrated limited. Water samples from Trichonis Lake were analyzed twice in 2013 and 2014 regarding concentrations of Chl-*a*, ammonium and CDOM concentration, which was determined as the absorption at 420 nm, $a_{CDOM(420)}$, extrapolated from the absorption spectra. According to literature and lab measurements, Trichonis Lake is not only characterized as an oligotrophic lake but also illustrates a relatively low quantitative, temporal and spatial variability.

MLRs were conducted among available data and the majority of models were characterized by insignificant statistical correlations. Optimal models were selected based on statistical criteria but presented low coefficients and unsuccessful results. The selected predictive model of Chl-*a* concentration involves the combination of Ultra-Blue (B1), Blue (B2) and Green (B3) OLI bands of L8 satellite sensor. These results are in accordance with those of Pahlevan et al. (2014), who attempted to map OLI's spectroradiometric sensitivity to changes in optically active components (OACs), such as Chl-*a*, for a nominal solar zenith angle $\theta s=40^{o}$, (solar zenith angle in our study equals to $\theta s=35^{o}$). According to their results, the Blue band (B2) shows the highest sensitivity to changes in Chl-*a*, in particular on average for changes higher than 0.5 μg/l. This implies difficulties in detecting changes smaller than 0.5 units of Chl-*a* on the focal plane using this single band. While the Ultra Blue (B1) and the green bands (B3), on average, exhibit similar sensitivity to the changes in Chl-*a*, the B1 band is slightly better for waters with low Chl-*a* concentrations.

Addittionally, the logChl-*a* predictive model based on spectral indices incorporated OLI bands 2 (Blue), 3 (Green), 4 (Red), 5 (NIR) and 7 (SWIR2) with R equal to 0.58 (training data). Brezonik et al. (2005); Olmanson et al. (2008); Fadel et al. (2016) and Bohn et al. (2017) used similar bands for the estimation of Chl-*a* in lakes and reservoirs and more particular vegetation indices and bands TM and ETM 1 (Blue), 2 (Green) and 4 (NIR). Bohn et al. (2017) used the NDVI in Laguna Chascomús in relation to Chl-*a* estimation for its optical characteristics and because it is sensitive to the pigment absorption. NDVI has been found to be very sensitive to changes in the environment (Kahru et al., 1993; Bohn et al., 2017). Moreover, its use is more successful in zones with moderate wind speeds without developing waves, which is not the case in Trichonis Lake. Furthermore, water indices include SWIR band and according to Barrett and Frazier (2016) all significant band

combinations for chlorophyll included at least one of the short-wave infrared bands (SWIR), although most WQ studies to this point have not included SWIR bands. Ratios between either chlorophyll absorption bands (Red and Blue) or chlorophyll reflectance bands (Green and NIR) with either of the two SWIR bands are expected to emphasize the portion of the spectrum affected by chlorophyll, thereby making estimated values more readily correlated with actual sample values (Barrett and Frazier, 2016).

As far as the utilized satellite sensor is concerned, Bonansea et al. (2018) tried to generate a different Chl-*a* model for different Landsat sensors (5 TM, 7 ETM+ and 8 OLI). Although OLI sensor has better radiometric sensitivity and signal to noise ratio, they could not prove that OLI is better than TM and ETM+ sensors. Overall, they observed that each Landsat sensor can be used to estimate Chl-*a* in the reservoir while the best model for TM sensor included a combination of Green, Red and NIR band, and the ratio green/red ($R^2 = 0.92$). A three-variable model using Green and SWIR-1 bands and the ratio red/green was the best model to predict Chl-*a* using EMT+ sensor ($R^2 = 0.91$).

Concerning the assessment of CDOM absorption via RS, Pahlevan et al. (2014) explored the detection limits associated with CDOM absorption at 440 nm. While hereby study resulted in a predictive model for CDOM absorption at 420 nm combining the Blue, Green and Red bands, Pahlevan et al. (2014) found out that in waters with relatively low CDOM concentrations, ($a_{CDOM(440)}<0.5\ m^{-1}$), the Blue and the Green bands exhibit the highest sensitivity whereas the Red band was found insensitive to the changes in CDOM absorption. In general, it was found that OLI is, on average, sensitive to changes in CDOM absorption larger than 0.1 m^{-1}. Although actual retrievals can be improved by the use of multiple bands, the fact that in Trichonis lake detected changes in Chl-*a* concentrations and to a lesser extent in CDOM absorption are marginally equal to the aforementioned threshold values (0.5 μg/l and 0.1 m^{-1}, respectively), could be the main reason of not managing high-precision assessment results. Furthermore, Pahlevan and Schott (2013) applied a physics-based approach to fully examine the potential of OLI in CDOM absorption mapping. Based on their observations they concluded that the disparity between the response functions of OLI is more noticeable in turbid waters than in clearer waters when mapping CDOM absorption.

Development of reliable methods to retrieve CDOM information from spectral reflectance data is difficult. Indeed, among the major WQ variables measurable by remote sensing (e.g., suspended solids, chlorophyll, Secchi depth), for several reasons CDOM may be the most difficult to measure accurately in inland waters. CDOM absorbs but does not scatter or reflect light while it has no absorbance troughs or peaks, such as are found for plant pigments; instead light absorption by CDOM follows a simple quasi-

exponential decrease with increasing wavelength. There are no wavelength bands in the visible spectrum uniquely associated with CDOM that can be used for measurement purposes. Thus, measurement of low to moderate levels of CDOM in optically complex Case-2 waters is especially difficult because light scattering by these particles dominates their reflectance spectra (Brezonik and Olmanson, 2015).

Predicting ammonium concentration in inland waters can be a hard task since very few studies have attempted to monitor data with non-optical properties, such as nutrient concentrations. Furthermore, not many previous studies have been able to provide total nitrogen models with statistically significant results or reasonable adjusted R^2 values (Isenstein and Park, 2014). Hereby research resulted in the ammonium predictive model incorporating Ultra-Blue, Green and Red bands yielding a regression coefficient equal to 0.7, regarding the validation process. Similar results, concerning the utilized wavelengths, were presented by Dewidar and Khedr (2001) and Isenstein and Park (2014), who detected the strongest correlation among TN and Landsat TM bands 1 (Blue) and 2 (Green). Chen and Quan (2012) predicted TN concentrations with Landsat TM bands 1 (Blue), 2 (Green), 3 (Red), and 4 (NIR), however these results were not very successful (R^2 = 0.24). Effective and precise WQ determination is dependent on the satellite sensor used, the methodology followed and also on the nature of the waters studied (Case-1, Case-2). Based in these premises, Gons et al., (2008) attempted to estimate Chl-*a* concentration by using MERIS images and they concluded that the application of MERIS FLH algorithms in oligotrophic waters may indeed be precluded because of too low signal to noise ratio.

All in all, in this study results showed that WQ monitoring of oligotrophic freshwater bodies through RS tools can be a really challenging task. Landsat 8 has been widely used in eutrophic lakes and even fewer studies have managed to estimate nutrients, particularly ammonium concentrations. Season of water samplings, lake trophic status and the spatial homogeneity may be the greatest limitations that prevented a better and more accurate prediction.

5.2.5 Conclusions

Hereby study explored the use of RS technology and specifically of Landsat 8 OLI sensor, to accurately quantify certain WQ parameters in Trichonis lake.

According to the *in-situ* data analysis and their spatial distribution, it has been strongly ascertained that Trichonis Lake is characterized by particularly low concentrations and the lack of any spatial or temporal value

differentiation across the twenty-two sampling stations, case that inhibited a greater predictive potential.

Furthermore, weak correlations were detected among *in-situ* and satellite data while those correlations, particularly in autumn and summer, may also be due to the lake turnover effect. When the equalization of the thermal gradient in the lake induces mixing of surface and bottom waters, remote monitoring is made difficult due to instability (Barrett and Frazier, 2016). Moreover, the incorporation of the SWIR band into Chl-*a* estimation (in contrast to other studies) suggests that there may be a relationship between SWIR reflection and algae/plant production, which deserves further investigation.

Additional water samplings should be made during different time periods concerning specific mixing boundaries (surface-bottom waters) in order to investigate whether the feasibility of remote monitoring increases. In case strong relationships are found, this may help improve prediction capabilities by providing researchers with bounded time periods (according to region) (Barrett and Frazier, 2016).

Further research is required towards the investigation of more water parameters or using sensors of different spatial and geometrical analysis in order to be able to compare the outcomes among all different cases.

Even though early results demonstrated the vulnerability of the Landsat 8 imagery to precisely determine certain WQ components in an inland oligotrophic body, it is generally accepted that those models may initially increase the knowledge of Trichonis lake's WQ and then be utilized as warning indicators of WQ deterioration.

4.2.6 References

Albright, T.P.; Ode, D.J. Monitoring the dynamics of an invasive emergent macrophyte community using operational remote sensing data. *Hydrobiologia* **2011**, vol. 661, pp. 469-474.

Baret, F.; Guyot, G. Potentials and limits of vegetation indices for LAI and PAR assessment. *Remote Sens. Environ.* **1991**, 35, 161–173.

Barrett, D.; Frazier, A. Automated Method for MonitoringWater Quality Using Landsat Imagery. *Water* **2016**, 8, 257.

Barsi, J.A.; Lee, K.; Kvaran, G.; Markham, B.L.; Pedelty, J.A. The Spectral Response of the Landsat-8 Operational Land Imager. *Remote Sens.* **2014**, vol. 6, pp. 10232-10251, doi: 10.3390/rs61010232.

Bohn, V.; Carmona, F.; Rivas, R.; Lagomarsino, L.; Diovisalvi, N.; Zagarese, H. Development of an empirical model for chlorophyll-a and Secchi Disk Depth estimation for a Pampean shallow lake (Argentina). *Egypt. J. Remote Sens. Space Sci.* **2017**.

Bonansea, M.; Fernandez, R.L. Remote sensing of suspended solid concentration in a reservoir with frequents wildland fires on its watershed. *Water Science and Technology* **2013**, vol. 67, no.1, pp. 217–223.

Bonansea, M.; Rodriguez, C.; Pinotti, L. Assessing the potential of integrating Landsat sensors for estimating chlorophyll-a concentration in a reservoir. *Hydrol. Res.* **2018**, 49 (5), 1608-1617.

Bonansea, M.; Rodriguez, M.C.; Pinotti, L.; Ferrero, S. Using multi-temporal Landsat imagery and linear mixed models for assessing water quality parameters in Río Tercero reservoir (Argentina). *Remote Sens. Environ.* **2015**, 15, 28–41.

Brezonik, P. L; Olmanson, G. L.; Finlay, J.C.; Bauer, M. E. Factors affecting the measurement of CDOM by remote sensing of optically complex inland waters. Remote Sensing of Environment, **2015**, 157, pp. 199-215, ISSN 0034-4257, http://dx.doi.org/10.1016/j.rse.2014.04.033.

Brezonik, P. L; Olmanson, G. L.; Finlay, J.C.; Bauer, M. E. Factors affecting the measurement of CDOM by remote sensing of optically complex inland waters. *Remote Sensing of Environment*, **2015**, 157, pp. 199-215, ISSN 0034-4257, http://dx.doi.org/10.1016/j.rse.2014.04.033.

Brezonik, P.; Menken, D.; Bauer, M. Landsat-based remote sensing of lake water quality characteristics, including chlorophyll and colored dissolved organic matter (CDOM). *Lake and Reservoir Management* **2005**, 21(4), 373–382.

Carlson, R.E.; Simpson, J. A Coordinator's Guide to Volunteer Lake Monitoring Methods; North American Lake Management Society: Madison, WI, USA, **1996**, 96p.

Chen, J.; Quan, W.T. Using Landsat/TM imagery to estimate nitrogen and phosphorus concentration in Taihu Lake, China. *IEEE J-STARS* **2012**, 5(1), 273–280, DOI: 10.1109/JSTARS.2011.2174339.

Dewidar, K.; Khedr, A. Water quality assessment with simultaneous Landsat-5 TM at Manzala Lagoon, Egypt. *Hydrobiologia* **2001**, vol. 457, no.1–3, pp. 49–58.

Doña, C.; Chang, N.B.; Caselles, V.; Sánchez, J.M.; Camacho, A.; Delegido, J.; Vannah, B.W. Integrated satellite data fusion and mining for monitoring lake water quality status of the Albufera de Valencia in Spain. *J. Environ. Manag.* **2015**, 151, 416–426.

Eleveld MA. Wind-induced resuspension in a shallow lake from Medium Resolution Imaging Spectrometer (MERIS) full-resolution reflectances. *Water Resour. Res.* **2012**, 48: W04508.

EPA—US Environmental Protection Agency. Nutrient criteria, technical guidance manual, lakes and reservoirs. First Edition, EPA-822-B00-001, **2000**.

Fadel, A.; Faour, G.; Slim, K. Assessment of the trophic state and chlorophyll-a concentrations using Landsat OLI in Karaoun reservoir, Lebanon. *Leban. Sci. J.* **2016**, 17, 130–145.

Gao, B.C. NDWI—A Normalized Difference Water Index for Remote Sensing of Vegetation LiquidWater from Space. *Remote Sens. Environ.* **1996**, 58, 257–266.

Gholizadeh, M.; Melesse, A.; Reddi, L. A comprehensive review on water quality parameters estimation using remote sensing techniques. *Sensors* **2016**, 16 (8), 1298. https://doi.org/10.3390/s16081298.

Giardino, Claudia; Bresciani, Mariano; Stroppiana, Daniela; Oggioni, Alessandro; Morabito, Giuseppe. Optical remote sensing of lakes: an overview on Lake Maggiore. *J. Limnol.*, **2014**, *73(s1): 201-214, DOI: 10.4081/jlimnol.2014.817*

Gitelson, A.A.; Kaufman, Y.J.; Merzlyak, M.N. Use of a green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sens. Environ.* **1996**, 58, 289–298.

Gons, H. J.; Auer, M. T.; Effler, S. W. MERIS satellite chlorophyll mapping of oligotrophic and eutrophic waters in the Laurentian Great Lakes", *Remote Sensing of Environment* **2008**, 112, pp. 4098–4106.

Guan, X.; Li, J.; Booty, W.G. Monitoring lake Simcoe water clarity using Landsat-5 TMimages. *Water Resources Management* **2011**, 25(8), 2015–2033.

Guang, J.; Wey, Y.; Huang, J. A model for the retrieval of suspended sediment concentrations in Taihu Lake from TM images. *Journal of Geographical Sciences* **2006**, 16(4), pp. 458–464.

Gunn, J. M.; Snucins, E.; Yan, N. D.; Arts, M. T. Use of water clarity to monitor the effects of climate change and other stressors on oligotrophic lakes. *Environmental Monitoring and Assessment* **2001**, 67, 69–88, 2001.

Hadjimitsis, D.G.; Hadjimitsis, M.G.; Clayton, C.; Clarke, B.A. Determination of turbidity in Kourris dam in Cyprus utilizing Landsat TM remotely sensed data. *Water Resources Management* **2006**, 20(3), 449–465.

Holm-Hansen, O.; Lorenzen, C. J.; Hormes, R. W.; Strickland, J. D. H. Fluorometric determination of chlorophyll. *Journal du Conseil/Conseil Permanent International pour l'Exploration de la Mer* **1965**, 30, 3–15, 1965, doi:10.1093/icesjms/30.1.3.

Hu, C.M.; Muller-Karger, F.E.; Andrefouet, S.; Carder, K.L. Atmospheric correction and cross-calibration of Landsat-7/ETM+ imagery over aquatic environment: A multiplatform approach using SeaWiFS/MODIS. *Remote Sens. Environ.* **2001**, 78, 99–107.

Huang, C.; Zou, J.; Li, Y.; Yang, H.; Shi, K.; Li, J.; Chena, X.; Zheng, F. Assessment of NIR-red algorithms for observation of chlorophyll-a in highly turbid inland waters in China. *ISPRS J. Photogramm. Remote Sens.* **2014**, 93, 29–39.

Isenstein, E. M.; Park, Mi-H. Assessment of nutrient distributions in Lake Champlain using satellite remote sensing. *Journal of Environmental Sciences* **2014**, 26, 1831 – 1836, http://dx.doi.org/10.1016/j.jes.2014.06.019.

Kahru, M.; Leppanen, J.M.; Rud, O. Cyanobacterial blooms cause heating of the sea surface. *Mar. Ecol. Prog. Ser.* **1993**, 101, 1–7.

Karakaya, N.; Evrendilek, F.; Aslan, G.R.; Gungor, K.; Karakas, D. Monitoring of lake water quality along with trophic gradient using landsat data. *International Journal of Environmental Science and Technology* **2011**, 8(4), 817–822.

Kutser, T. The possibility of using the Landsat image archive for monitoring long time trends in coloured dissolved organic matter concentration in lake waters. *Remote Sensing of Environment* **2012**, 123, 334–338.

Kutser, T.; Paavel, B.; Metsamaa, L. Mapping coloured dissolved organic matter concentration in coastal waters. *Int. J. Remote Sens.* **2009**, 30, pp. 5843-5849.

Kutser, T.; Pierson, D.C.; Kallio, K.; Reinart, A.; Sobek, S. Mapping lake CDOM by satellite remote sensing. *Remote Sens. Environ.* **2005**, 94, 535-540.

Le, C.; Li, Y.; Zha, Y.; Sun, D.; Huang, C.; Lu, H. A four-band semi-analytical model for estimating chlorophyll-a in highly turbid lakes: The case of Taihu Lake, China. *Remote Sens. Environ.* **2009**, 113, 1175–1182.

Liu, H.Q.; Huete, A.R. A feedback-based modification of the NDVI to minimize canopy background and atmospheric noise. *IEEE Trans. Geosci. Remote Sens.* **1995**, 33, 457–465.

Lyu Lili, Song Kaishan, Wen Zhidan, Liu Ge, Shang Yingxin, Li Sijia, Tao Hui, Wang Xiang, Hou Junbin. Estimation of the lake

trophic state index (TSI) using hyperspectral remote sensing in Northeast China. *Opt. Express* **2022**, 30, 10329-10345. https://doi.org/10.1364/OE.453404.

Matthews, M.W. A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters. *International Journal of Remote Sensing*, **2011**, 32(21), 6855–6899.

McCullough, I.M.; Loftin, C.S.; Sader, S.A. Combining lake and watershed characteristicswith Landsat TMdata for remote estimation of regional lake clarity. *Remote Sensing of Environment* **2012a**, 123, 109–115.

McFeeters, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, 17, 1425–1432.

Memon, A.A.; Muhammad, S.; Rahman, S.; Haq, M. Flood monitoring and damage assessment using water indices: A case study of Pakistan flood-2012. *Egypt. J. Remote Sens. Space Sci.* **2015**, 18, 99–106.

Nikolaidis, N. P., Phillips, G., Poikane, S., Várbíro, G., Bouraoui, F., Malago, A., & Lilli, M.A. River and lake nutrient targets thatsupport ecological status: European scale gap analysis and strategiesfor the implementation of the Water Framework Directive.Science ofthe Total Environment **2022**, 813(March), 151898, https://doi.org/10.1016/j.scitotenv.2021.151898.

Odermatt, D.; Gitelson, A.; Vittorio, E.V.; Schaepman, M. Review of constituent retrieval in optically deep and complex waters from satellite imagery. *Remote Sens. Environ.* **2012**, 118, 116–126.

Olmanson, L.G.; Bauer, M.E.; Brezonik, P.L. A 20-year Landsat water clarity census of Minnesota's 10,000 lakes. *Remote Sens. Environ.* **2008**, 112, 4086–4097.

Onderka, M.; Pekárová, P. Retrieval of suspended particulate matter concentrations in the Danube River from Landsat ETM data. *Science of the Total Environment* **2008**, 397(1), 238–243.

Pahlevan, N.; Schott, J.R. Leveraging EO-1 to evaluate capability of new generation of Landsat sensors for coastal/inland water studies. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2013**, 6, 360–374.

Pahlevan, N.; Wei, J.; Schaaf, C. B.; Schott, J. R. Evaluating Radiometric Sensitivity of Landsat 8 over coastal / inland waters. Geoscience and Remote Sensing Symposium (IGARSS), **2014** IEEE International, 13-18 July 2014, Quebec City, QC, Canada. DOI: 10.1109/IGARSS.2014.6946695.

Pizani, F.; Maillard, P. The determination of water quality parameters by remote sensing technologies: 2000 -2020. Universidade Federal de Minas Gerais, [S. l.] **2022**, 1–30. DOI: 10.13140/RG.2.2.32203.87849.

Randolph, K.; Wilson, J.; Tedesco, L.; Li, L.; Pascual, D.L.; Soyeux, E. Hyperspectral remote sensing of cyanobacteria in turbid productive water using optically active pigments, chlorophyll-a and phycocyanin. *Remote Sens. Environ.* **2008**, 112, 4009–4019.

Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. Monitoring Vegetation Systems in the Great Plains with ERTS; NASA SP-351; NASA: Washington DC, USA, 1974, 1, 309–317.

Ruddick, K.G.; Ovidio, F.; Rijkeboer, M. Atmospheric correction of SeaWiFS imagery for turbid coastal and inland waters. *Appl. Opt.* **2000**, 39, 897–912.

Ruiz-Verdu, A; Simis, SGH; de Hoyos, C.; Gons, H.J.; Peọa- Martvnez, R. An evaluation of algorithms for the remote sensing of cyanobacterial biomass. *Remote Sens. Environ.* **2008**, 112, 3996-4008.

Seegers, N. Bridget; Werdell, P. Jeremy; Vandermeulen, A. Ryan, Salls, Wilson, Stumpf, P. Richard; Schaeffer, A. Blake; Owens, J. Tommy; Bailey, W. Sean; Scott P. Joel; Loftin A. Keith. Satellites for long-term monitoring of inland U.S. lakes: The MERIS time series and application for chlorophyll-a, *Remote Sensing of Environment* **2021**, 266, 112685, ISSN 0034-4257, https://doi.org/10.1016/j.rse.2021.112685.

Song, K.S.; Li, L.L.; Tedesco, S.; Li, H.; Duan, T.; Liu, D.W.; Hall, B.; Du, J.; Li, Z.C.; Shi, K.; et al. Remote estimation of chlorophyll-a in turbid inland waters: Three-band model versus GA-PLS model. *Remote Sens. Environ*. **2013**, 136, 342–357.

Sun, D.; Hu, C.; Qiu, Z.; Cannizzaro, J.P.; Barnes, B.B. Influence of a red band-based water classification approach on chlorophyll algorithms for optically complex estuaries. *Remote Sens. Environ.* **2014**, 155, 289–302.

Tebbs, E.J.; Remedios, J.J.; Harper, D.M. Remote sensing of chlorophyll-*a* as a measure of cyanobacterial biomass in Lake Bogoria, a hypertrophic, saline–alkaline, Flamingo Lake, using Landsat ETM+. *Remote Sensing of Environment* **2013**, 135, 92–106.

Vincent, R.K.; Qin, X.; McKay, R.M.; Miner, J.; Czajkowski, K.; Savino, J. et al. Phycocyanin detection from LANDSAT TM data for mapping cyanobacterial blooms in Lake Erie. *Remote Sensing of Environment* **2004**, 89(3), 381–392.

Welschmeyer, N. A. Fluorometric analysis of chlorophyll a in the presence of chlorophyll b and pheopigments. *Limnol. Oceanogr.* **1994**, 39. http://dx.doi.org/10.4319/lo.1994.39.8.1985.

Xu, H. Modification of Normalized Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery. *Int. J. Remote Sens.* **2006**, 27, 3025–3033.

Yacobi, Y.Z.; Moses, W.J.; Kaganovsky, S.; Sulimani, B.; Leavitt, B.C.; Gitelson, A.A. NIR-red reflectance-based algorithms for chlorophyll-a estimation in mesotrophic inland and coastal waters: Lake Kinneret case study. *Water Research* **2011**, 45, 2428–2436.

Yang, W.; Matsushita, B.; Chen, J.; Fukushima, T.; Ma, R.H. An enhanced three band index for estimating chlorophyll-a in turbid case-II waters: Case studies of Lake Kasumigaura, Japan, and Lake Dianchi, China. *IEEE Geosci. Remote Sens. Lett.* **2010**, 7, 655–659.

Yang, Z.; Reiter, M.; Munyei, N. Estimation of chlorophyll-a concentrations in diverse water bodies using ratio-based NIR/Red indices. *Remote Sens. Appl. Soc. Environ*. **2017**, 6, 52–58.

Zacharias, I.; Dimitriou, E.; Koussouris, Th. Developing sustainable water management scenarios by using thorough hydrologic analysis and environmental criteria. *Journal of Environmental Management* **2003**, 69, 401–412.

Zacharias, I.; Dimitriou, E.; Koussouris, Th. Integrated water management scenarios for wetland protection: application in Trichonis Lake. *Environmental Modelling & Software* **2005**, 20, 177-185.

Zhao, D.; Cai, Y.; Jiang, H.; Xu, D.; Zhang, W.; An, S. Estimation of water clarity in Taihu Lake and surrounding rivers using Landsat imagery. *Advances in Water Resources* **2011**, 34(2), 165–173.

Zhou, W.; Wang, S.; Zhou, Y.; Troy, A. Mapping the concentrations of total suspended matter in lake Taihu, China, using Landsat-5 TM data. *Int. J. Remote Sens.* **2006**, 27, 1177-1191.

Zhu, W.; Yu Q.; Tian, Y.Q.; Becker, B.L.; Zheng, T., Carrick, H. J. An assessment of remote sensing algorithms for colored dissolved organic matter in complex freshwater environments. *Remote Sensing*

5. General Conclusions and Limitations

This chapter recapitulates the final key conclusions drawn in the framework of this PhD thesis, accompanied by the scientific contributions offered and limitations accrued per main objective pursued.

Objective 1: Establish a methodological framework that aims to model WQ and trophic status of optically diverse Greek lakes (Case 2 waters) by assessing key WQ elements with fine spatial resolution (10-30 m) RS image data. Ultimate goal of this methodology is the accurate spatial assessement of WQ and trophic status over various types of lakes, thus acquiring the valuable information of their variability in near-real time.

Conclusions:

(a) WFD application in Greece concerning Lake Waterbodies yielded so far significant WQ datasets whose statistical elaboration indicated the great **significance of lakes' nature concerning the constituents' variance.** Furthermore, hereby developed WQ models initially **accommodated the spectral configuration differences among the Landsat sensors** (7 ETM+ and 8 OLI) and then **managed to assess adequately WQ and trophic status of Greek lakes**. Studied WQ parameters include Chl-*a* and TP concentrations and Secchi Disk depths while their employment into Carlson's trophic state index (TSI) equation facilitated the lakes' trophic status assessment. The most optimal Chl-*a* quantitative models include the ratios Blue to Green and Red, Red to Green and Blue, and the ln transformed bands SWIR1 and SWIR2. The $Secchi_{general}$ model incorporated a combination of bands Blue, Red, Green and SWIR2 while models developed for natural and artificial lakes were accompanied by the insertion of logchl-*a* as a significant Secchi predictor. The general TP assessment model includes the logarithmic transformation of Chl-*a* and the band ratio of Ln-Red and Ln-SWIR1 bands while the $TP_{natural}$ model incorporates also the ratio of Green and Red bands. Based on those results, the background knowledge on whether a lake is natural or artificial proved to be valuable concerning the models' predictions's accuracy; hence, it was concluded that **WQ and trophic status assessment of a) artificial and b) shallow (mean depth <5 m) lakes was less successful.** Consequently, a deeper limnological research regarding the primary limiting factors of Chl-*a* production and the predominant sources of

turbidity (algal/non-algal), particularly in reservoirs, is considered essential. This prerequisite is further strengthened by the fact that none TP statistically acceptable model was generated for artificial lakes, partly attributed to the time lag that has been observed for phytoplankton to consume TP in this type of lakes.

Above all, herewith **delivered WQ models were proved capable of supporting perpetual lake WQ monitoring and sustainable management at a national scale, at near real time** and at finer spatial resolution compared to similar, large-scale applications-based on EO techniques- offered worldwide. Furthermore, lake WQ monitoring can be more effectively pursued after obtaining the valuable information of spatial variability, while there is the possibility for WQ and trophic status classification no longer be controlled by only one (1) sampling station. WQ models can potentially be delivered to competent public authority which is responsible for WQ monitoring of Greek lakes; in this way there will be the opportunity to inspect lakes' trophic status at times independently of the scheduled upcoming sampling campaigns.

(b) Development of WQ models and their validation indicated the **high contribution of SWIR bands in WQ monitoring of Case-2 waters**, although they have not been widely used by other studies. It was proven that when Case 2 waters is the case, sediment reflectance exceeds the absorptive properties of water in the SWIR wavelengths, thus a common assumption made in Case 1 waters is abandoned.

Scientific contribution:

This is the first time, to the author's knowledge, that 50 representative Greek lakes' WQ has been modeled efficiently by using *in-situ* data and RS technology. Spatial distribution of lakes' Chl-*a*, Secchi depth, TP values and by extension trophic status can by assessed and monitored continuously in near-real time and in fine spatial resolution, constituting a valuable lake management tool - at a country level- in the hands of scientists and competent authorities.

Given the complexity that characterizes the mapping of WQ elements in Case 2 waters (coastal, lakes, rivers) in combination with the wide study area which includes a broad range of limnological conditions, hereby delivered lake WQ models contribute essentially to sustainable water resources management of Greece.

Limitations:

(a) A more accurate lake WQ assessment was hindered by the intense optical diversity characterizing the under study Greek lakes. Considering the high number (50) of the studied lakes and the fact that their optical properties may vary resembling either to both Case 1 or

Case 2 waters or classes within Case 2 waters, there has been an effort to distinguish distinct optical water types (OWTs) among them based on statistical analysis. However, hierarchical cluster analysis incorporating *in-situ* WQ constituents (concentrations of Chl-*a*, TP and Secchi depths) and Landsat response (respective reflectance at bands blue, green, red, NIR and SWIRS) did not manage to yield reliable results. Classification efforts employed datasets including all, average and median values of WQ data, and median values of WQ data grouped by seasonal sampling. The factor Season was considered since it affects indirectly the eutrophication degree of waterbodies, e.g. the growth of algae, the influx of nutrients caused by rainfall, the re-suspension of suspended matter caused by wind and so on.

(b) One more factor that has potentially prevented a better prediction of lakes' WQ is the existence of an uncertainty of accuracy regarding the location of sampling stations. Water sampling in lakes requires special attention as winds and other external factors (e.g. season, lake depth and changes in water level, ease of proximity) contribute to potential transpositions even when revisiting the same sampling sites.

(c) In addition to the aforementioned, the location of sampling stations plays a major role in WQ monitoring, particularly of artificial lakes. One of the main differences between artificial and natural lakes is that artificial characteristically exhibit a trophic gradient as it may grade from eutrophic (in its upper reaches) to oligotrophic (close to the dam). As reservoirs lose nutrients (in particular P) through settling in a downstream direction, utilized training *in-situ* data of a WQ model may not reflect the actual WQ conditions of an artificial lake. Based on this rationale, water samples from reservoirs is proposed to be collected from at least two (2) stations with optimal quantity the three (3) ones. The first one should be near to the point where each river/stream drains into the lake, the second one in the middle of the lake (as is now the case with the studied 50 lakes) and the third one close to the dam.

(d) It is reported that water transparency in artificial lakes is notably influenced by non-algal sources of turbidity (Canfield and Bachmann, 1981). Moreover, Lind (1986) documented that the use of Chl-*a* to estimate Secchi depth is inappropriate for waters where even moderate amounts of non-algal turbidity are present; this presumption surely affects hereby results, since the $\text{Secchi}_{\text{artificial}}$ model developed herein, employs the logarithmic transformation of Chl-*a*.

(e) The implementation of the DOS1 atmospheric correction method has not been validated in order to assure that atmosphere biases have been completely removed. Given the high number of utilized satellite imageries, though and the absence of atmospheric measurements, this

method was considered as adequate since it is widely used by the EO community.

(f) One more limitation concerning hereby attempted lake WQ monitoring through RS techniques, is the dependence on good climatic conditions. More particularly, a quite large number of satellite images with dates matching those of *in-situ* measurements, could not be used due to atmospheric effects. This fact induced the loss of a significant amount of field data, which otherwise could have contributed to the development of even stonger WQ models.

(g) Employment of RS techniques is also restricted by the utilized sensor 's temporal resolution or revisit time. In the context of this thesis, this feature has surely prevented a more effective lake WQ monitoring for those cases when the frequency desired was greater than the revisit capacity of Landsat sensors. Hence, it is further suggested that the EKBY's WQ monitoring field trips will be planned to coincide with the dates of Landsat and/or Sentinel 2 sensors overpass over the studied lakes. In this way, more *in-situ* data with matching satellite images can be employed in WQ models' validation.

(h) Concerning the distinction between land and water throughout the satellite imageries, no prior classification has been conducted. WQ models have been applied to lake surfaces accrued from lake shapefiles, hence there is the possibility that some pixels, while covering land, were defined as water; thus, may have prevented WQ parameters' quantification with higher accuracy.

Objective 2: Explore the spatio-temporal transferability of Landsat-developed WQ models across sensors; initially across Sentinel 2 and then across multi-sensor image data (Landsat 7 ETM+, 8 OLI and Sentinel 2 MSI). The transferability was tested along the National Lake Network Monitoring of Greece (WFD) and concerns the sampling campaigns of 2018.

Conclusion:

In general, herewith developed WQ models are proposed to employ principally Landsat images; however, the employment of Sentinel 2 data potentially produces reliable enough results with some (not significant) deviations from both corresponding Landsat-derived and *in-situ* reference lake WQ. Furthermore, the employment of multi-sensor (Landsat/ Sentinel 2) image data offered some improvement on a case by case basis while joint use facilitates in those cases the performance of high-frequency time series analyses. This effort highlighted a match between the corresponding spectral bands of Landsat and Sentinel 2 sensors; however, the slight inferiority of

Sentinel 2 images indicated the existence of differences in the final recorded radiometric values.

Scientific contribution:
As it has been already mentioned, the transferability of published WQ models across different sensors has been poorly examined, due to inherent differences (radiometrical, orbital, spatial, spectral). Towards strengthening the facilitation of WQ and trophic status monitoring across Greek lakes, employment of Sentinel 2 and multi-sensor image data (Landsat 7 ETM; Landsat 8; Sentinel 2 MSI) permits the integration among existing and historical missions while contributing to long-term time series data collection. In this way, the smoothly and well-operating performance of hereby WQ models will facilitate the multi-temporal lake WQ analyses, supporting further integrated lake management in the framework of national environmental policy.

Limitation concerning the WQ models' application on Sentinel-2 images:
WQ empirical models were developed based on Landsat-7 ETM+ and 8 OLI images; hence it is expected to be affected by the corresponding spectral composition and perform better when employing Landsat rather than Sentinel 2 reflectance.

Limitations concerning the combined use of Landsat and Sentinel-2 images:

(a) Mandanici and Bitelli (2016) documented that Sentinel-2 MSI band 8A (vegetation red edge) is the optimal option from the radiometric point of view when Sentinel-2 images are associated with Landsat 8 ones. Instead, MSI band 8 (NIR) is highly recommended for a joint use with older Landsat series, such as Landsat 5. However, in this thesis, the utilization of the Sentinel 2 B08 (NIR) band to match with B4 (L7 ETM+) and B5 (L8 OLI), may constituted an obstacle in achieving better and more accurate WQ quantifications when employing Sentinel 2 data.

(b) Another factor that plays a major role in the combination of Landsat and Sentinel 2 data are the residual effects of water specular reflections. Those are usually derived from the different zenith and azimuth angles and spacecraft altitude of the different sensors.

Objective 3:
An examination of the influence of different atmospheric correction methods to WQ models' performance after employing differently-atmospherically corrected SR values. The harmonization between the different SR products is based on the development of corrected sensor-specific (Landsat/ Sentinel 2) WQ models by utilizing dataset of 2018. Analysis and validation processes are

performed across the National Lake Network Monitoring of Greece (WFD) and the comparison of results is based on the *in-situ* WQ data of years 2019 and 2020.

Conclusion:
Double employment (dataset of 2018) of differently atmospheric corrected reflectance values (DOS1, LaSRC, LEDAPS and Sen2Cor) in WQ models indicated the DOS1 as the most effective method for the quantification of lake WQ elements in almost all cases and for all sensors (Landsat/Sentinel 2). In this way, empirically-developed WQ models, were proved to be applicable to both archived and future Landsat and Sentinel 2 image data embedded in GEE platform, exempting the user from the time-consuming pre-processing procedure. Furthermore, sensor-corrected models provided better assessment accuracy on a case by case basis depending on each water constituent's behavior with Secchi models (general, natural, artificial) illustrating the highest improvement, followed by TP and Chl-*a* models. WQ models' evaluation of years 2019 and 2020, indicated the sufficiency of GEE public data for mapping Chl-*a*, Secchi depth and TP concentrations in a large geographical region and particularly at a national scale (Greece).

Scientific contribution:
This is the first attempt, to the author's knowledge, to facilitate the quantification of spatiotemporal lake WQ across the Greek Lake Monitoring Network of WFD, by using multi-sensor reflectance values retrieved from GEE platform. Moreover, this research demonstrated the WQ models' temporal stability when employ SR retrieved from the GEE platform, offering scientists and Greek competent authorities the opportunity to exploit this massive warehouse of satellite data combined with the on-going WFD application.
All in all, exploitation of GEE image data promotes the long-term, near real-time, national-scale lake WQ and trophic status monitoring by mapping long-term WQ trends in less time and fine spatial resolution.

Limitations: Satellite imageries from Landsat and Sentinel 2 sensors, manually downloaded, have been subjected to different preprocessing procedures (DOS1) compared to corresponding, embedded in GEE platform (LaSRC, LEDAPS and Sen2Cor). Different atmospheric correction methods make harmonization of those differences essential to eventually generate reflectance values that can be comparable and combined. Otherwise, RS scientists are strongly recommended to apply the same AC algorithms as those adopted in GEE platform in cases where manually-derived satellite products need to be compared or/and combined with GEE-integrated images.

Objective 4: Assess WQ models' performance in a distinct category of optically complex Case-2 waters, oligotrophic Trichonis and Amvrakia lakes.

Conclusion:
Chl-*a*, Secchi and TP quantitative models have been empirically built based on a wide range of concentrations measured throughout 50 Greek lakes, representing almost all trophic status classes. However, their application in oligotrophic Trichonis and Amvrakia lakes was ineffective, confirming in absolute agreement with the relevant literature, the underachievement of universal WQ models at the lowest chlorophyll concentrations (oligotrophic waterbodies) and at cases where the optical contribution is non-algal and inorganic (e.g. sediments). In this way, the necessity of the development of special oligotrophic algorithms for a more efficient and comprehensive lake management is underlined. A more accurate WQ assessment in oligotrophic bodies requires more appropriate satellite bands (blue-to-green ratios), more and narrower wavelengths specifically in NIR spectrum, additional water samplings during different time periods and the refinement of AC processors. Additional to those implications, a preceding and thorough research on phytoplankton community composition, Chl-*a* distributions, particles and CDOM is required for a more accurate WQ monitoring of oligotrophic waterbodies along Greece.

Scientific contribution:
The unique contribution of this objective lies in fact of making the final decision on whether national WQ models support adequately the perpetual WQ monitoring of Greek oligotrophic lakes or special oligotrophic algorithms should be developed. Based on the above, this research facilitates the refinement of lake WQ monitoring in Greece by laying the foundation stone of further discriminating Case-2 Greek waterbodies into distinct optical water types (OWT). Hence, the increase of WQ assessment's accuracy per OWT is promoted and in particular of oligotrophic lakes by providing the background knowledge required.

Limitations:

a) The extremely low measurements of Chl-*a* concentrations in Trichonis and Amvrakia lakes; those are usually connected with decreased turbidity (lack of suspended particles) and by extension with decreased energy received by the satellite sensor.
b) Oligotrophic lake WQ monitoring by using Landsat (7 ETM+/8 OLI) sensors may be problematic. Sensors delivering images with more and narrower bands in the red-NIR spectral region may perform better in oligotrophic waterbodies. Furthermore, Chl-*a* and turbidity are

documented to be better quantified with sensors having red-edge band (680-710 nm) such as Sentinel 2, Sentinel 3 and Rapideye (Pizani and Maillard, 2022). Additionally, the alternative proposed for oligotrophic waters is the development of blue-to-green ratio algorithms (O'Reilly and Werdell, 2019).

c) Water samples initially from Trichonis/Amvrakia lakes and then throughout the WFD lake network have been collected by different institutions's staff, more probably using different methodology. Furthermore, the determination of Chl-*a* and nutrients' concentrations has been conducted employing different laboratory techniques, equipment and calibrations in different facilities, e.g. EKBY and HCMR. Those differences generate errors and biases concerning their intercomparison.

Objective 5: Model WQ of oligotrophic Trichonis lake by assessing WQ key elements (Chl-*a*, nutrient concentrations and CDOM absorption at 420 nm) through satellite RS. Trichonis is the largest freshwater lake of Greece while the available *in-situ* and satellite data concern years 2013 and 2014.

Conclusion: The ability of Landsat 8 OLI imagery was proven limited concerning the establishment of WQ models in the oligotrophic Trichonis lake. Furthermore, observed weak correlations among *in-situ* and satellite data, particularly in autumn and summer, may also be due to the lake turnover effect. When the equalization of the thermal gradient in the lake induces mixing of surface and bottom waters, remote monitoring is made difficult due to instability (Barrett and Frazier, 2016). Further research is required towards the investigation of more water parameters or using sensors of different spatial and geometrical analysis in order to be able to compare the outcomes among all different cases.

Scientific contribution:
Even though early results demonstrated the vulnerability of the Landsat 8 imagery to precisely determine certain WQ components in an inland oligotrophic body, it is generally accepted that those models may initially increase the knowledge of Trichonis lake's WQ and then be utilized as warning indicators of its WQ deterioration.

Limitations:

(b) The main factors disabling WQ assessment in Trichonis lake are the extremely low constituents' concentrations and the lack of any value differentiation among the sampling stations. Water samplings have

been conducted in autumn and summer seasons; additional water samplings should be made during different time periods concerning specific mixing boundaries (surface-bottom waters) in order to investigate whether the feasibility of remote monitoring increases. In case strong relationships are found, this may help improve prediction capabilities by providing researchers with bounded time periods.

(c) Another significant factor is the lack of any prior knowledge about regional phytoplankton community composition, namely whether the sampling stations are diatom- or cyanobacteria- dominated. When assessing the retrieval accuracy of satellite observations, phytoplankton community composition should be considered as spatial and temporal variations in composition may result in uncertainty in the inference of bloom severity. Binding et al. (2019) documented that if satellite Chl-*a* algorithms are calibrated primarily on one-species-cyanobacteria-dominated conditions, it may lead to significant uncertainty in derived phytoplankton biomass within mixed bloom assemblages or blooms dominated by other cyanobacteria. This may lead to potential misrepresentation of bloom severity while such uncertainty is particularly important for long term trend analysis.

5.1 Future work

Monitoring of lake WQ is a key priority topic, in terms of national environmental policy, which, ideally, should operate continuously while yielding updated and accurate results. During the accomplishment of this PhD thesis arose obstacles and particular factors hindering the achievement of more precise outcomes; on the other hand, in this way future pathways were configured towards the direction of managing an ever-enhanced and succesful WQ assessment throughout Greek lakes based on EO applications.

Overall, studies have concluded that there are atmospheric correction (AC) methods that perform better when employed for the assessment of certain WQ parameters compared to others. Since the DOS1 atmospheric correction method has not been validated in the context of this thesis, one of priority future actions is the application of alternative atmospheric correction methods with principle goal the exploration of their wider effect on WQ models' predictive ability and the selection of the most optimal one. A robust statistical analysis of how effective a AC method is, requires ground truth Rrs values, collected using a hand-held or shipborne spectroradiometer, to compare and evaluate AC corrected pixels. With a much larger dataset of observations on the optical properties, it will be possible to separate the match-ups by water type, giving further information and possible targeting of

water bodies with specific AC processors. AC methods' intercomparison will manage to imply the most optimal AC algorithm and even distinguish the most appropriate one not only for each WQ constituent but also for each OWT (optical water type).

Based on the rationale of connecting suitable AC method with respective OWT, the classification of hereby-studied 50 lakes in distinct OWTs is considered necessary. In general, lakes and reservoirs may present WQ variability not only across the study area (Greece) but also among regions and within some water bodies. The aim of this classification is to ensure that every OWT will be linked to specific bio-optical conditions in order to reflect the dominance of individual or group WQ elements' concentrations. The most commonly used inland water OWT classification system has been developed by Spyrakos et al. (2018) and is based on freely shared *in-situ* lake hyperspectral data, across global range. More particularly, resampling of OWT classes spectra to Landsat and Sentinel 2 spectra- except for WQ constituents' concentrations as was herewith performed- and the detection of their inter similarities would facilitate their OWT classification in an efficient and accurate manner. The use of OWTs contributes to the development of WQ algorithms for optically complex waters while in parallel helps to choose an efficient AC processor for a specific region of interest.

Moreover, authors intend to expand their research horizons and experiment with non-linear methods concerning the WQ models' development. Empirical methods and in particular machine learning techniques are able to handle complicated non-linear relationships which typically characterize the WQ remote sensing data. Considering the non-parametric nature of the hereby-utilized data, a sensitivity analysis among artificial neural network, genetic algorithms and support vector machine might be valuable for a more accurate assessment of WQ constituents. Additionally, the indication of the most optimal non-linear method and its comparison with herewith WQ linear models would overall contribute to a deeper knowledge of WQ parameters' behaviour.

Concerning the hereby delivered lake WQ models, one of the most significant tasks that will definitely be accomplished in the near future, is their integration in the GEE platform. Separate Chl-*a*, Secchi and Total Phosphorus models developed for all, natural and artificial lakes accompanied by the respective TSI equations for the trophic state assessment will be unified in an automated tool for near real time, national-scale WQ and trophic status monitoring.

As far as the utilized satellite sensors are concerned, most of the inland water applications developed for WQ monitoring and management have been established by using images from multispectral and medium to high spatial resolution satellites (e.g., Landsat, Sentinel-2-MSI). Nowadays, spaceborne

hyperspectral sensors have been attracting a significant attention since they offer simultaneous collection of hundreds of narrow spectral bands, facilitating the retrieval of WQ parameters. Hence, experimentation with hyperspectral image data such as PRISMA, DESIS, HISUI and EnMAP constitutes one more future task which will be conducted concerning the WQ parameters' estimation in Greek lakes. Especially EnMAP is proposed as a suitable choice for remote estimation of Case-2 WQ properties as it has a spatial resolution of 30 m, 4 days of minimum temporal resolution and its products consist of 224 bands.

Above all, it should be emphasized that the most significant precondition in order all of the above-mentioned future tasks be successfully achieved, is the on-going operation of WFD in Greece.