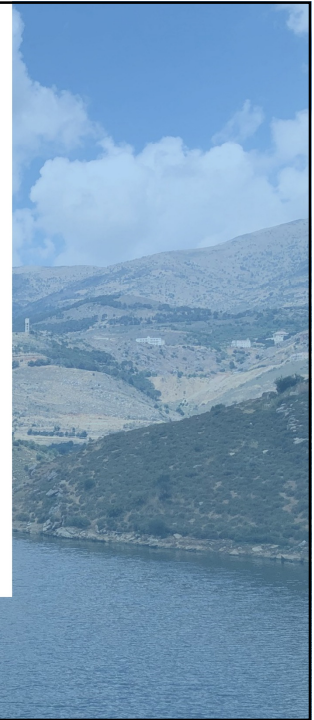


Predicting Water quality Variability in a Mediterranean Hypereutrophic Monomictic Reservoir using Sentinel-2 MSI: The Importance of Model Functional Form

Ibrahim Alameddine

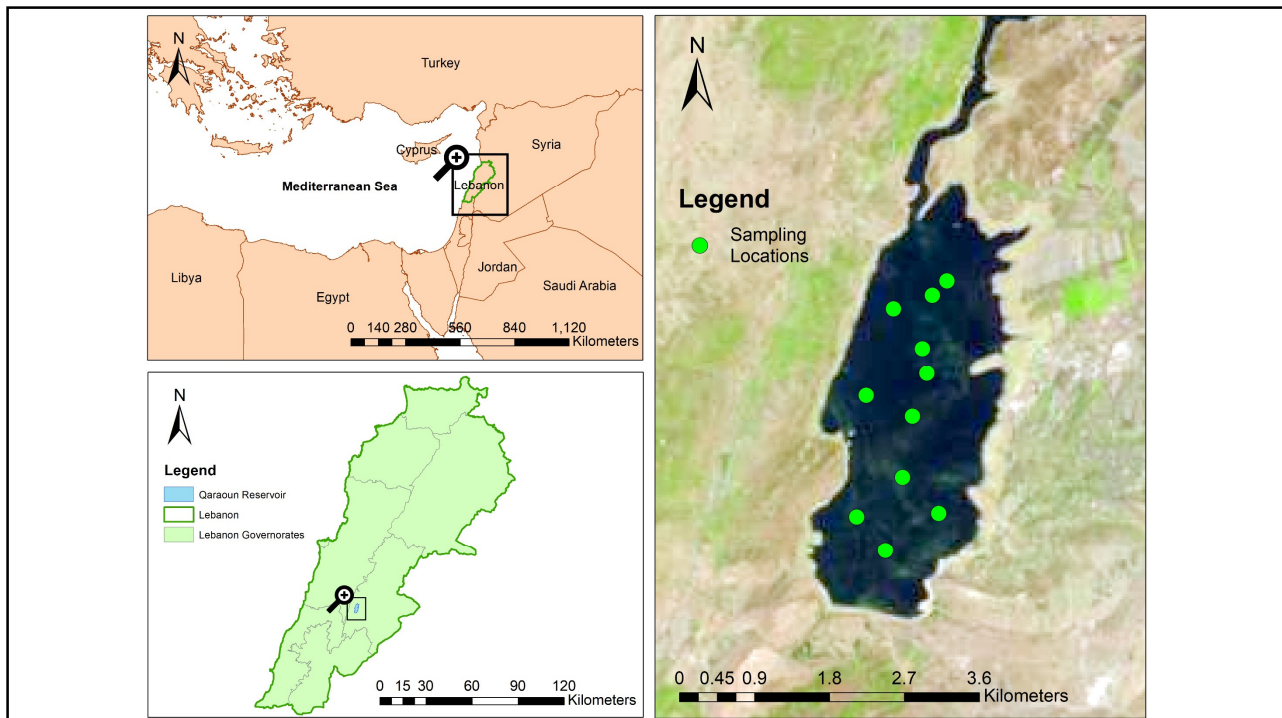
Mohamad Abbas

April 2021



Qaraoun Reservoir

- Qaraoun Reservoir (33° 34' N, 35° 52' E) is a **hypereutrophic** reservoir located in the semi-arid **Bekaa valley in Lebanon** along the **Eastern Mediterranean** coastline
 - **Surface area** varies between **4 and 10 Km²**
 - **Maximum depth** of the reservoir is **45 m**, while the median depth varies between **10 and 20 m** depending on the season



Qaraoun Reservoir: Water Quality

- The reservoir's water quality deteriorated over the past 2 decades as a result of the discharge of **untreated domestic and industrial wastewater** along with **agricultural runoff**
- High level of external nutrient loading reaching Qaraoun have accelerated the **eutrophication** process and promoted the dominance of the cyanobacteria → **hypereutrophic**
- Harmful Algal Blooms dominate throughout the growing season
 - *Microcystis aeruginosa*



Aims of the Study

1. Assess the potential of using **Sentinel 2 MSI** to predict and evaluate the spatio-temporal variability in the water quality of the Qaraoun Reservoir
2. Assess the performance and transferability of existing water quality models calibrated to the lake based on Landsat 7 and 8 surface reflectance data towards predicting Chlorophyll-a (**Chl-a**), total suspended solids (**TSS**), Secchi Disk Depth (**SDD**), and Phycocyanin (**PC**) levels from **Sentinel 2 MSI** data
3. Develop Sentinel 2 specific models for Chl-a, TSS, SDD and PC
4. Test different model functional forms:
 - Multiple linear regressions (**MR**)
 - Multivariate adaptive regression splines (**MARS**)
 - Support vector regressions (**SVR**)

Field Sampling Data

- 23 in situ water sampling campaigns between **July 10, 2018** and **October 3, 2019**
- During each campaign 5 to 11 monitoring stations visited
- 133 of a total of 153 water quality samples collected coincided with a **Sentinel 2 MSI** overpass with minimal cloud cover
 - SDD measured with a 20-cm diameter disk in the field
 - Grab-samples collected from the water surface and analyzed in the lab for Chl-a, TSS, and PC concentrations



Sentinel 2 MSI

- Sentinel 2 MSI surface reflectance images downloaded from the European Space Agency (ESA) Copernicus Open Access Hub website <https://scihub.copernicus.eu/dhus/#/home>
- Sentinel 2 MSI images were atmospherically corrected using the **Sen2Cor** atmospheric correction program
- The Sentinel 2 MSI images were processed using the Raster 3.0.12 package in the software R



Methods

- Transferability of Landsat model to Sentinel 2 assessed based on several previously calibrated Landsat-based models developed specifically for the Qaraoun Reservoir

Parameter	Satellite used	Model	Reference
Chlorophyll-a	Landsat 8	$Chl - a = \frac{\frac{Blue}{Red} \times NIR - 98}{0.75}$	Fadel et al. (2016)
Phycocyanin	Landsat 8	$PC = 109.63 \frac{NIR}{Red} + 17.873$	Sharaf et al. (2019)
Chlorophyll-a	Landsat 7+8	$\ln(Chl - a) = 4.7 + 1.1 \ln\left(\frac{NIR}{Red}\right) - 2.1 \left(\frac{Blue - Red}{Green}\right)$	Deutsch et al. (2018)
Total Suspended Matter	Landsat 7+8	$\ln(TSM) = 6.2 + 0.8 \ln\left(\frac{NIR}{Blue}\right) + 1.1 \ln(Red)$	
Secchi Disk Depth	Landsat 7+8	$\ln(SDD) = 0.4 + 0.5 \frac{Blue}{Red} - 22.5 Red$	

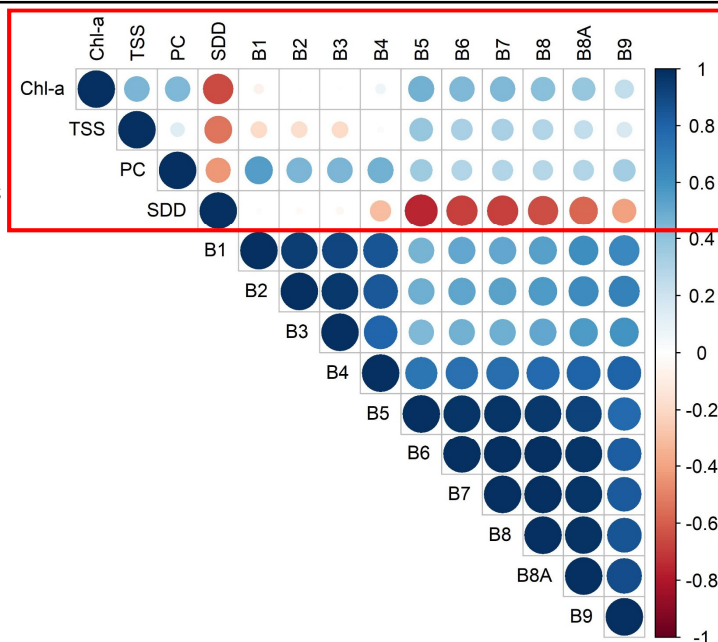
Methods

- Sentinel 2 specific models calibrated using:
 - Step-wise multiple linear regression
 - Support Vector Regression (SVR)
 - The radial basis function Kernel was used
 - Tuning parameters were the cost of constraints violation (cost) and the insensitive-loss function (epsilon), with a 10-Fold cross validation
 - Multivariate adaptive regression spline (MARS)
 - The tuning parameters were the degree of interaction (degree) and the maximum number of terms in the pruned model (nprune), with a 10-Fold cross validation

Results: Reservoir Water Quality

- Median Chl-a level was 97.64 $\mu\text{g/L}$ ($n = 125$, range = 6.94 – 674.85 $\mu\text{g/L}$, mean = 137.12 $\mu\text{g/L}$)
- Median TSS concentration was 16.33 mg/L ($n = 130$, range = 0 – 126 mg/L, mean = 20.74 mg/L)
- Median SDD was 1 m ($n = 92$, range = 0.5 – 2 m, mean = 1.12 m)
- PC concentrations ranged between 19.4 and 148.0 $\mu\text{g/L}$ ($n = 35$, median = 69.67 $\mu\text{g/L}$, mean = 70.42 $\mu\text{g/L}$)

- Correlation between all water quality parameters, except between TSS and PC
- Chl-a and TSS showed high correlation
- Chl-a and SDD field measurements were **negatively** correlated
- Chl-a and PC are **positively** correlated
- Chl-a and SDD correlated with the bands B5, B6, B7, B8, B8A
- TSS showed a correlation with the same bands but with lower levels
- PC showed a higher correlation with the bands B1, B2, B3 and B4



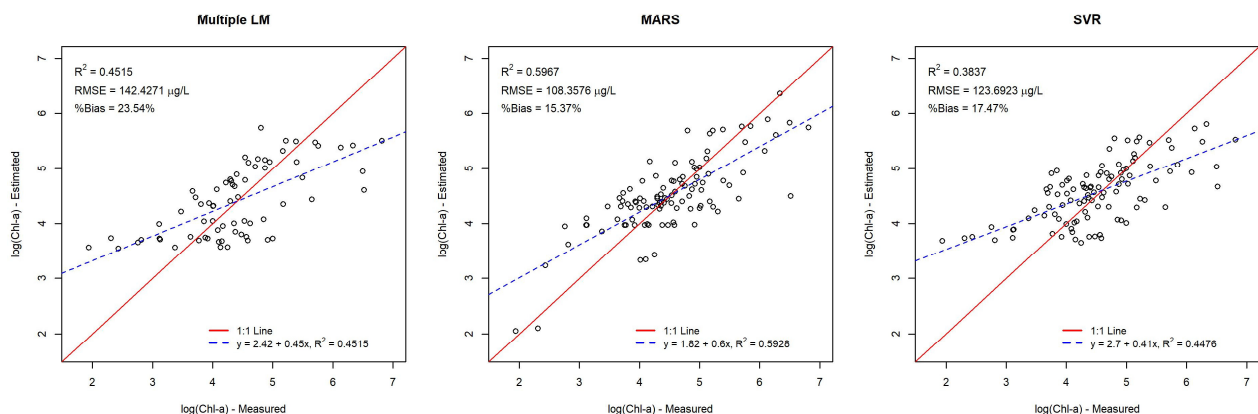
Spearman correlation between the different water quality parameters and Sentinel 2 MSI bands

Results: Water Quality Model Development

- Poor transferability of the Landsat models to Sentinel 2, with all models experiencing a significant drop in their predictive skill

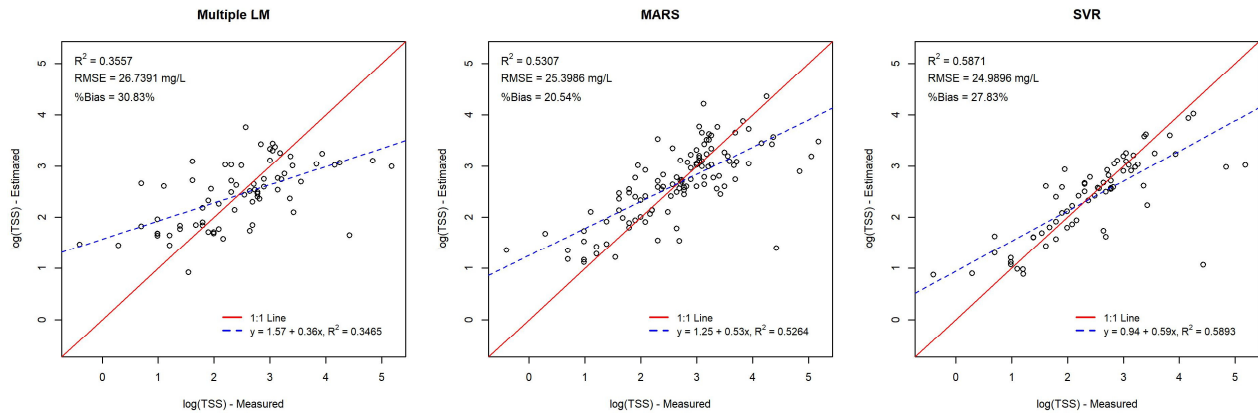
Parameter	Model	Drop in R ² (in %)
Chlorophyll-a	$Chl - a = \frac{\frac{Blue}{Red} \times NIR - 98}{0.75}$	88
Phycocyanin	$PC = 109.63 \frac{NIR}{Red} + 17.873$	95
Chlorophyll-a	$\ln(Chl - a) = 4.7 + 1.1 \ln\left(\frac{NIR}{Red}\right) - 2.1 \left(\frac{Blue - Red}{Green}\right)$	70
Total Suspended Matter	$\ln(TSM) = 6.2 + 0.8 \ln\left(\frac{NIR}{Blue}\right) + 1.1 \ln(Red)$	70
Secchi Disk Depth	$\ln(SDD) = 0.4 + 0.5 \frac{Blue}{Red} - 22.5 Red$	82

Results: Sentinel 2 Models – Chl-a



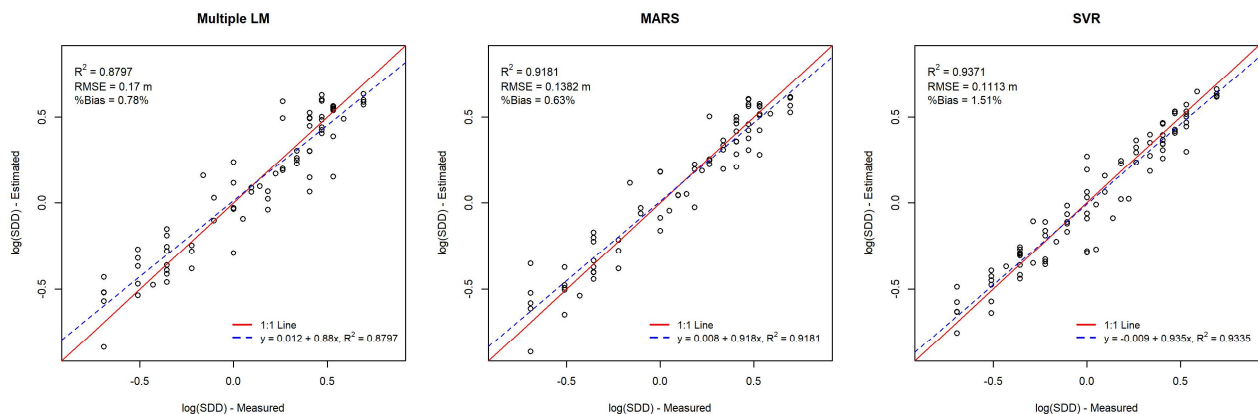
Selected bands were consistent across the three model functional forms: B4, B5 and B8A

Results: Sentinel 2 models – TSS



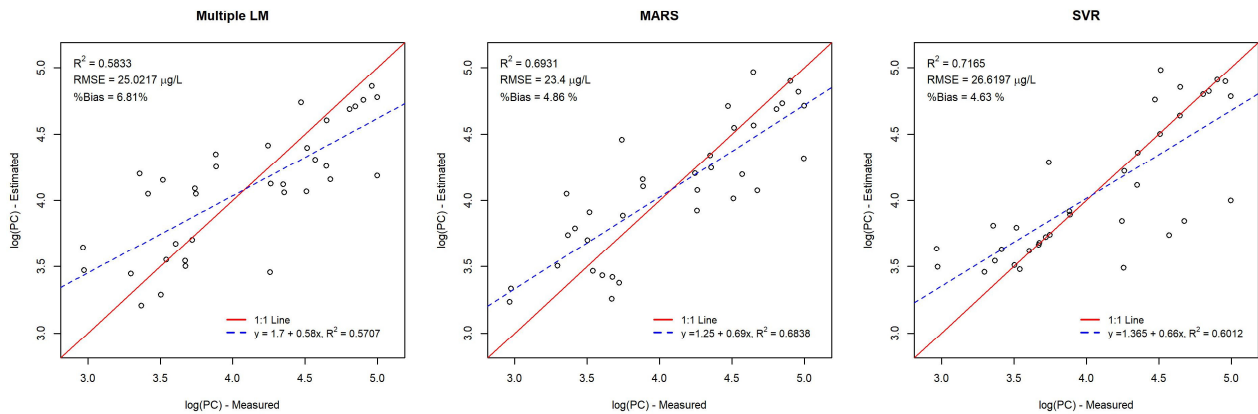
Significant differences in band selection across the three modeling techniques

Results: Sentinel 2 models – SDD



Selected bands were consistent across the three model functional forms: B2, B5, B6, B8 and Temp

Results: Sentinel 2 models – PC



Significant differences in band selection across the three modeling techniques



Conclusions

- Poor transferability between Landsat and Sentinel 2 based water quality models
- For Chl-a, the MARS model outperformed MR and SVR, with an R^2 of 60%
- For SDD, all three model forms (MR, MARS and SVR) had a high performance ($R^2 = 88\%$, $R^2 = 92\%$, and 94% respectively)
- The SVR-based models outperformed MR and MARS for TSS and PC ($R^2 = 59\%$, and 72% respectively)
- For Chl-a and SDD, all three models shared the same bands
- For TSS and PC, models diverged on model selection