Article

Not peer-reviewed version

Remote Sensing Imagery and Machine Learning-Based Methods for Quantifying Total Dissolved Solids and Total Suspended Solids Concentration in River Systems: A Case Study of the Colorado River Basin

Godson Ebenezer Adjovu *, Haroon Stephen, Sajjad Ahmad

Posted Date: 5 March 2026

doi: 10.20944/preprints202603.0403.v1

Keywords: ensemble models; machine learning (ML); monitoring; remote sensing (RS); total dissolved solids (TDS); total suspended solids (TSS); sustainability; standalone models

 doi:10.20944/preprints202603.0403.v1

Article

Remote Sensing Imagery and Machine Learning-Based Methods for Quantifying Total Dissolved Solids and Total Suspended Solids Concentration in River Systems: A Case Study of the Colorado River Basin

Godson Ebenezer Adjovu *, Haroon Stephen and Sajjad Ahmad

Department of Civil and Environmental Engineering and Construction, University of Nevada Las Vegas, Las Vegas, NV 89154, USA

* Correspondence: adjovg1@unlv.nevada.edu

Abstract

The Colorado River and its tributaries housed in the Colorado River Basin (CRB) are the primary source of water to the western United States and the Republic of Mexico. The river system is under intense stress due to prolonged drought and anthropogenic activities which have worsened its water quality. Total dissolved solids (TDS) and total suspended solids (TSS) are two water quality parameters (WQPs) that are crucial to the sustainability of the river system. These parameters are noted to have caused varied severity to the sustenance of the basin's water. Monitoring of these WQPs has been conventionally conducted using field and laboratory analysis which are cost and labor-intensive. This study utilized a novel method to effectively develop machine learning (ML) models to estimate TDS and TSS concentrations in the CRB by utilizing the potential of optically sensitive multispectral Sentinel 2 A/B Multispectral Scanners (MSI) and Landsat 8 Operational Land Imager (OLI) remote sensing (RS) data retrieved from the Google Earth Engine (GEE) and in situ measurements collected from 2013-2022. Several standalone models such as linear regressions (LR), ridge regressions (Ridge), lasso regressions (Lasso), and k-nearest neighbor (KNN), and ensemble methods including the gradient boosting machines (GBM), random forest (RF), adaptive boosting (AdaBoost), eXtreme gradient boosting (XGBoost), and bagging were applied for the accurate estimation of TDS and TSS. Results found ensemble models like the XGBoost as the most optimal model estimating TDS using images from both Sentinel-2 MSI and Landsat 8 OLI with performance on the external validation dataset derived as 0.99, 26.52 mg/L, and 19.19 mg/L, respectively for R^2, RMSE, and MAE for Sentinel-2 images. The XGBoost yielded R^2, RMSE, and MAE of 0.97, 35.82 mg/L, and 27.90 mg/L, respectively. The AdaBoost was found to be best model for TSS estimations with values of 0.92, 29.48 mg/L, and 24.64 mg/L, respectively, for R^2, RMSE, and MAE for the Sentinel-2 image on the external validation dataset. The RF model was found to be the optimal model for TSS estimations with the Landsat 8 OLI with reported R^2, RMSE, and MAE of 0.90, 32.80 mg/L, and 22.91 mg/L, respectively on the external validation dataset. These findings show great potential of utilizing RS data to produce cost-efficient spatiotemporal changes on the WQPs which has an important implication for the continuous management of the limited water resources. Further study should consider the effect of land use land cover, runoff, and other climatic data to understand the complexity and dynamics of these parameters on TDS and TSS in the river.

Keywords: ensemble models; machine learning (ML); monitoring; remote sensing (RS); total dissolved solids (TDS); total suspended solids (TSS); sustainability; standalone models

 doi:10.20944/preprints202603.0403.v1

1. Introduction

The Colorado River and its tributaries are housed in the about 657,000 sq. km area Colorado River Basin (CRB) located in the southwestern United States [1]. The Colorado River is a complex regulated river system ranked 6th out of 135 rivers in the US with lengths of more than 160.93 km. The river is joined by several tributaries with major ones as San Juan, Gunnison, White, Yuma, Virgin, Green, Little Colorado, Gila, and Delores Rivers [2]. The river houses several large reservoirs including Lake Powell and Lake Mead whose water levels are drastically declining due to significant drought, climate change, and increased water demand by the growing population [1].

The issues of drought and stress in the river systems are aggravated by increased water impairment caused by anthropogenic activities resulting from the negative impacts of urbanization. Water impairment has negative impact on the recreational and commercial activities of the water body [3,4]. Total dissolved solids (TDS) or Salinity and total suspended solids (TSS) or suspended sediment concentration (SSC) are two water quality parameters (WQPs) that degrade the quality of many water systems including the Colorado River.

TDS and TSS concentrations have a direct impact on the water use, the ecological health and ecosystem of the river basin and require monitoring to ensure the resilience and sustenance of the basin's water. High levels of TDS and TSS in the Colorado River affect the appearance, taste, and safety of the river's water for human consumption and support of the aquatic ecosystem [5–7]. Although TDS concentration is a secondary drinking water regulations with no enforcement, the US Environmental Protection Agency (EPA) limits TDS concentration in drinking water to be 500 mg/L [8–10]. TSS concentration is regulated differently by states. Some states use numerical or narrative criteria while others use both. Some states have no criteria. Those states that use numerical criterion have limited TSS in waterbodies have been limited to 30 to 158 mg/L [11]. The tolerance level of TSS concentration for bottom invertebrates and fishes in waterbodies is limited to 13 mg/L and 90 mg/L, respectively [10,11].

TDS concentration is most naturally occurring as significant amount of salts sediments in the river is said to have been deposited in the prehistoric marines. An estimated 55% of the TDS in the river is accounted for by natural sources including runoff and groundwater intrusion. Agricultural and irrigation activities account for 37% of the salt contribution while 8% is accounted for my evaporation, municipal, and municipal sources [12]. The river transports significant amounts of erodible salt originating from sedimentary rock and agricultural runoff as it flows from it headwaters to its terminus with the salt concentration increasing from about 50 mg/L to 850 mg/L as it crosses its US-Mexico border.

The US and Mexico signed an international water treaty in 1944 which allows the delivery of 1.85 billion m3 of water annually from the Colorado River to the Republic of Mexico in compliance with the Minutes 242 [13–15]. The water discharged to Mexico from the Morelos Dam must however comply with an additional increment in TDS of 115± 30 mg/L in addition to the average annual flow-weighted TDS concentration sets at the Imperial Dam (located at the California/Arizona boundary). The seven states within the basin namely California, Arizona, Nevada grouped as Lower Colorado River Basin (LCRB), Colorado, Utah, Wyoming, and New Mexico grouped as the Upper Colorado River Basin (UCRB) have adopted numerical limits for TDS concentrations for different locations namely at the Imperial, below Parker, and below Hoover Dams with concentrations (mg/L) of 879, 747, and 723 respectively [7,16–18]. The flow weighted average annual TDS concentrations at these numerical criterial locations have, however over the past years been considerably reduced drastically to values below 750 mg/L TDS levels since the early 2000s due to the salinity control programs implemented [17]. The implemented programs is reported to reduce the annual salt loads in the river to values exceeding 1.3 million tons with 100 mg/L reduction at the Hoover, Parker, and Imperial Dams [19].

Measures put in place to control the amount of salinity and associated effects of the salinity in the Colorado River is said to have been causing the US in excess of USD 300 million annually in

 doi:10.20944/preprints202603.0403.v1

economic damages [6,10,12,14,20,21] with the amount expected to rise to an estimated USD 471 million by 2025 [22].

Extensive sedimentation caused by increased load and concentration of TSS can reduce water clarity, reduce light penetration, and negatively impact the photosynthetic process of aquatic plants reducing the dissolved oxygen in the water. As a result, the entire food chain of the aquatic ecosystem is distributed, which has an impact on the wellbeing and biodiversity of their ecosystems. Increased in TSS in the water has the potential to increase water clogging in the gills of fishes and other aquatic in the river causing them to die prematurely [23–26]. Additionally, silt buildup in the river's reservoirs and water delivery infrastructure may result in reduced storage capacity and higher maintenance and operational expenses.

Monitoring, management, and assessment of TDS and TSS is therefore necessary to understand the changes in the water and to protect and restore the quality of the river's water in serving its use [27–29]. Monitoring and assessment of these WQPs are carried out using conventional method measurements such as grab sampling and laboratory analysis. These approaches are, however, cost and labor intensive and require more time to conduct. Conventional approaches of measuring these parameters are also limited in providing spatiotemporal assessment and real time monitoring of these WQPs [30–33]. Remote sensing (RS) approaches however provide synoptic, repetitive measurements, and cost-effective approach for monitoring these WQPs as compared to traditional field sampling methods [34]. Most RS sensed images are publicly available and come with no cost for developing models for WQP estimations. RS models offers a cost-effective tool for the quantification of several WQPs including TDS and TSS and have been widely used in several studies [35–38]. The accuracy of models developed from RS images depends on several factors including spectral, spatial, and temporal resolutions of the sensors as well as the mechanisms applied.

WQP signal detection by remote sensors is based on the optical characteristics of the parameter under study. WQPs are categorized as optically active or optically inactive based on their ability to directly influence the optical properties of water measured by the sensors. TSS is an optically active parameter meaning it directly impacts the optical properties of the water and can easily be detected [39]. Suspended particles in the water can interact with light to produce signal reflectance which can be sensed by the RS sensor. TSS also generates a strong backscatter effect that allows light to scatter in the water column, enabling it to be detected by RS sensors [40]. TDS on the other hand is a spectrally (optically) inactive parameter, and its determination is based on the fact that it has a strong correlation with other WQPs with which there may be co-variation [30]. Consequently, spectral signatures from images captured by RS sensors are therefore used in conjunction with conventional field measurement to establish various models including empirical, analytical, semi-empirical, and more recent novel machine learning (ML) or artificial intelligence (AI) models for the accurate estimations of WQPs [30,41–51].

ML is a form of AI which applies implicit algorithms to capture linear and non-linear relationships to solve unlike the conventional or physical models. They works by approximating functions to describes patterns in studied phenomenon [30,42,52,53]. Compared to traditional statistical approaches, ML models are more recent, recurrent, and gaining tremendous considerations in recent studies [54,55]. ML provides a better approach for solving complex and challenging problems without prior knowledge compared to physical modes. They provide efficient and intelligent extraction of WQPs with high dimensional spectra properties [56]. The recent advancement in RS analytics, ML, and cloud computing has provided a novel approach for monitoring these WQPs [57]. Integrated remote sensed ML models have emerged as promising, novel, and cost-effective estimation of both optically active and inactive WQPs including TSS and TDS to address the drawbacks associated with conventional monitoring techniques including limitation with spatiotemporal assessment on large scale [52]. ML models are categorized as standalone or single and ensemble models. Standalone models are single or individual models which are trained and used in independent estimations of WQPs. Some of the widely used standalone models include support vector machine (SVM) or regression (SVR), artificial neuron network (ANN),

 doi:10.20944/preprints202603.0403.v1

Linear regressors (LR), and K-Nearest Neighbor Model (KNN). Ensemble models on the other hand combine multiple single models to boost the accuracy and resilience of predictions. The ensemble models work by leveraging on the diversity of each single model and aggregation of their individual predictions through series of algorithms to obtain an overall precise and reliable predictions. Examples of ensemble models include bagging, gradient boosting machine (GBM) or regression (GB), stacking, adaptive boosting (AdaBoost), extra trees (ET), eXtreme gradient boosting (XGBoost), and random forest [21,58–60].

Researchers in [61] used ML and RS spectral indices (band combinations) to identify water quality index (WQI) in the Ebinur Lake Watershed in China. WQPs studied include TDS, various anions, and cations. The study utilized the particle swarm optimization (PSO)-SVR model and RS spectral indices of a hyperspectral RS data (collected using FieldSpec③3 ASD Spectroradiometer) including difference index (DI), ratio index (RI), and normalized difference index (NDI). Findings from the result showed the potential of using PSO-SVR model for accurate estimation of WQI with an obtained R^2 of 92%.

Researchers in [53] used Landsat 4 and 5 TM, Landsat 7 Enhanced Thematic Mapper Plus (ETM+), and Landsat-8 OLI in the US and a feed-forward neural network (FFNN), a cascade forward neural network (CFNN), and an extreme learning machine (ELM) to estimate the SSC along major fluvial systems, namely Mississippi and Missouri Rivers in the US. The findings show that ELM-based techniques have the potential to accurately predict SSC for the Landsat 4-5, 7, and 8 sensors with $R^2 > 90\%$ with minimal to no overfitting.

A study by [62] used Landsat-8 OLI RS data and ML algorithms including RF, SVM, XGBoost, AdaBoost, GB, ANN for predicting of Chlorophyll-a, turbidity, and TDS in the Lake Tana located in the Tropical Highlands of Ethiopia. Results demonstrated the capability of utilizing ML models for accurate prediction of these WQPs with XGBoost found to be the best predictor for Chlorophyll-a with R^2 of 78%, the RF model performed best for the TDS and turbidity estimation with R^2 of 79% and 80%, respectively. Other researchers in [39] used ML-based strategy (including SVR, neural network (NN), and the RF) and Sentinel-2 RS data to estimate optically inactive WQPs namely total nitrogen (TN), total phosphorous (TP), and chemical oxygen demand (COD) in small urban lake located in the City of Tianjin, China. The results obtained demonstrate the potential of using ML-RS based model to accurately estimate optically inactive WQPs with obtained $R^2 > 85\%$ for all the 3 WQPs.

Researchers in [63] also used the Sentinel-2 based spectral indices and empirical regression models to estimate TDS concentrations in freshwater wetlands located in Guartinaja and Momil wetlands in Northern Colombia. Results obtained shows a possible opportunity of utilizing RS models for estimation of TDS with a derived normalized RMSE of < 10%.

Additionally, researchers in [40] used Sentinel-2 MSI and a semi-empirical model to quantify SSC in the Yangtze main stem located in China. Findings from the study revealed a great potential of utilizing satellite derived images for the estimation of SSC with reported $R^2 > 79\%$.

Although ML and RS are fast gaining attention by the research community, only few studies till date have attempted these applications in monitoring particularly in the retrieval of TDS with very sparse studies in North America.

This study therefore aims to leverage on the standalone and ensemble ML algorithms and multispectral RS images from Landsat 8 OLI and Sentinel-2 A/B MSI to estimate TDS and TSS concentration in the Colorado River System. This will aid in the revolutionization of novel techniques for the assessment of water quality to further aid in efforts in the monitoring, management, and sustenance of the limited river's water resource amid the current drought and ever-shifting environmental issues confronting the river system.

The goal of the study is to advance knowledge of water quality dynamics in this crucial area by offering insightful information for efficient water resource management, conservation initiatives, and well-informed decision-making.

 doi:10.20944/preprints202603.0403.v1

This study hypothesized that spectral signatures from RS multispectral images coupled with standalone and ensemble ML models can be used to develop RS-ML models for the accurate estimations of TDS and TSS in the Colorado River. The study consequently addresses the following research questions:

1. How can RS multispectral images be utilized in developing robust ML models for the estimation of TDS and TSS in the CRB?
2. How do the spatial resolutions of multispectral RS images affect its potential in effectively estimating TDS and TSS concentration in the CRB?
3. How accurate are RS-ML based models in the estimation of TDS and TSS concentrations in the CRB?

2. Materials and Methods

2.1. Study Area

The Colorado River and its tributaries are housed in the CRB located in the western United States as shown in Figure 1. The river is approximately 2318 km long [64] and runs from its headwaters which are located in rocky mountainous and characterized by copious snowfall, high altitude, and alpine habitant in Colorado and Wyoming to the terminus in the Gulf of California and influenced by interactions of several natural processes, human activities and water management strategies [2]. The entire basin is about 657,000 sq. km [1,64] and is divided into upper and lower CRB. The UCB encompasses the Colorado, Wyoming, Utah, and New Mexico while the LCRB encompasses Arizona, Nevada, California, and the Gila River between Arizona and New Mexico [16]. The Colorado River has several tributaries joining it as at different locations with the major tributaries as Gunnison, White, Yampa, San Juan Delores, Little Colorado, Gila, and Virgin Rivers [2]. The river with tributaries is source of water to about 24 billion m^2 of farmland in the United States and Republic of Mexico with US lands constituting more than 90%. About 35 million people and 3.3 million people, respectively, in the US and Republic of Mexico depend on the river and its tributaries for their water needs. The river systems also provide about 16 million jobs with economic support of more than USD 1.4 trillion to the US [16,65,66].

The basin is known for its known for its breathtaking scenery, which ranges from snow-capped mountainous terrains to the arid desert regions. is a well-known and ecotourism-rich region because of its famous natural wonders including the Grand Canyon, Glen Canyon, and other national parks. The basin is predominantly a semi-arid region with varying elevations to up > 4267 m above mean sea level. The area receives an average annual precipitation of 10.16-152.40 cm, with the areas in the UCRB mountainous areas receiving the most precipitation. Most areas in the LCRB are predominantly dry due to the orographic barriers and prolonged high-pressure weather cells. The area receives widespread temperatures with values ranging from below freezing point during wintertime particularly in the UCRB to over 37.77 °C during the summertime in the LCRB. This precipitation and the hydrologic cycle are driven by the intense evaporation rate in the area. Precipitation is evaporated from the snowpacks, soils, open water surface, crops, and other vegetations through evapotranspiration processes [67]. Climate change and changes in natural processes in the river is said to have led to the reduction in streamflow particularly in the UCRB by 10% with 6 to 14% uncertainty reductions [68].

The CRB houses several reservoirs and dams for water storage and hydroelectric power generation. The main stem Colorado River has 15 dams whose combined storage is estimated to be about 72 billion m^3 [69] and several major reservoirs including Lake Mead and Lake Powell. Lake Mead is the largest reservoir of the river located in the Mohave Desert specifically within the Arizona-Nevada boundary formed after the construction of Hoover Dam. Lake Powell is formed by the Glen Canyon Dam and its located between southern Utah and northern Arizona [2,18,22,70]. Both reservoirs store up to four times of the basin's historic average yearly flow [71]. Releases from these reservoirs are regulated and may result in significant changes to the downstream water quantity and

 doi:10.20944/preprints202603.0403.v1

quality. The water allocation from the basin is based on several acts including Colorado River Compact of 1922, the 1944 Water Treaty, and the UCRB Compact of 1948, and the US Supreme Court rulings [72]. The area is, however, faced with severe drought coupled with water quality impairment which are induced by natural factors such as climate change geology, natural variability and mineral discoveries, and anthropogenic activities like excessive demand for water resources, the vast agricultural activities, mining activities, wastewater discharges, and energy development [16]. Known water quality issues in the basin includes TDS and TSS. Monitoring, management, and understanding the occurrence and levels of these WQPs is therefore key to ensuring the sustainability of the limited water resources and improving the overall ecosystem of the basin. The broaden and vast landscape, the intricacy, and the possible difficulty in sampling the entire Colorado River water system makes it ideal to apply RS-ML based model for the estimation of these WQPs.

Figure 1. Colorado River Basin with States and Urbanized Areas Created with the ESRI ArcGIS Pro Software.

2.2. Data Collection and Description

This subsection describes the data collected and used for this study. The study used in situ or field data obtained from the United States Geological Survey (USGS) and the RS data obtained from the USGS and European Space Agency (ESA) website respectively, for Landsat 8 OLI and Sentinel 2 A/B MSI. These datasets were used to develop robust and accurate RS-ML models for the TDS and TSS concentrations.

2.2.1. Field or In Situ Data

WQP monitoring efforts are often carried out by professionals and volunteers using several methods including automated situ sensors [28].

This study utilizes WQPs collected by the USGS. TDS and TSS data were retrieved from the USGS database. USGS carried out and maintained field and or laboratory analysis of several WQPs including TDS and TSS and are available for public access through its website [73]. In situ water quality measurements for the period of 1 January 2013, to 15 July 2022, were obtained for this study. January 2013 was to have data availability to correspond to the Landsat 8 OLI sensor which was launched in 2013, and July 2022 was to the period of the analysis of the data. To retrieve the data for TDS and TSS, steps presented in the flow chart presented in Figure 2 were used by inputting the USGS parameter codes of 70301 for TDS and 80154 for TSS.

 doi:10.20944/preprints202603.0403.v1

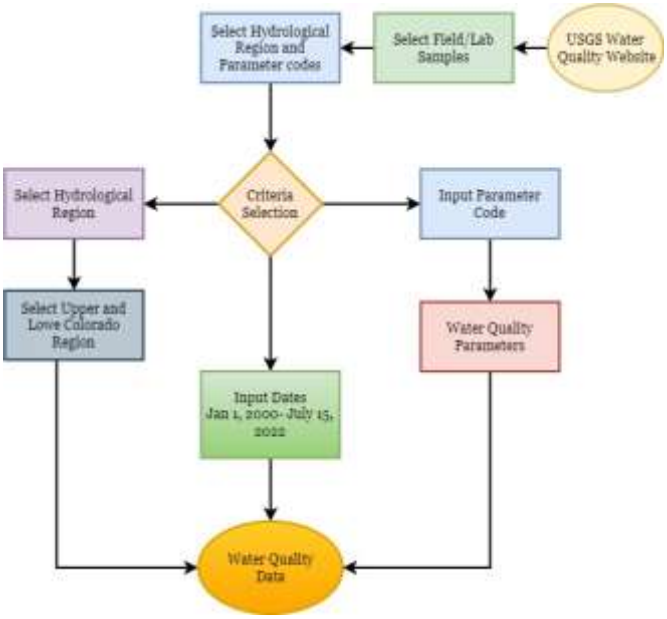


Figure 2. Schematic Procedures for the Retrieval of TDS and TSS from the USGS Data Repository.

2.2.2. Multispectral Satellite Data

Landsat 8 OLI and Sentinel-2 A/B MSI images were utilized for this study. The Landsat 8 OLI sensor was developed by a collaborative effort of the USGS and the National Aeronautics and Space Administration (NASA) and launched in February 2013. The Sentinel 2 A/B MSI sensor on the other hand was developed by the ESA and launched on June 23, 2015, and March 7, 2017, respectively. [30,74–77]. These sensors have been selected for the study because of their application and usefulness in the assessment of WQPs in freshwater systems including rivers. Images from these sensors are publicly available and free for users. Although these images have moderate resolutions compared to SPOT, IKONOS, and ASTER images, they have more advantages including their extensive temporal record. Despite having a spatial resolution of 2.5-20 m, SPOT images have a revisit time that is over twice as long as Landsat 8 OLI and more than five times as long as Sentinel-2 images [78].

The Landsat 8 OLI used in this study has resolutions ranging from 15 to 30 m resolution and a cycle of 16 days. The long temporal resolution of 16 days may pose limitations on intra-seasonal monitoring of the WQPs [32]. Consequently, additional sensor of a shorter temporal resolution of 10 days for Sentinel 2 A and 2 B(and combined 5 days) and a high spatial resolution of 10 m resolution which provides the opportunity to accurately and carefully discriminate between the water and the land surface frequently was used [18,39]. Although Sentinel 2 A/B has a higher spatial resolution compared to Landsat 8 OLI, the signal-to-noise ratio (S/N) of the Landsat 8 OLI visible bands offers a comparable or better radiometric quality to the Sentinel-2 images [59] making it an additional preferred choice for this study.

Researchers have estimated TDS and TSS concentrations using Landsat 8 OLI and reported $R^2 >$ 50% [37,38,79,80]. Researchers in [56,63] also applied Sentinel 2A/B MSI to quantify TSS concentration in waterbodies and reported $R^2 > 70\%$. Although these studies have been carried out with these sensors only a few compared the comparisons of these sensors and there exist gaps in the utilization of novel ML applications in the estimation of these WQPs particularly TDS in freshwater systems. The spectral and spatial resolutions of the Landsat 8 OLI and Sentinel-2 A/B MSI are presented in Table 1 and Table 2 with comparisons between these two sensors presented in Table 3.

Table 1. Band Characterization of the Landsat 8 OLI Bands. Adapted from [81–83].

Band	S/N	Wavelength (μm)	Spatial resolution (m)
Band 1 – Coastal aerosol	284	0.430 – 0.450	30

 doi:10.20944/preprints202603.0403.v1

Band 2 – Blue	321	0.450 – 0.510	30
Band 3 – Green	223	0.530 -0.590	30
Band 4 – Red	113	0.640 – 0.670	30
Band 5 – Near Infrared (NIR)	45	0.850 – 0.880	30
Band 6 – Short Wave Infrared (SWIR) 1	10.1	1.570 – 1.650	30
Band 7 – SWIR 2	7.4	2.110 – 2.290	30

Table 2. Band Characterization of the Sentinel 2-MSI A/B Sensor. Adapted from [78,83,84].

Band	S/N	Central wavelength (μm)	Bandwidth (nm)	Res. (m)
B2-Blue	102	0.490	65	10
B3-Green	79	0.560	35	
B4-Red	45	0.665	30	
B8-NIR	20	0.842	115	
B5-Red Edge 1	45	0.705	15	20
B6-Red Edge 2	34	0.740	15	
B7-Red Edge 3	26	0.783	20	
B8A-Red Edge 4	16	0.865	20	
B11-SWIR 1	2.8	0.161	90	
B12-SWIR 2	2.2	0.219	180	

Table 3. Summary Comparison of Landsat 8 OLI and Sentinel-2 Sensors. Adapted from [78,85,86].

Characteristics	Satellite Sensor Landsat 8 OLI	Satellite Sensor Sentinel-2 A/B MSI
Sensor	OLI	MSI
Launch date	Feb 11, 2013	June 23, 2015 (Sentinel-2A) March 7, 2017 (Sentinel-2B)
Temporal resolution (days)	16	10 (5 for the combined Sentinel A and B)
Spatial resolution (m)	15,30, 60 (band specific)	10, 20, and 60 (band specific)
Radiometric resolution (bits)	12	12
Equatorial crossing times (Local time)	10:00 a.m. ± 15 min	10:30 a.m.
Swath	185	290
Altitude (km)	705	786
Inclination (◦)	98.22	98.62

2.3. Preprocessing

The search results retained 38,254 data points. The data was then taken through a series of cleaning to refine the data for the study. The data was first cleansed using the parameter code for TDS (70301) by deleting rows with blanks, rows with M (presence verified but not quantified), rows with < signs (since the exact value is unknown), rows with E (estimated values), and then rows with A (average values). After this initial deletion, other agencies included in the data AZ014 and USEPA

 doi:10.20944/preprints202603.0403.v1

which have very little data for their stations were subsequently deleted. This led to a total of 102 datasets of 5,311.

The refined data were further filtered using the sediment code (80154). At the end of filtering for the sediment data, the total amount of data was 2,232 with 49 stations. Stations located at sections where the river is narrower are eliminated to avoid contamination by land surface reflectance. An additional seven stations with their data were randomly selected and added to ensure each state on the CRB has some representative values for the analysis. In all a total of 56 stations with 2,525 and 2,265 datapoints respectively for TDS and TSS were proposed following the cleaning process which is presented in Figure 3. The USGS stations are presented in Figure 4.

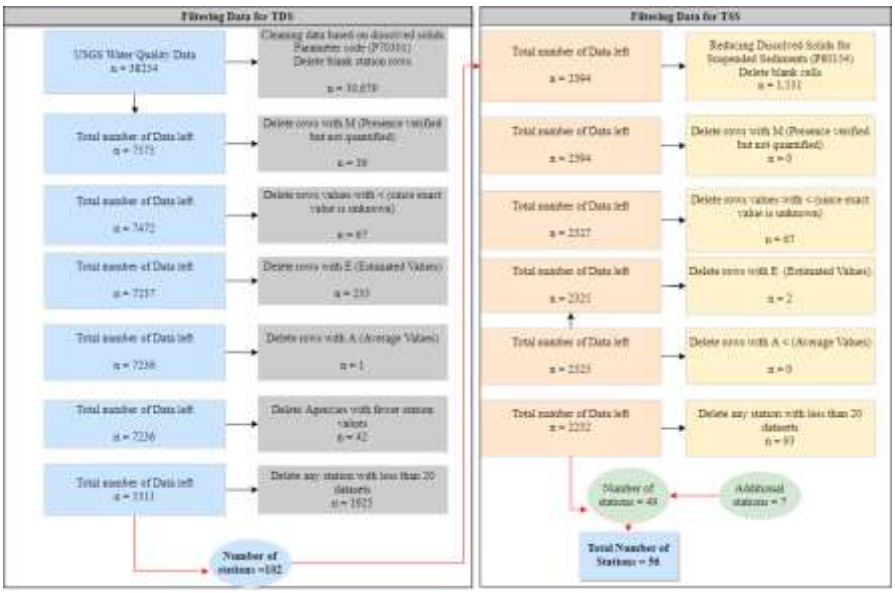


Figure 3. Data Filtering Process for USGS Datasets.

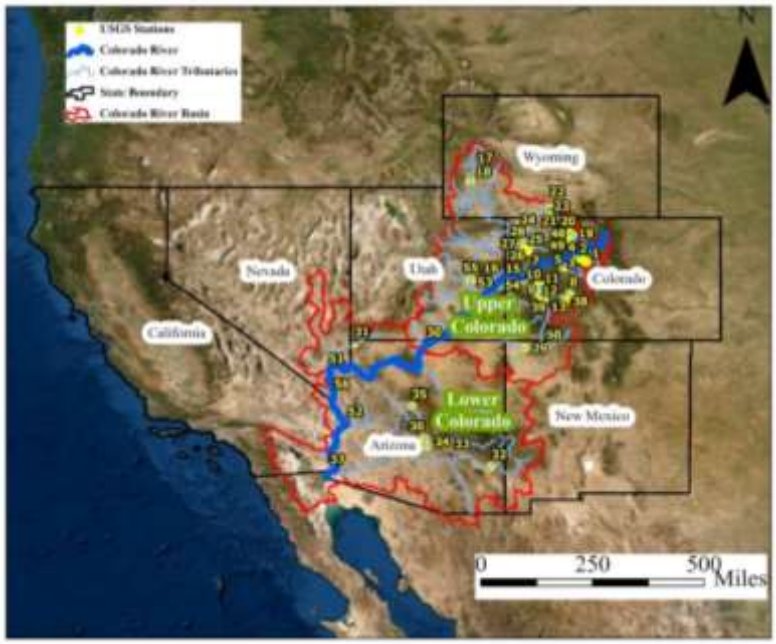


Figure 4. Distribution of Selected USGS Stations for Water Quality Data on the CRB River System.

The RS image collections for Landsat 8 OLI and Sentinel 2 A/B MSI were imported into Google Earth Engine (GEE). The GEE is a powerful cloud-based platform to access and use a comprehensive

 doi:10.20944/preprints202603.0403.v1

collection of satellite images including Landsat 8 OLI and Sentinel-2 A/B MSI images. The platform was launched in 2010 and uses JavaScript language. The GEE combines data storage, computing capacity, and a variety of analysis tools for rapid and effective processing of large RS datasets and has been used for monitoring and assessment of environmental phenomena such as water quality, wetland, and forest fire monitoring among others [75,87–91]. In this study, the coordinates of Level-2 atmospheric correlated surface reflectance data of the Landsat 8 OLI (available from April 2013) and Sentinel 2 A/B MSI (available from March 2017) were "called" into the GEE. The shapefile of the coordinates of the location of the USGS sampling stations was created in Google Earth Pro version 7.3.6.9345 and ArcGIS Pro version 3.1.0 software into the GEE. Using the point coordinates of the station and the images, stations within narrower sections of the river were eliminated to ensure only spectral signatures of water are returned reducing any possible contamination by values of land surface through visual inspections at each site for each satellite pass through the GEE platform. One pixel reflectance value was extracted for each station point. The time series of the point reflectance values were then exported as a .csv file to be matched with the water quality data.

The matching of the dataset was based on a 7-day offset between the two datasets. To ensure the spectral signals of the sensor captured the actual impact of light and water quality interactions, the dataset with WQPs sampled before the satellite overpasses used were matched based on the 7-day difference. This is important to ensure no significant occurrences occurred between the sampling efforts and the following satellite overpass. Studies have used varying days with days ranging from 0, ±1, to ±15 days offsets [92–97]. The 7-day offset was used to ensure that enough RS data and WQPs were retained, and the actual water characteristics were captured by the RS signals for further analysis. Due to the narrower river channels at sections of the river coupled with the 7-day limit to synchronize the field and RS data, the initial 56 stations were reduced to 33 stations for 10 m resolution and 16 for the 20 and 30 m resolution. Average values were computed for rows with images with more than one spectral value for a particular day to represent the value for that day. After the match, missing rows in the TSS data were filled using the linear interpolation technique. Interpolation allows for the estimation of values at unsampled dates for the specific station using information on earlier and succeeding dates to compute the missing value [98]. Accurate imputation results in estimators with good large-sample characteristics and reliable acceptable large-sample inferences [99]. Due to its simplicity and effectiveness in terms of computing, linear interpolation was chosen as the preferred interpolation method for this study [100,101].

The Landsat 8 OLI surface reflectance presents some negative and zero reflectance values which occurs as a result of error with the atmospheric correction and retrieval algorithms of the Landsat surface reflectance [102]. These values were replaced with a small negligible value of 0.0001 [103] to prevent possible errors which may arise in the computation of indices by way of division. It is important to add negligible values to ensure there is no general significant alteration of the scalability and biases to the computation performed [104].

Aside from performing the analysis on the spectral signatures of single bands, band combinations or indices that are sensitive to the WQPs and have been used in past studies were also computed and used. Separate indices utilized for TDS and TSS and used for this study are presented in Table 4 and Table 5.

Table 1. RS Indices Used in Estimating Concentrations of TDS and Salinity.

Index	Image/Data	Equation	Metrics	Study Area (Country/Region)	References
Salinity Index 1	Landsat 8 OLI	$\sqrt{R * B}$	R^2 = 0.72	Colorado River (USA)	[6]
			$R^2 \geq$ 0.72	Shatt al-Arab River (Iraq)	[38]
Salinity Index 2		$\sqrt{G * B}$	R^2 = 0.73	Colorado River (USA)	[6]

 doi:10.20944/preprints202603.0403.v1

	Landsat 8 OLI		R^2> 0.79	Shatt al-Arab River (Iraq)	[38]
Salinity Index 3	Landsat 8 OLI	$\sqrt{R^2 + G^2}$	R^2 = 0.72	Colorado River (USA)	[6]
Salinity Index 4	Landsat 8 OLI	$\sqrt{R^2 + G^2 + NIR^2}$	R^2 = 0.70	Colorado River (USA)	[6]
Salinity Index 5	Landsat 8 OLI	$G * R$	R^2 = 0.73	Colorado River (USA)	[6]
			R^2 ≥ 0.44	Shatt al-Arab River (Iraq)	[38]
	Landsat 5 TM	$G * R$	R^2 > 0.65	Coastal surface water (Bangladesh)	[105]
	Landsat 8 OLI		R^2 > 0.49		
Salinity Index 6	Landsat 8 OLI	$\sqrt{G^2 * NIR^2}$	R^2 = 0.71	Colorado River (USA)	[6]
Salinity Index 7	Landsat 5 TM	$\frac{B}{G}$	R^2 > 0.06	Coastal surface water (Bangladesh)	[105]
	Landsat 8 OLI		R^2> 0.46		
-	ASTER	$\frac{B5}{B3} * B7$	R^2 > 0.57	Qaroun Lake (Egypt)	[106]
NDSI	Landsat 8 OLI	$\frac{R - NIR}{R + NIR}$	R^2 > 0.51	Mekong Delta (Vietnam)	[107]
TDS 1		$G + R + R1$	r ≥ 0.79		
TDS 2		$\frac{G + R - R1}{G + R + R1}$	r ≥ 0.00		
TDS 3	Sentinel-2 MSI	$\frac{R1 - B}{R1 + B}$	r ≥ 0.52	Guartinaja and Momil wetlands (Colombia)	[63]
TDS 4		$R1 + R - B$	r ≥ 0.79		
TDS 5		$R1 + R$	r ≥ 0.79		
TDS 6		$\frac{R + R1}{SWIR1 + SWIR2}$	r ≥ 0.00		
TDS 7		$\frac{R * R1}{SWIR2}$	r ≥ 0.00		

R = red band; B = blue band; G =green band; NIR=near-infrared band; R1= Red_edge1 band; SWIR1= Shortwave Infrared 1 band, and SWIR2 = Shortwave Infrared 2 band. B1, B2, B3, B4,… stands for bands 1, 2,3,4, and so on.

Table 2. RS Indices Used in Estimating Concentrations of TSS, SS, and Suspended Material.

Index	Image/Data	Equation	Metrics	Study Area (Country/Region)	References
NSMI	Landsat 7 ETM+	$\frac{R + G - B}{R + G + B}$	$R^2 \geq 0.96$	Lake Mead (USA)	[5]
	Landsat 8 OLI		-	Cabo Rojo (Puerto Rico)	[108]

 doi:10.20944/preprints202603.0403.v1

				$R^2 > 0.51$	Barito Delta (Indonesia)	[109]
				$R^2 > 0.70$	Dams (South Africa)	[37]
					Lake Mead (USA)	[5]
			$\frac{B - \text{NIR}}{\text{B} + \text{NIR}}$	-		
				-	Cabo Rojo (Puerto Rico)	[108]
NDSSI	Landsat ETM+	7		$R^2 = 0.01$	Barito Delta (Indonesia)	[109]
					Mississippi River	
				$R^2 > 0.66$	Lake Pontchartrain (USA)	[110]
			$\frac{G}{\text{B}}$	-	Lake Mead (USA)	[5]
					Cabo Rojo (Puerto Rico)	[108]
BR	Landsat ETM+	7		$R^2 = 0.05$	Barito Delta (Indonesia)	[109]
WSRI	Landsat 8 OLI		$1 - (\frac{SWIR - B}{R})$	$R^2 > 0.70$	Dams (South Africa)	[37]
EGRI	Landsat 8 OLI		$\frac{G}{\text{C} + B}$	$R^2 > 0.70$	Dams (South Africa)	[37]
NDVI	Landsat 8 OLI		$\frac{NIR - R}{NIR + R}$	$R^2 > 0.70$	Dams (South Africa)	[37]

2.4. Summary Data Processing and Analysis

Descriptive analysis such as the measures of tendency minimum, mean, and maximum) was used to gain an overall understanding of the trend of the river's water quality [64]. Additionally, the Pearson correlation analysis was performed to measure the strength of relationships between the WQPs and spectral signatures (bands and their combinations) [111–113]. The strength of the correlation is quantified by the correlation coefficient (r) whose values range from -1 to 1, with values nearing ±1 indicating the existence of a strong relationship. Negative values mean there is a negative relationship between the variables [38,114,115].

2.5. Machine Learning Modeling

Before performing the ML modeling, the z-score was used in the identification of outliers due to the spikes in the WQPs contributed by the tributaries of the river. Rows with WQPs identified as outliers were used to ensure the optimal performance of the model. There were spikes in the WQPs noted in some of the dates contributed by the tributaries which have the potential to impact the overall performance of the ML. The z-score, which measures the deviation of the data point from the mean value in relation to the spread of the data was used in the identification and removal of outliers. The Z-score is given by Equation 1 [116,117]. Values can be positive or negative with positive values indicating that values are above the mean and negative indicating values below the mean. The absolute z-threshold of this study was set to 3 in accordance to [117] which categorized data points with z-scores greater than 3 as exceedingly atypical value.

$$Z = |\frac{X - \mu}{\sigma}| \quad (1),$$

where Z is the Z-score, μ is the mean of the dataset and σ is the standard deviation.

Following the preprocessing and analysis, the datasets and the stations initially identified for this study outlined in Section 2.3 and presented in Figure 4 reduced drastically from 2525 to values as presented in Tables 6.

 doi:10.20944/preprints202603.0403.v1

Table 3. Summary of Data after Matching and Filtering of RS and Field Data.

WQP	Sentinel 2	Landsat 8 OLI
TDS	220	581
TSS	142	492

In this study both standalone and ensemble ML models were used for accurate estimation of TDS and TSS concentrations. Standalone or single models used in this study to analyze RS and WQPs for the water quality estimation and monitoring includes LR and the KNN [21,58,59,118]. To offer a potential improvement in the accuracy of standalone models in certain instances, studies combine multiple individual ML algorithms called ensemble models to make predictions. This study utilized some of these ensemble models, namely GBM, AdaBoost, RF, XGBoost, and Bagging [21,62,118–120]. These models are described in detail in subsections 2.5.1 and 2.5.2.

2.5.1. Standalone ML Models

i. LR is the most widely used algorithm used in prediction. LR are simple but powerful algorithms used to establish linear relationships between variables used in many decision-making processes.LR models are grouped into simple or MLR based on the number of explanatory variables in the model. A simple uses a single explanatory and response variable while a MLR utilizes multiple predictors for the prediction of the response variable [58,121]. LR models attempt to minimize the sum of squared errors between the true and estimated values by computing the coefficients of the predictor variables. There are several forms of LR based on how the coefficients are treated. This includes the ordinary least square regression (OLR), Least Absolute Shrinkage and Selection Operator (lasso), and the ridge. The OLR minimizes the sum of the squared errors by assuming that the model's errors are normally distributed and have a fixed variance. The ridge and Lasso are used to address multicollinearity while preventing overfitting in LR models. Lasso applies a penalty term to the LR through a process of regularization to shrink the less significant coefficients to zero. The ridge on the other hand applies a penalty term to mitigate the effects of multicollinearity by shrinking all coefficients and lowering their variability to avoid overfitting. The LR models are therefore controlled by a hyperparameter called regularization hyper parameter α which works to prevent the model from overfitting [58,122,123]. The three forms of LR, namely OLR, ridge, and LASSO as described were used to estimate the WQPs in this study.

ii. KNN model is a non-parametric technique which carries out predictions based on how similar the k nearest data points is to one another. It identifies the k nearest datasets to the target during the data training and uses their values to categorize the input features to make prediction. Hyperparameters for the KNN model include the number of neighbors (which influences the number of data points to be used for prediction), and the distance metrics which are used assess the degree of similarity between the data points. The known form of distance metrics used are the Manhattan and the Euclidean distance. The KNN model is highly influenced by the voting function and the distance selected for the optimal value [119,124,125].

2.5.2. Ensemble ML Models

i. GBM ML model: builds prediction by combining several standalone ML models in sequence with each iterative model learning from the previous weak models to enhance the overall reliability and accuracy of the model. Hyperparameters associated with the GBM include the number of estimators used in determining the number of weak learners or decision trees included in the ensemble, the learning rate which regulates the contribution of each weak learner to the ensemble, and the maximum tree depth which restricts the maximum depth of each decision tree needed to control model complexity and to prevent overfitting [124,126,127].

ii. AdaBoost ML model: developed by Freund and Schapire combines several weaker learners which are trained in sequence with each model learning from the errors of the previous models

 doi:10.20944/preprints202603.0403.v1

through interactive to form a stronger model based on the weighted majority voting of the accuracy of weaker learners. Hyperparameters of the AdaBoost includes the number of estimators, learning rate, and the base estimators to use [62,126,128,129].

iii. RF: works by adding several individual decision trees to produce a robust and accurate model for the estimation of a response variable. The ultimate prediction is obtained by averaging the predictions of all separate trees, each of which is trained on a different random subset of the data [130,131]. Each of the decision trees in this algorithm is created using a randomly selected set of training data and features. This randomization improves variability in the decision trees further aiding in minimizing model overfitting. The RF is associated with various hyperparameters including the number of decision trees and the maximum tree depth which restricts the maximum depth of each tree needed to control the model complexity to prevent overfitting [130–133].

iv. XGBoost ML model: developed in 2016 at the University of Washington, operates by adding multiple decision trees to weaker learners produced as a result of errors from previous decision trees to enhance the overall performance of the model estimations [60,126,134]. The XGBoost shares similar features and benefits to the RF algorithm in estimation performance, interpretability, and simplicity. The main difference existing between the two algorithms is that the RF builds decisions trees independently while those trees are built sequentially in the case of the XGBoost [134]. The hyperparameters associated with the XGBoost includes the maximum tree depth which controls overfitting, the learning rate, the regularization parameter (Ɣ), and the number of estimators that determines the number of weak learners in the ensemble model [135].

v. Bagging or bootstrapping is an ensemble ML model that aggregates the predictions of several models to enhance the overall performance of the model. It uses the bootstrap sampling mechanisms to derive random subsets of the training data of the original model for an accurate and robust estimation with minimal overfittings. Bagging works to lower the variance in the prediction model by controlling the bias-variance tradeoffs associated with a model. Final predictions for the bagging are based on the mean outcome of all the decision trees created on the training subsets during the bootstrapping steps. Hyperparameter for the bagging ML includes the number of estimators, the base estimator i.e., decision trees, and the maximum number of features for each estimator [126,133,136–138].

2.5.3. Model Feature and Hyperparameter Optimization

The data for 2020 was set aside (this year was chosen to ensure enough data was returned to validate the model on both satellite images) as 'unseen data' for external validation before performing feature selection and the hyperparameter optimization. The external validation processes are necessary to ensure unbiased evaluation and ensure model outputs are true in any real-life setting [23,119,136,139–142]. The rest of the data was randomly split using the 80%/20% proportion with the 80% used in the model training and 20% to model testing [55,143,144].

Selection of optimal features

Feature optimization was performed to select the best features for each model using the training dataset and each base model. The powerful Recursive Feature Elimination with the Cross-Validation (RFECV) approach was utilized for the selection of optimal features for each model. The RFECV technique works with cross-validation to eliminate weaker features from all model features consisting of the spectral bands and respective indices, subsequently generating the most optimal features required to reduce the model overfitting, speed up the model learning algorithm, improve the model interpretability, and enhancing the general performance of the ML models [62].

This study uses k= 5 folds for an extensive hyperparameter tuning needed to minimize any issue of overfitting and biases. Several studies have utilized varying folds including the five folds in the hypertuning approaches [145–149]. The k = 10 folds have also been widely used in several studies to produce robust and accurate results with no biases [52,150,151]. The five folds were used in this study due to the limited data size and to ensure a reduction in computation time. Researchers in [148] found the usage of the k = 5-fold as being successful in the estimation of WQI in a timely and cost-efficient

 doi:10.20944/preprints202603.0403.v1

manner. The cross-validation process is carried out through a series of steps which are i.) Separate the training dataset into independent, equal-sized k-groups or folds with no replacement such that each observation is utilized precisely once for training and validation, ii.) train the model using the k-1 fold and use the rest to validate the model, iii,) repeat step ii. till each part is used to achieve the k number of performance metrics, and iv.) the final performance is obtained by computing the average of the k-performance metrics [55,143,144,150,152,153].

The bands and their combination as described in Section 2.2.2 for Sentinel-2 MSI and Landsat 8 OLI images were used as initial features for the feature optimization. Initial features proposed and used for retrieving TSS and TDS includes the spectral signatures of the 10 and 20 m resolutions Sentinel-2 images which includes Band 2, Band 3, Band 4, Band 8, representing the B, G, R, and NIR bands, respectively, and the 30 m resolution Landsat 8 OLI image Bands 1-7, representing the C, B, G, R, NIR, SWIR, and SWIR bands, respectively. In addition to the spectral bands, the band combinations as outlined in Tables 4 and 5 which include the TDS and salinity indices (i.e., TDS indices 1 to 7 and Salinity Indices 1 to 7, NDSI) for TDS estimations from the Sentinel-2 images were also considered. Estimations of TDS from Landsat 8 OLI did not utilize the TDS indices outlined in Table 4 due to the absences of R1 bands in the Landsat 8 OLI images.

Estimation of TSS also makes use of band combinations including NDSSI, NSMI, BR, NDVI, EGRI, and WSRI for Landsat 8 OLI as presented in Table 5. Estimations of the TSS with Sentinel-2 images also involve band combinations as used for the Landsat 8 OLI except for EGRI which requires the C band which is not present in the Sentinel-2 image catalog. This study therefore proposed and utilized 15 and 25 features for the estimations of the TSS and TDS, respectively using the Sentinel-2 images, and 13 and 15 features, respectively for TSS and TDS estimations using the Landsat 8 OLI.

Optimization of Model Hyperparameters

Model parameter optimization or tuning is essential in ML modeling since it is used to determine the optimal parameters needed to maximize the performance of the model. The grid search technique was used in the model parameter tuning to find the optimal values for the parameters for the model [62,136].

Grid search is a computer-intensive schematic technique used in fine-tuning hyperparameters. It works by thoroughly and iteratively evaluating the performance of the model based on the combination of a defined grid of potential values to find the optimal values for the best performance of the model [55,143,144]. The grid search was performed using the cross-validation mechanism with k folds.

ML Modeling and analysis were carried out using the Python Programming Language version 3.9.12 within the Jupyter Notebook web-based processing environment [154]. To utilize the Python Programming Language, several libraries NumPy, Pandas, and Matplotlib were imported. These libraries and software used in the study are summarized in Table 7. The specification of the main computer used in the study is as follows: Windows 10 Pro, 64-bit operating system, x64-based processor with Intel(R) Core (TM) i3-4150 CPU @ 3.50 GHz, with 16.0 GB RAM.

Table 4. Summary of ML Packages, Libraries, and Software Used in the Study.

Software/Package/Libraries	Brief Description	Version
Python	Python is a general-purpose programming language.	3.9.12
Pandas	Pandas is a Python package for data analysis and manipulation.	1.4.2
Matplotlib	Matplotlib is a library for creating visualizations in Python	3.5.1
Seaborn	Python visualization library that provides a high-level interface for statistical graphics.	0.11.2

 doi:10.20944/preprints202603.0403.v1

NumPy	NumPy is a package for scientific computing with Python.	1.21.5
Scikit-learn (sklearn)	Python module for ML built on top of SciPy.	1.0.2
SciPy	SciPy is an open-source software to carry out statistical, optimization, integration, and linear algebraic analysis, among others.	1.7.3
PerMetrics	Python library to carry out performance metrics of ML models.	1.3.1
XGBoost	An optimized distributed library for gradient boosting.	1.7.4

2.5.4. Model Evaluation Metrics

Several evaluation metrics have been widely used in past studies. This study utilized some of these metrics to assess the effectiveness of the models in estimating the WQPs such as coefficient of determination (R^2), Mean Absolute Error (MAE), Percent Mean Absolute Relative Error (PMARE), Root Mean Square Error (RMSE), the and Ratio of RMSE to Standard Deviation of the observed data (RSD), and scatter index (SI) as represented in Equations 2 to 8 [21,125,141,155–159]. The R^2 indicates the proportion of variations in the WQPs explained by RS data. Its values range from - ∞ to 1, with 1 being a great explanation [160]. Negative R^2 values are obtained in equations without constant terms and in situations where the model fit is worse than fitting a horizontal line because of model overfitting. Values of R^2 obtained in this situation cannot be interpreted as the square of the coefficient of correlation [161].

The MAE and PMARE estimate the mean absolute difference between the predicted and observed concentrations of the WQPs, while the RMSE calculates the square root of the mean squared variations of concentrations of the WQPs. Values of the MAE, RMSE, and RSD ranges 0 to ∞ with lower values closer to zero indicating an accurate model performance in estimating the WQPs [3,125,141,155–157]. The SI is a representation of the difference in the RSME in relation to the mean value of observation. Lower values of SI is an indication of accurate prediction [158]. The Nash–Sutcliffe Efficiency (NSE) coefficient and Percent Bias (PBIAS) shown in Equations 7 and 8, respectively, were also used to quantify the performance of the ML models in estimating the WQPs. The NSE and PBIAS are hydrological evaluation metrics used to assess the goodness of fit of the models. They have been employed here to assess the accuracy of the RS-ML based models. The NSE compares the relative magnitude of the residual variance to that of the observed values. The values of NSE ranges from − to 1 with values around 0.75 to 1.00 said to be very good. Values less than 0.4 are hover categorized as unsatisfactory [21,151,155,156]. The PBIAS is a metric which measures the probability of the model to either underestimate or overestimate the measured observed data. It is expressed as percentage with lower values closer to zero indicating model accurate estimations. Negative and positive PBIAS indicates over and underestimation, respectively [21,155,156].

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_{sim} - Y_{obs})^2}{\sum_{i=1}^{n}(Y_{obsmean} - Y_{obs})^2} \quad (2)$$

$$PMARE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{Y_{obs} - Y_{sim}}{Y_{obs}}\right| * 100 \quad (3)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|Y_{obs} - Y_{sim}| \quad (4)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_{obs} - Y_{sim})^2} \quad (5)$$

 doi:10.20944/preprints202603.0403.v1

$$SD = \frac{RMSE}{STD_{obs}} \frac{\sqrt{\sum_{i=1}^{n}(Y_{obs} - Y_{sim})^2}}{\sqrt{\sum_{i=1}^{n}(Y_{obs} - Y_{obs}mean)}} \times 100 \quad (6)$$

$$NSE = \frac{\sum_{i=1}^{n}(Y_{obs} - Y_{sim})^2}{\sum_{i=1}^{n}(Y_{obs} - Y_{obsmean})^2} \quad (7)$$

$$PBIAS = \frac{\sum_{i=1}^{n}(Y_{obs} - Y_{sim})}{\sum_{i=1}^{n}(Y_{obs})} x\ 100\%, \quad (8)$$

$$SI = \frac{\sqrt{\frac{1}{n}\sum_{i}^{n}((Y_{sim} - Y_{simmean}) - (Y_{obs} - Y_{obsmean}))^2}}{\sqrt{\frac{1}{n}\sum_{i}^{n}(Y_{obs})}} \quad (9)$$

where Y_{obs}, Y_{sim}, $Y_{obs}mean$, $Y_{obs}sim$ are the observed, modeled, mean of the observed values, and mean of predicted values respectively.

The evaluation metrics were implemented in the Python programming software version 3.9.12 in the Jupyter Notebook environment using standard Python codes from libraries and documentation such as Scikit Learn.

3. Results and Discussion

This section presents the results of the study as well as discussion of the obtained results. Results presented include the summary statistics to assess the temporal variability in the TDS and TSS as well as correction analysis, and the performance evaluations of the ML techniques in estimating these WQPs using Sentinel-2 and Landsat 8 OLI images.

3.1. Summary Statistical Analysis

3.1.1. Summary Analysis of WQPs

Descriptive statistics for TDS and TSS levels obtained in the 56 stations (before elimination due to the width of the river and satellite overpass) in the river for the period under study are performed for the WQPs. Sample size varies from year to year owing to the varying frequency of sampling for the various stations sampled as described in Section 2.2.1.

The statistical summary shows considerable variability of TDS concentration with values ranging from 23 to 17,200 mg/L obtained in June 2017 and July 2021, respectively. Higher values of TDS concentrations were found in the tributaries of the river compared to the main channel stations.

For example, the USGS station 9180000 located at the Dolores River Near Cisco, UT recorded the highest TDS of 17,200 mg/L under the period of study. This tributary river drains about 11,862 km2 area and flows about 388 km can be traced to San Juan Mountains, a well-known location for trout fishing [162], the activities of the river watershed may be a contributing source to this spike in TDS [64]. Studies have corroborated the increase in the mean TDS contribution from some tributaries compared to the main channel [17].

The huge variability in the TDS in the river system is largely due to contributions from the tributaries, the geological processes underlying the basin via which the river runs from its headwaters to the terminus, and seasonal variabilities. The activities in both the UCRB and LCRB have also been noted to be a contributing factor to the variability in TDS in the river system. There basin contains several reservoirs diversions, and dams that collect and store water for several uses including agriculture, drinking water intake, and the generation of electricity. Impoundment of the water by these structures changes the natural flow of water in the river system which may impact TDS concentrations along the river. Additionally, treatment processes such as water softening may also influence the level of TDS levels in the river [64]. For example, residents in Las Vegas Valley employ ion-exchange water softeners composed of salt brines of KCl and NaCl ions to treat their water withdrawn from Lake Mead located on the river system. The natural TDS level in the lake is estimated

 doi:10.20944/preprints202603.0403.v1

to be more than 600 mg/L [10,21,64,72] which is higher than the EPA guideline of 500 mg/L in drinking water [15,17] hence the need for treatment. The regenerant salts are then discharged into the wastewater treatment plants and ultimately discharged into Lake Mead raising the level of TDS in the river [72].

Analysis of the summary of TSS shows a much higher variability in the river with values ranging from 1 to 69,200 mg/L with average values typically above the numerical limit of 30 mg/L to 158 mg/L set by states with TSS. Typical mean values are also above the mean upper tolerance level of TSS of about 13 mg/L for bottom invertebrates and 90 mg/L for fish [11]. Changes in the huge variability in TSS along the river system could be attributed to variations in the watershed activities such as mining, dredging, constructions, agriculture (livestock grazing, irrigation, and other related-activities), and natural processes including climate change (resulting in monsoon rainfall and runoff), soil and river bank erosions [10,16,25,34,64].

The highest TSS of 69,200 mg/L was recorded in October 2017 at the USGS Station 9367540 located at the San Juan River near Fruitland, NM which drains about 20,590 km2 area. The spike in the TSS in this area could be due to the soil, geological compositions, land use characteristics, and changes in the stream channel of the San Juan River Basin which overlaps four corner areas of AZ, UT, CO, and NM with a drainage area of 64,577 km2. Cañon Largo, a main drainage to the Blanco Canyon Subbasin of the San Juan River Basin is noted to be one of the major contributors of salinity and suspended sediments to the San Juan River watershed [163].

In summary, increases in TDS and TSS variations in the river systems could be because of the prolonged drought in the river systems resulting in intense evaporation in known large reservoirs on the river system. Additional contributing factors to the large variability could be due to seasonal variations in streamflow, energy and mineral extraction, groundwater discharges, rainfall events, snowmelt, abiotic and biotic interactions, and anthropogenic activities including irrigation and releases from untreated wastewater treatment plants [16,66,70,72,164].

3.2. ML Model Assessment and Analysis

The surface reflectance of each spectral band was extracted from the Sentinel-2 and Landsat 8 OLI images through the GEE platform. The time series of the band spectral signature was exported as a csv and used for further analysis. The exported band spectral signature matched with the WQPs based on the 7-day offset as detailed in Section 2.3. The spectral indices are described in Table 4 and Table 5.

3.2.1. Model Feature Selection and Model Hyperparameter Optimization

Feature Selection

The RFECV was used to select the most optimal features for each model as discussed in Section 2.5.3 employing the base models [62]. The feature selection resulted in varying amounts of optimal features for each model as outlined in Tables 8 and 9 for TDS and TSS, respectively. All the models had a reduction in the number of features for TDS estimations from the 25 and 13 features for the Sentinel-2 and Landsat 8 OLI images, respectively, as presented in Tables 8 and 9. Similarly, there were varying reduction in the number of features for the TSS estimations from the initial 15 and 13 features for the Sentinel-2 and Landsat 8 OLI images, respectively, except for the case of the Lasso, RF, and Bagging models used on the Sentinel-2 images as shown in Table 8-9.

The features eliminated in the feature selection process signify their almost no or negligible importance on the overall performance of the models. Each of the models in the tables shows varying sets of features which is an indication the features extracted from the RS images obtained from the two sensors are not only impacted by their correlations with TDS and TSS but also by the underlying operations of the different ML models [62]. It is, however, not established if the number of optimal features has a direct link to the overall estimation capability of the ML model.

 doi:10.20944/preprints202603.0403.v1

Table 5. Optimal Features for each of the ML Models for Sentinel 2 Images.

WQP	ML model	Optimal features	No. of features
TDS	LR	NIR, R4, SWIR 1, SWIR 2, TDS 2, TDS 3, TDS 6, TDS 7, Salinity Index 2, Salinity Index 6, Salinity Index 7, NDSI	12
	Lasso	R4, SWIR 2, TDS 2, TDS 3, TDS 6, Salinity Index 6, Salinity Index 7, NDSI	8
	Ridge	R4, SWIR 2, TDS 2, TDS 3, TDS 6, Salinity Index 6, Salinity Index 7, NDSI	8
	KNN	NIR, R4, SWIR 1, SWIR 2, TDS 2, TDS 3, TDS 6, Salinity Index 6, Salinity Index 7, NDSI	10
	RF	NIR, R4, SWIR 1, SWIR 2, TDS 2, TDS 3, TDS 6, Salinity Index 6, Salinity Index 7, NDSI	10
	GBM	B, G, R, NIR, R2, R3, R4, SWIR 1, SWIR 2, TDS 1, TDS 2, TDS 3, TDS 4, TDS 6, TDS 7, Salinity Index 1, Salinity Index 2, Salinity Index 4, Salinity Index 5, Salinity Index 6, Salinity Index 7, NDSI	22
	XGBoost	NIR, R4, SWIR 2, TDS 2, TDS 6, Salinity Index 6, Salinity Index 7, NDSI	8
	Bagging	R4, SWIR 2, TDS 2, Salinity Index 6, NDSI	5
	AdaBoost	R4, SWIR 2, TDS 2, TDS 3, Salinity Index 6, Salinity Index 7, NDSI	7
TSS	LR	R4, NSMI, NDSSI	3
	Lasso	B, G, R, NIR, R1, R2, R3, R4, SWIR 1, SWIR 2, NSMI, NDSSI, BR, WSRI, NDVI	15
	Ridge	G, R, NIR, R1, R2, R4, SWIR 1, SWIR 2, NSMI, NDSSI, BR, WSRI, NDVI	13
	KNN	B, G, R, NIR, R2, R3, R4, SWIR 1, SWIR 2, NSMI, NDSSI, BR, WSRI, NDVI	14
	RF	B, G, R, NIR, R1, R2, R3, R4, SWIR 1, SWIR 2, NSMI, NDSSI, BR, WSRI, NDVI	15
	GBM	G, R, NIR, R1, R2, R3, R4, SWIR 1, NSMI, NDSSI, BR, WSRI, NDVI	13
	XGBoost	B, G, R, R2, R3, R4, SWIR 1, NSMI, NDSSI, BR, WSRI, NDVI	12
	Bagging	B, G, R, NIR, R1, R2, R3, R4, SWIR 1, SWIR 2, NSMI, NDSSI, BR, WSRI, NDVI	15
	AdaBoost	G, R, R1, R2, R3, R4, SWIR 1, NSMI, NDSSI, BR, WSRI, NDVI	12

Table 6. Optimal Features for each of the ML Models for the Landsat 8 OLI Images.

WQP	ML model	Optimal features	No. of features
TDS	LR	C, B, G, R, NIR, SWIR 1, SWIR 2, Salinity Index 1, Salinity Index 2, Salinity Index 4, Salinity Index 5, Salinity Index 6, Salinity Index 7, NDSI	14
	Lasso	C, B, G, R, NIR, SWIR 1, SWIR 2, Salinity Index 2, Salinity Index 4, Salinity Index 6, Salinity Index 7, NDSI	12

 doi:10.20944/preprints202603.0403.v1

	Ridge	C, B, G, R, NIR, SWIR 1, SWIR 2, Salinity Index 1, Salinity Index 2, Salinity Index 4, Salinity Index 5, Salinity Index 6, Salinity Index 7, NDSI	14
	KNN	C, G, R, SWIR 1, SWIR 2, Salinity Index 4, Salinity Index 7, NDSI	8
	RF	C, B, G, R, NIR, SWIR 1, SWIR 2, Salinity Index 4, Salinity Index 7, NDSI	10
	GBM	B, G, R, NIR, SWIR 1, SWIR 2, Salinity Index 4, Salinity Index 7, NDSI	9
	XGBoost	C, B, G, R, NIR, SWIR 1, SWIR 2, Salinity Index 1, Salinity Index 2, Salinity Index 4, Salinity Index 5, Salinity Index 6, Salinity Index 7, NDSI	14
	Bagging	C, B, G, R, NIR, SWIR 1, SWIR 2, Salinity Index 2, Salinity Index 4, Salinity Index 6, Salinity Index 7, NDSI	12
	AdaBoost	C, B, G, R, NIR, SWIR 1, SWIR 2, Salinity Index 1, Salinity Index 2, Salinity Index 4, Salinity Index 5, Salinity Index 6, Salinity Index 7, NDSI	14
TSS	LR	R, SWIR 1, NDVI	3
	Lasso	R, SWIR 1, NDVI	3
	Ridge	R, SWIR 1	2
	KNN	R, SWIR 1, NDVI	3
	RF	R, SWIR 1	2
	GBM	R, SWIR 1, SWIR 2, NDSSI, WSRI, NDVI	6
	XGBoost	R, SWIR 1, NDSSI	3
	Bagging	R, SWIR 1, NDVI	3
	AdaBoost	R, SWIR 1, NDSSI	3

Hyperparameter optimization

The results of the hyperparameter optimization of each model are presented in Tables 10 and 11. The optimal hyperparameters obtained to achieve the best performance of the model [143,165,166] are presented in Tables 10 and 11, for TDS and TSS, respectively.

Table 7. Optimal Values of the Hyperparameters of the ML Models for Sentinel 2 Images.

WQP	ML model	Hyperparameter	Optimal Value
TDS	LR	Fit intercept	False
	Lasso	α	1.0
	Ridge	α	1.0
	KNN	Number of neighbors	5
	RF	Number of estimators	100
		Learning rate	0.01
	GBM	Number of estimators	200
	XGBoost	Learning rate	0.1
		Maximum depth	3
	Bagging	Number of estimators	20
	AdaBoost	Number of estimators	150
TSS	LR	Fit intercept	False
	Lasso	α	10.0
	Ridge	α	10.0

 doi:10.20944/preprints202603.0403.v1

KNN	Number of neighbors	10
RF	Number of estimators	200
GBM	Learning rate	0.01
	Number of estimators	100
XGBoost	Learning rate	0.01
	Maximum depth	3
Bagging	Number of estimators	15
AdaBoost	Number of estimators	150

Table 8. Optimal Values for each of the ML Models for the Landsat 8 OLI Images.

WQP	ML model	Hyperparameter	Optimal Value
TDS	LR	Fit intercept	True
	Lasso	α	0.1
	Ridge	α	0.1
	KNN	Number of neighbors	10
	RF	Number of estimators	100
	GBM	Learning rate	0.01
		Number of estimators	300
	XGBoost	Learning rate	0.1
		Maximum depth	3
	Bagging	Number of estimators	20
	AdaBoost	Number of estimators	50
TSS	LR	Fit intercept	False
	Lasso	α	0.1
	Ridge	α	0.1
	KNN	Number of neighbors	10
	RF	Number of estimators	200
	GBM	Learning rate	0.01
		Number of estimators	100
	XGBoost	Learning rate	0.01
		Maximum depth	3
	Bagging	Number of estimators	15
	AdaBoost	Number of estimators	50

3.2.2. Model Evaluation Analysis

The model accuracy evaluation was performed on the three different datasets i.e., a) the training, b) the testing, and iii) external validation or unseen datasets as detailed in Section 2.5.3. Model assessment was conducted on all these different datasets in order to provide a thorough knowledge of the effectiveness, limitations, and overall generality of the ML models, which may be impacted by factors like the intricacy of the relationships present in the data being investigated [136].

Assessing the performance of the model on only the training phase do not typically serve as a benchmark for the overall model performance [167] hence there is a need to perform analysis all phases of the dataset to ensure effective prediction and decision-making. The accuracy evaluation of each of these datasets helps in the selection, optimization, and deployment of highly effective models that have the potential to accurately imitate real-life scenarios without overfitting and preventing poor generalization [136,144,150].

The results were obtained and summarized for the standalone and ensemble models for the TDS as shown in Table 12 and Table 16, respectively for Sentinel-2 and Landsat 8 OLI images while Tables 14 and 18, for the TSS estimations using the, respectively, for Sentinel-2 and Landsat 8 OLI images.

 doi:10.20944/preprints202603.0403.v1

The findings show that the standalone models can explain approximately 7 to 63% variations in TDS across the standalone models using the Sentinel images. The Landsat 8 OLI images could only explain about 11 to 31% variations in TDS across all the standalone models which is an indication of superior performance of Sentinel 2 in estimating TDS compared to Landsat 8 OLI. The ensemble models resulted in a superior prediction across the two images with the Sentinel-2 images capturing about the 81 to 100% variance in TDS across the three datasets. The ensemble models also show significant improvement in the model performance using the Landsat 8 OLI images compared to the standalone models with the models accounting for approximately 32 to 97% variance in TDS in the river system across all the datasets.

The models show similar but reduced performance in the estimations of TSS. While the Sentinel-2 images were able to account for about -2 to 14% of the variation across the standalone models, Landsat 8 OLI accounted for approximately -16 to 35% variations in TSS. The ensemble models show an improved performance with the spectral signatures explaining about 36 to 99% in TSS for the Sentinel-2 images. Landsat 8 OLI on the other hand saw a reduced explanation for the variance in TSS of about 11 to 90% across all the three datasets.

In general, the ensemble models show superior performance for the estimation of the TDS in the river system using the Sentinel 2 MSI and the Landsat 8 OLI images compared to the standalone models which is to be expected given their complexity in preventing overfitting and model generalizations. Sentinel 2 images were found to offer much accurate estimations potentially due to their finer spatial of 10 and 20 m resolution of combined 5 days temporal resolutions [18,39]. This result is supported by the study performed by [168] which have determined the Sentinel-2 images to be greater predictors of most of the WQPs studied. WQPs considered in the study include but are not limited to temperature, conductivity, turbidity, and Secchi disk depth. Additionally, researchers in [169] similarly found the Sentinel-2 images to be superior compared to the Landsat 8-OLI in retrieving WQPs such as chlorophyll-a concentrations in the King Talal Dam in Jordan.

Other researchers in the past have however noted the Landsat 8 OLI to have a higher accuracy in estimations as in the findings of researchers in [86] who found the Landsat 9 OLI-2 to have higher accuracy in the estimation of TSM (R^2 = 89%) compared to the 71% obtained with the Sentinel-2 images in Four Italian lakes. The greater signal-to-noise of the Landsat 9 OLI-2 (14 bit) compared with the 12-bit Sentinel-2 images was attributed to the improved performance of the Landsat-9 OLI-2 images using physics-based inversion models.

The models generally perform better in estimating TDS compared to TSS possibly due to the nature and the variability in these datasets in the river system. Unlike TDS which shows largely consistent variability in the river, TSS data presents underlying issues including spikes in the level of sediments in the river system contributed by the tributaries.

Possible reasons for the poor performance on the TSS data could include:

Limited amount of data: Matching up of the field and RS data resulted in insufficient amount of data needed by the models to learn, train, and understand the relationships and patterns of TSS and RS data which may result in the poor performance [120]. This is evident in the performance of ensemble models. Example R^2 of 0.90, RMSE = 32.80 mg/L, and MAE = 22.91 mg/L obtained on the external validation for the RF model and R^2 of 0.45, RMSE = 76.84 mg/L, and MAE = 42.60 mg/L for the XGBoost model as presented in Table 18.

Over or underfitting of the ML: the ML models may have been over or underfitted which occurs because of randomness “noise” or limitation of the data [91]. The models therefore fail to accurately capture the fundamental trends and patterns needed to ensure the model has the balance for the estimations.

Limitation on the representation of variation of TSS data: The huge fluctuations in the TSS across the river basin because of seasonal and environmental changes, anthropogenic contributions, and the varying watershed conditions results resulting limited representation of the TSS data making it difficult for the model to generalize the data [170].

 doi:10.20944/preprints202603.0403.v1

Absence of contextual or environmental factors in the model: the RS data can be said to not be sufficient to offer all the necessary contextual data that influences TSS concentrations in the river system. Other environmental factors including land use land cover changes, percent of imperviousness, runoff volumes, precipitation, antecedent dry days [118] which may be necessary to estimate TSS accurately in addition to RS.

Researchers [120] compared the various ML algorithms such as ANN, RF, SVR, and Cubist regression (CB) for the retrieval of WQPs including SSC utilizing band combinations of B1-B4, (B3)2, B3/B1, B1*B3, and B2*B3 obtained from Landsat 5, 7, and 8 for coastal waters in Hong Kong. The maximum observed SSC was 33.0 mg/L which is far less than the highest value of 69,200 mg/L observed for the Colorado River water system. The authors found the RF models to be the poorest model among the other models in estimating SSC with reported R^2, MAE, and RMSE of 0.51, 1.78 mg/L, and 3.11 mg/L, respectively, were obtained as the cross-validation results with a data size of 120. The reason for the poor performance of the RF was attributed to the possibly small amount of data utilized for model training. This research however found the RF to be the best performing model for TSS estimations with Landsat 8 OLI with R^2 of 0.90, MAE of 22.91 mg/L, and RMSE 32.80 mg/L on the external dataset with a total size of 492 split for training, testing, and external validation as shown in Table 18.

The ML models show varying performance from the training stage to the external validation stages for the estimation of TDS as shown in Tables 12 and 16 for Sentinel-2 and Landsat 8 OLI images, respectively. The results presented in the tables had the LR seeing its R^2 increase from 0.48 (training) to 0.59 in the external validation stage for the Sentinel-2 images, the R^2 for the models like KNN reduced from 0.63 in the training phase to 0.60 in the external validation stage for the Sentinel-2 images. The XGBoost, however, maintains a similar R^2 from the training stage (0.98) to the external validation stage (0.99) with the Sentinel-2 image. Similar trends were seen with the Landsat 8 OLI data for the TDS estimations where the LR saw R^2 increase from 0.14 to 0.31. The XGBoost, however, saw a significant increment in R^2 from 0.70 to 0.97 with the Landsat 8 OLI images in the TDS estimations.

Models that perform consistently over all phases of datasets are more likely to make accurate predictions in real-life situations. An example is the case of XGBoost for TDS estimations using the Sentinel-2 images with slight R^2 variations of 0.98, 1.00, and 0.99, respectively for the training, testing, and external, validation stages of the modeling. Models like GBM saw its performance in terms of R^2 rise from 0.44 in the training stage to 0.74 and RMSE reducing from about 190.8 mg/L to 110.4 mg/L with the Landsat OLI images for TDS estimations.

The variations in the models could result from model underlying algorithms, bias-variance trade-offs and overfitting in the training dataset, and the spectral, spatial, and temporal resolutions of the images acquired from the two sensors; as a result, it is necessary to test the model on the external or unseen datasets to create an unbiased assessment of the models' accuracy and ensure validation of their ability for generalization [23,119,136,139–141].

TSS estimations with Sentinel-2 images saw ensemble models such AdaBoost obtained highest R^2, and least RMSE across all the datasets with obtained values of (0.91,52.62 mg/L), (0.99, 24.34 mg/L), and (0.92, 29.48 mg/L), respectively, for the training, testing, and external validation which is somewhat consistent with the findings of researchers in [118] which found the AdaBoost and RF as the best models for the estimation of TSS in urban watersheds in some 17 states in the USA. R^2 of 0.77 and 0.67 were obtained, respectively, for the training and prediction steps of the AdaBoost model. Researchers [171] in however found the AdaBoost to produce R^2 of 0.784, 0.845, and 0.695, respectively, for the estimation of chlorophyll-a concentration, turbidity, and NH3-N using images from an unmanned aerial vehicle (UAV) for the Nanfei River, China.

Results obtained from this study show some level of agreement with previous work. Researchers in [62] studied three WQPs including TDS in Lake Tana located in Ethiopia using Landsat 8 OLI images. The researchers found RF as the best performing model for the estimation of TDS with the retrieved R^2 of 0.79, NSE of 0.80, RMSE of 12.30 mg/L, and MARE of 0.082 from 18 optimal features

 doi:10.20944/preprints202603.0403.v1

obtained from the images including (B4 + B3)/2, (B4 + B2)/2, (B3 + B2)/2, (B2 + B3 + B4)/3. No external validation was reported in the study. In contrast, this study resulted in R^2 of 0.85, NSE of 0.85, RMSE of 84.04 mg/L, and MARE of 0.23 for the unseen (external) dataset with the obtained with optimal features outlined in Table 16. The varying outputs between the study in the Tana Lake and this study particular with the RMSE is as results of the conditions in the two waterbodies. TDS concentrations in Tana Lake varies from 7.30 and 113.3 mg/L while in this study range from 23 to 17,200 mg/L.

Furthermore, researchers in [167] applied models such as principal component regression (PCR), Gaussian process regression (GPR), and backpropagation neural network (BPNN) for modeling TDS from a variety of WQPs including pH, TSS, turbidity, and EC among others obtained from surface, ground, and drinking waters in the Ghanaian mining town of Tarkwa, the authors discovered GPR average values of R^2, MAE, and RMSE of 98.7%, 7.910, and 4.090 mg/L, respectively. The TDS values obtained in the water sampled for the study ranged from 8.81 to 534.00 mg/L.

Assessing the overall performance of the ML models on just one or two evaluation indicators as reported in most studies may not present a broad understanding of the overall performance of the model. Additionally, evaluating performance based on various specific measures independently may be challenging.

In TDS estimations using Sentinel-2 images shown in Table 12, for instance, AdaBoost produced R^2 of 81% on the external validation, which was the fifth best index, but its performance when assessed with PBIAS was -12.51% which was the third best. The same model produced PMARE 31.97 mg/L (6th best) indicating that focusing on just one metric may not be sufficient to distinguish one model's performance from another. Considering this, ranking summary that includes several performance factors and provides a thorough evaluation of each model's overall performance was utilized, as shown in Tables 13 and 17 for Sentinel-2 and Landsat 8 OLI images, respectively.

These rankings offer a well-organized and consistent view of the model evaluation across the three datasets. The models are rated for each dataset according to how well they perform on the metrics employed, with the greatest R^2 and NSE both allocated 1 and the least values assigned 9 for the nine models. Models with lower RMSE, PMARE, MAE, SI, absolute PBIAS, and RSD values given 1; those with higher values are given a value of 9 [3,18,125,141,155–158].

The ranking results demonstrate that the performance of the model varies from the training to external validation stages of the datasets, with a clear change being observed in some models from the training stage to another dataset stage. The average rank of the model for the Bagging used in TDS estimation with Landsat 8 OLI image changed from 2.0 in the training phase to 2.9 in the testing phase and 2.9 in the external validation phase, showing that performance assessments on the training dataset may not be the ideal baseline for future model performance considering the relatively similar performance ranks in the testing and external validation datasets [167].

The rank summary analysis demonstrate that the ensemble models are generally superior to the standalone models in estimating TDS for images from both sensors, with XGBoost being the best model with average rank of 1.0 across all the three datasets for the Sentinel-2 images and average rank of 3.0, 1.0, and 1.0, respectively, for the training, testing, and the external validation phases of the modeling for the Sentinel and Landsat 8 OLI images, as shown in Tables 13 and 17, respectively. The performance of the XGBoost model particularly for the Landsat 8 OLI were similar to that of the RF which had an average of 1.0, 2.1, and 2.5, respectively, for the training, testing, and the external validation phases of the modeling as shown in Table 17. Similar overall ranks of 1.7 and 1.9 were obtained for the XGBoost and the RF, respectively, which corroborate the similar underlying algorithms in these models as described in Section 2.5.2. The primary distinction between the two models is that whereas XGBoost constructs decision trees sequentially, RF constructs its trees independently [134]. The XGBoost was thus discovered to be the most optimal model estimating TDS using images from both Sentinel-2 and Landsat 8 OLI. This finding may be as a result of the underlying algorithms of the XGBoost models, which incorporate multiple decision trees to weaker learner decision trees to handle or resolve errors created from prior trees, resulting in a more accurate and robust estimations [60,126,172]. The optimal hyperparameters and features obtained in Section

 doi:10.20944/preprints202603.0403.v1

3.2.1 may also have played a substantial role in determining the model's overall efficient performance due to its potential to reduce complexity and improve the generalizability [55,150].

The rank summary outlined in Table 15 and 19 estimations showed the AdaBoost to be the best model for estimating TSS models with an average rank of 2.0, 1.3, and 1.8, respectively on the training, testing, and external validation datasets using the Sentinel-2 images. The RF was however found to be best in estimating TSS from the Landsat 8 OLI with average ranks of 1.3, 2.3, and 1.4, respectively, for the training, testing, and external validation phases of the modeling. The obtained results for the TSS particularly by the standalone models were poor compared to those produced for the TDS estimations possibly due to the nature of the TSS datasets.

The results of the performance evaluations showed that the combination of RS images and ML models could offer a time and cost-effective method for estimating WQPs which is essential for monitoring water quality and for making well-informed decisions about the distribution of available water resources.

Table 12. Performance Metrics of ML Models in Retrieving TDS Using Sentinel-2 Images.

External Validation (Unseen Data) (Sample size = 58)							
R^2	RMSE (mg/L)	MAE (mg/L)	PMARE (%)	SI	NSE	PBIAS (%)	RSD
0.59	221.84	152.53	36.18	0.33	0.59	-16.62	0.64
0.21	307.29	217.79	57.40	0.46	0.21	-36.08	0.89
0.16	317.40	222.76	62.17	0.48	0.16	-41.34	0.92
0.60	219.45	142.21	31.82	0.33	0.60	-19.17	0.63
0.86	131.42	81.33	15.79	0.20	0.86	-8.38	0.38
0.90	109.65	82.89	21.25	0.17	0.90	-12.94	0.32
0.99	26.52	19.19	3.66	0.04	0.99	-1.39	0.08
0.84	138.08	85.48	20.45	0.21	0.84	-13.47	0.40
0.81	152.40	131.33	31.97	0.23	0.81	-12.51	0.44

Table 13. Summary of the Performance Ranks for the ML Models in Retrieving TDS Using Sentinel-2 Images.

	Training (Sample size =129)									Testing (Sample size = 33)								
Model	R^2	RMSE	MAE	PMARE	SI	NSE	PBIAS	RSD	Avg.	R^2	RMSE	MAE	PMARE	SI	NSE	PBIAS	RSD	Avg.
LR	7	7	7	7	7	7	6	7	6.9	7	7	7	6	7	7	6	7	6.8
Lasso	8	8	8	9	8	8	8	8	8.1	8	8	8	8	8	8	8	8	8.0
Ridge	9	9	9	8	9	9	9	9	8.9	9	9	9	9	9	9	9	9	9.0
KNN	6	6	6	6	6	6	7	6	6.1	6	6	6	7	6	6	7	6	6.3
RF	2	2	2	3	2	2	3	2	2.3	5	5	4	5	5	5	5	5	4.9
GBM	5	5	4	5	5	5	5	5	4.9	3	3	3	3	3	3	3	3	3.0
XGBoost	1	1	1	1	1	1	1	1	1.0	1	1	1	1	1	1	1	1	1.0
Bagging	3	3	3	2	3	3	2	3	2.8	4	4	5	4	4	4	4	4	4.1

 doi:10.20944/preprints202603.0403.v1

AdaBoost	4	4	5	4	4	4	4	4	4.1	2	2	2	2	2	2	2	2	2	2.0

External Validation (Unseen Data) (Sample size = 58)									
R^2	RMSE	MAE	PMARE	SI	NSE	PBIAS	RSD	Avg.	Overall Avg.
7	7	7	7	7	7	6	7	6.9	6.8
8	8	8	8	8	8	8	8	8.0	8.0
9	9	9	9	9	9	9	9	9.0	9.0
6	6	6	5	6	6	7	6	6.0	6.1
3	3	2	2	3	3	2	3	2.6	3.3
2	2	3	4	2	2	4	2	2.6	3.5
1	1	1	1	1	1	1	1	1.0	1.0
4	4	4	3	4	4	5	4	4.0	3.6
5	5	5	6	5	5	3	5	4.9	3.7

Table 14. Performance Metrics of ML Models in Retrieving TSS Using Sentinel-2 Images.

	Training (Sample size =84)								Testing (Sample size = 22)							
Model	R^2	RMSE (mg/L)	MAE (mg/L)	PMARE (%)	SI	NSE	PBIAS (%)	RSD	R^2	RMSE (mg/L)	MAE (mg/L)	PMARE (%)	SI	NSE	PBIAS (%)	RSD
LR	-0.02	177.92	112.11	2077.26	2.15	-0.02	-2050.99	1.01	0.06	248.22	123.86	2053.70	2.59	0.06	-1510.55	0.97
Lasso	0.00	176.52	113.03	2054.96	2.13	0.00	-2030.07	1.00	0.00	256.10	116.79	1879.68	2.68	0.00	-1862.39	1.00
Ridge	0.02	174.98	112.35	2103.13	2.11	0.02	-2078.24	0.99	0.01	254.66	115.53	1835.47	2.66	0.01	-1818.72	0.99
KNN	0.14	163.71	96.41	1355.50	1.97	0.14	-1329.07	0.93	0.11	242.07	94.18	681.36	2.51	0.11	-652.84	0.94
RF	0.84	70.27	42.61	588.66	0.85	0.84	-578.23	0.40	0.78	119.89	59.70	648.49	1.25	0.78	-635.58	0.47
GBM	0.52	121.83	80.51	1410.55	1.47	0.52	-1390.02	0.69	0.86	96.68	51.63	891.42	1.01	0.86	-877.74	0.38

 doi:10.20944/preprints202603.0403.v1

XGBoost	0.48	127.28	61.48	340.02	1.48	0.48	-308.73	0.72	0.58	165.44	56.46	120.20	1.64	0.58	-63.95	0.65
Bagging	0.82	75.53	44.84	743.26	0.91	0.82	-731.65	0.43	0.71	137.65	75.49	956.74	1.41	0.71	-941.29	0.54
AdaBoost	0.91	52.62	45.40	984.17	0.60	0.91	-971.47	0.30	0.99	24.34	15.75	271.69	0.25	0.99	-253.72	0.09

External Validation (Unseen Data) (Sample size = 36)							
R^2	RMSE (mg/L)	MAE (mg/L)	PMARE (%)	SI	NSE	PBIAS (%)	RSD
0.04	104.59	72.23	1051.36	1.22	0.04	-1023.21	0.98
0.00	106.67	76.09	1143.63	1.24	0.00	-1114.34	1.00
0.01	106.07	75.52	1124.44	1.24	0.01	-1095.46	0.99
0.03	105.08	76.41	1051.67	1.22	0.03	-1022.28	0.99
0.82	45.56	29.82	448.15	0.53	0.82	-438.38	0.43
0.65	63.40	47.56	645.42	0.74	0.65	-625.75	0.59
0.36	85.08	47.95	250.52	0.89	0.36	-201.29	0.80
0.85	41.53	26.10	319.58	0.48	0.85	-306.92	0.39
0.92	29.48	24.64	466.70	0.32	0.92	-456.19	0.28

Table 15. Summary of the Performance Ranks for the ML Models in Retrieving TSS Using Sentinel-2 Images.

	Training (Sample size =84)									Testing (Sample size = 22)								
Model	R^2	RMSE	MAE	PMARE	SI	NSE	PBIAS	RSD	Avg.	R^2	RMSE	MAE	PMARE	SI	NSE	PBIAS	RSD	Avg.
LR	9	9	7	8	9	9	8	9	8.5	7	7	9	9	7	7	7	7	7.5
Lasso	8	8	9	7	8	8	7	8	7.9	9	9	8	8	9	9	9	9	8.8
Ridge	7	7	8	9	7	7	9	7	7.6	8	8	7	7	8	8	8	8	7.8
KNN	6	6	6	5	6	6	5	6	5.8	6	6	6	4	6	6	4	6	5.5
RF	2	2	1	2	2	2	2	2	1.9	3	3	4	3	3	3	3	3	3.1
GBM	4	4	5	6	4	4	6	4	4.6	2	2	2	5	2	2	5	2	2.8
XGBoost	5	5	4	1	5	5	1	5	3.9	5	5	3	1	5	5	1	5	3.8

 doi:10.20944/preprints202603.0403.v1

Bagging	3	3	2	3	3	3	3	3	2.9	4	4	5	6	4	4	6	4	4.6
AdaBoost	1	1	3	4	1	1	4	1	2.0	1	1	1	2	1	1	2	1	1.3

External Validation (Unseen Data) (Sample size = 36)									
R^2	RMSE	MAE	PMARE	SI	NSE	PBIAS	RSD	Avg.	Overall Avg.
6	6	6	6	6	6	7	6	6.1	7.4
9	9	8	9	9	9	9	9	8.9	8.5
8	8	7	8	8	8	8	8	7.9	7.8
7	7	9	7	7	7	6	7	7.1	6.1
3	3	3	3	3	3	3	3	3.0	2.7
4	4	4	5	4	4	5	4	4.3	3.9
5	5	5	1	5	5	1	5	4.0	3.9
2	2	2	2	2	2	2	2	2.0	3.2
1	1	1	4	1	1	4	1	1.8	1.7

Table 16. Performance Metrics of ML Models in Retrieving TDS Using Landsat 8 OLI Images.

	Training (Sample size =414)								Testing (Sample size = 104)							
Model	R^2	RMSE (mg/L)	MAE (mg/L)	PMARE (%)	SI	NSE	PBIAS (%)	RSD	R^2	RMSE (mg/L)	MAE (mg/L)	PMARE (%)	SI	NSE	PBIAS (%)	RSD
LR	0.14	236.91	182.87	67.73	0.43	0.14	-45.79	0.93	0.26	220.42	178.24	52.44	0.40	0.26	-29.53	0.86
Lasso	0.12	239.50	186.85	69.62	0.43	0.12	-47.45	0.94	0.17	232.55	190.63	58.20	0.42	0.17	-34.18	0.91
Ridge	0.11	242.00	190.95	71.33	0.44	0.11	-48.77	0.95	0.17	233.11	190.07	58.35	0.42	0.17	-34.44	0.91
KNN	0.26	219.92	175.49	60.00	0.40	0.26	-37.94	0.86	0.27	218.74	182.21	53.56	0.40	0.27	-29.13	0.85
RF	0.88	88.68	68.39	23.36	0.16	0.88	-15.00	0.35	0.88	89.85	69.82	20.90	0.16	0.88	-12.85	0.35
GBM	0.44	190.84	149.98	54.45	0.34	0.44	-36.27	0.75	0.72	134.44	106.75	31.48	0.24	0.72	-18.47	0.53

 doi:10.20944/preprints202603.0403.v1

XGBoost	0.70	140.92	109.06	38.13	0.25	0.70	-24.55	0.55	0.96	53.01	40.08	10.72	0.10	0.96	-5.65	0.21
Bagging	0.85	99.43	72.42	26.13	0.18	0.85	-16.81	0.39	0.85	100.68	74.04	21.01	0.18	0.85	-11.81	0.39
AdaBoost	0.32	211.66	175.56	61.95	0.38	0.32	-41.87	0.83	0.64	152.70	129.79	37.69	0.28	0.64	-22.36	0.60

External Validation (Unseen Data) (Sample size = 63)							
R^2	RMSE (mg/L)	MAE (mg/L)	PMARE (%)	SI	NSE	PBIAS (%)	RSD
0.31	180.43	149.43	43.34	0.33	0.31	-19.19	0.83
0.23	191.47	154.57	48.01	0.35	0.23	-27.57	0.88
0.15	200.25	171.57	54.33	0.36	0.15	-31.62	0.92
0.11	205.86	177.43	57.54	0.37	0.11	-33.16	0.95
0.85	84.04	71.28	23.35	0.15	0.85	-14.30	0.39
0.74	110.43	94.43	25.86	0.20	0.74	-13.31	0.51
0.97	35.82	27.90	7.59	0.07	0.97	-3.69	0.16
0.83	90.21	72.03	22.70	0.16	0.83	-13.24	0.41
0.67	125.47	110.77	25.92	0.23	0.67	-11.92	0.58

	Training (Sample size =414)									Testing (Sample size = 104)								
Model	R^2	RMSE	MAE	PMARE	SI	NSE	PBIAS	RSD	Avg.	R^2	RMSE	MAE	PMARE	SI	NSE	PBIAS	RSD	Avg.
LR	7	7	7	7	7	7	7	7	7.0	7	7	5	5	7	7	7	7	6.5
Lasso	8	8	8	8	8	8	8	8	8.0	8	8	9	8	8	8	8	8	8.1
Ridge	9	9	9	9	9	9	9	9	9.0	9	9	8	9	9	9	9	9	8.9
KNN	6	6	6	6	6	6	5	6	5.9	6	6	7	7	6	6	6	6	6.3
RF	1	1	1	1	1	1	1	1	1.0	2	2	2	2	2	2	3	2	2.1
GBM	4	4	4	4	4	4	4	4	4.0	4	4	4	4	4	4	4	4	4.0
XGBoost	3	3	3	3	3	3	3	3	3.0	1	1	1	1	1	1	1	1	1.0
Bagging	2	2	2	2	2	2	2	2	2.0	3	3	3	3	3	3	2	3	2.9
AdaBoost	5	5	5	5	5	5	6	5	5.1	5	5	6	6	5	5	5	5	5.3

 doi:10.20944/preprints202603.0403.v1

Table 17. Summary of the Performance Ranks for the ML Models in Retrieving TDS using Landsat 8 OLI Image.

External Validation (Unseen Data) (Sample size = 63)									
R^2	RMSE	MAE	PMARE	SI	NSE	PBIAS	RSD	Avg.	Overall Avg.
6	6	5	5	6	6	6	6	5.8	6.4
7	7	7	7	7	7	7	7	7.0	7.7
8	8	8	8	8	8	8	8	8.0	8.6
9	9	9	9	9	9	9	9	9.0	7.0
2	2	2	3	2	2	5	2	2.5	1.9
4	4	4	4	4	4	4	4	4.0	4.0
1	1	1	1	1	1	1	1	1.0	1.7
3	3	3	2	3	3	3	3	2.9	2.6
5	5	6	6	5	5	2	5	4.9	5.1

	Training (Sample size =351)								Testing (Sample size = 88)							
Model	R^2	RMSE (mg/L)	MAE (mg/L)	PMARE (%)	SI	NSE	PBIAS (%)	RSD	R^2	RMSE (mg/L)	MAE (mg/L)	PMARE (%)	SI	NSE	PBIAS (%)	RSD
LR	-0.02	265.04	134.02	1466.99	2.18	-0.02	-731.50	1.01	-0.15	206.79	112.46	1267.37	1.88	-0.15	-938.84	1.07
Lasso	0.05	256.61	146.46	1937.03	2.15	0.05	-1913.29	0.98	0.01	191.42	121.48	1872.09	1.81	0.01	-1844.04	0.99
Ridge	0.04	257.37	148.12	2159.21	2.15	0.04	-2134.76	0.98	0.01	191.80	122.19	1919.41	1.81	0.01	-1891.21	1.00
KNN	0.35	212.22	103.90	996.67	1.78	0.35	-970.48	0.81	0.33	157.71	91.16	1108.44	1.49	0.33	-1086.76	0.82
RF	0.86	97.79	48.40	407.74	0.82	0.86	-395.66	0.37	0.88	66.82	33.71	415.77	0.63	0.88	-403.25	0.35
GBM	0.37	209.42	114.57	1450.33	1.75	0.37	-1432.22	0.80	0.67	111.34	68.71	1154.34	1.05	0.67	-1136.57	0.58
XGBoost	0.30	220.42	93.22	571.69	1.81	0.30	-535.73	0.84	0.51	135.27	64.55	417.18	1.21	0.51	-366.98	0.70

 doi:10.20944/preprints202603.0403.v1

Bagging	0.81	114.58	50.02	396.06	0.96	0.81	-381.67	0.44	0.88	65.80	30.24	218.64	0.62	0.88	-202.55	0.34
AdaBoost	0.11	247.60	217.04	4410.96	1.53	0.11	-4404.68	0.94	0.87	69.87	59.69	1431.77	0.60	0.87	-1419.74	0.36

Table 18. Performance Metrics of ML Models in Retrieving TSS Using Landsat 8 OLI Image.

External Validation (Unseen Data) (Sample size = 53)							
R^2	RMSE (mg/L)	MAE (mg/L)	PMARE (%)	SI	NSE	PBIAS (%)	RSD
-0.16	111.08	73.15	458.57	1.34	-0.16	-183.54	1.08
0.08	98.92	73.68	862.89	1.27	0.08	-831.09	0.96
0.06	100.22	74.35	982.96	1.29	0.06	-951.28	0.97
0.18	93.42	70.19	914.83	1.20	0.18	-888.92	0.91
0.90	32.80	22.91	356.09	0.42	0.90	-342.41	0.32
0.66	59.85	44.51	672.55	0.77	0.66	-652.96	0.58
0.45	76.84	42.60	196.12	0.90	0.45	-144.49	0.74
0.84	41.11	26.65	489.58	0.53	0.84	-475.03	0.40
0.84	41.48	31.99	488.40	0.52	0.84	-475.01	0.40

	Training (Sample size =351)									Testing (Sample size =88)								
Model	R^2	RMSE	MAE	PMARE	SI	NSE	PBIAS	RSD	Avg.	R^2	RMSE	MAE	PMARE	SI	NSE	PBIAS	RSD	Avg.
LR	9	9	6	6	9	9	4	9	7.6	9	9	7	6	9	9	4	9	7.8
Lasso	7	7	7	7	7	7	7	7	7.0	7	7	8	8	7	7	8	7	7.4
Ridge	8	8	8	8	8	8	8	8	8.0	8	8	9	9	8	8	9	8	8.4
KNN	4	4	4	4	5	4	5	4	4.3	6	6	6	4	6	6	5	6	5.6
RF	1	1	1	2	1	1	2	1	1.3	2	2	2	2	3	2	3	2	2.3
GBM	3	3	5	5	4	3	6	3	4.0	4	4	5	5	4	4	6	4	4.5
XGBoost	5	5	3	3	6	5	3	5	4.4	5	5	4	3	5	5	2	5	4.3
Bagging	2	2	2	1	2	2	1	2	1.8	1	1	1	1	2	1	1	1	1.1
AdaBoost	6	6	9	9	3	6	9	6	6.8	3	3	3	7	1	3	7	3	3.8

 doi:10.20944/preprints202603.0403.v1

Table 19. Summary of the Performance Ranks for the ML Models in Retrieving TSS Using Landsat 8 OLI Image.

External Validation (Unseen Data) (Sample size = 53)									
R^2	RMSE	MAE	PMARE	SI	NSE	PBIAS	RSD	Avg.	Overall Avg.
9	9	7	3	9	9	2	9	7.1	7.5
7	7	8	7	7	7	7	7	7.1	7.2
8	8	9	9	8	8	9	8	8.4	8.3
6	6	6	8	6	6	8	6	6.5	5.5
1	1	1	2	1	1	3	1	1.4	1.6
4	4	5	6	4	4	6	4	4.6	4.4
5	5	4	1	5	5	1	5	3.9	4.2
2	2	2	5	3	2	5	2	2.9	1.9
3	3	3	4	2	3	4	3	3.1	4.5

The findings of the model's evaluation on the TDS external validation datasets, which reflect possible estimations from practical scenarios, are further displayed with scatter plots as shown in Figures 5 and 7, each with a 45° bisector line, respectively, for Sentinel-2 and Landsat 8 OLI images to aid in visualization and interpretation of the models by researchers and water resources managers. The best performing models have observed and estimated values lying around the 45° bisector line [140]. Results presented in the figures shows that ensemble models such as XGBoost, RF, Bagging producing superior model estimates of $R^2 > 83\%$ and $> 82\%$ for TDS using Sentinel-2 and Land-sat 8 OLI images, respectively. These findings are consistent with previous research, which discovered that ensemble models, such as the XGBoost and RF, outperformed standalone models, such as the SVM, ANN, and LR [62,173].

The scatterplot produced for the TSS estimations using the Sentinel-2 and Landsat 8 OLI images are shown in Figures 6 and 8, each with a 45° bisector line, respectively. Var-ying performance was observed for the TSS with ensemble models such as AdaBoost, RF, and Bagging producing $R^2 > 0.81$ with Senitnel-2 images which is consistent with previous research, which discovered that ensemble models, such as the XGBoost and RF, out-performed standalone models, such as the SVM, ANN, and LR [62,126,173]. Ensemble models such as RF and Bagging also outperformed standalone models such as LR, Lasso, and Ridge in the case of the Landsat 8 OLI images. Previous studies have however found standalone models such as ANN or neural networks to be good performing models [39,120].

The box and whisker plots were also used to visualize the performance of the models [174] on the external validation datasets. The plots presented in Figures 9 and 10, and Figures 11 and 12 representing estimations from the Sentinel-2 and Landsat 8 OLI images, respectively, are useful in comparing and analyzing the variations in the performance of the different models [140]. Results shown in the figures offer the opportunity to evaluate and assess each ML model's and RS image's effectiveness at estimating the WQPs under study more thoroughly.

The box and whisker plots show ranging percentiles which represent the variation in the observed and the estimated datasets. The horizontal line in the plot represents the 50th percentile or the median value. The interquartile range which measures the difference between the 25th and 75th percentiles is calculated using the values of the top and bottom horizontal lines of the box plots. The whisker lines are extensions of these percentile lines beyond the boxplots. The box and whisker plot may also be characterized by outliers which scattered dots seen below or above the whiskers [140].

 doi:10.20944/preprints202603.0403.v1

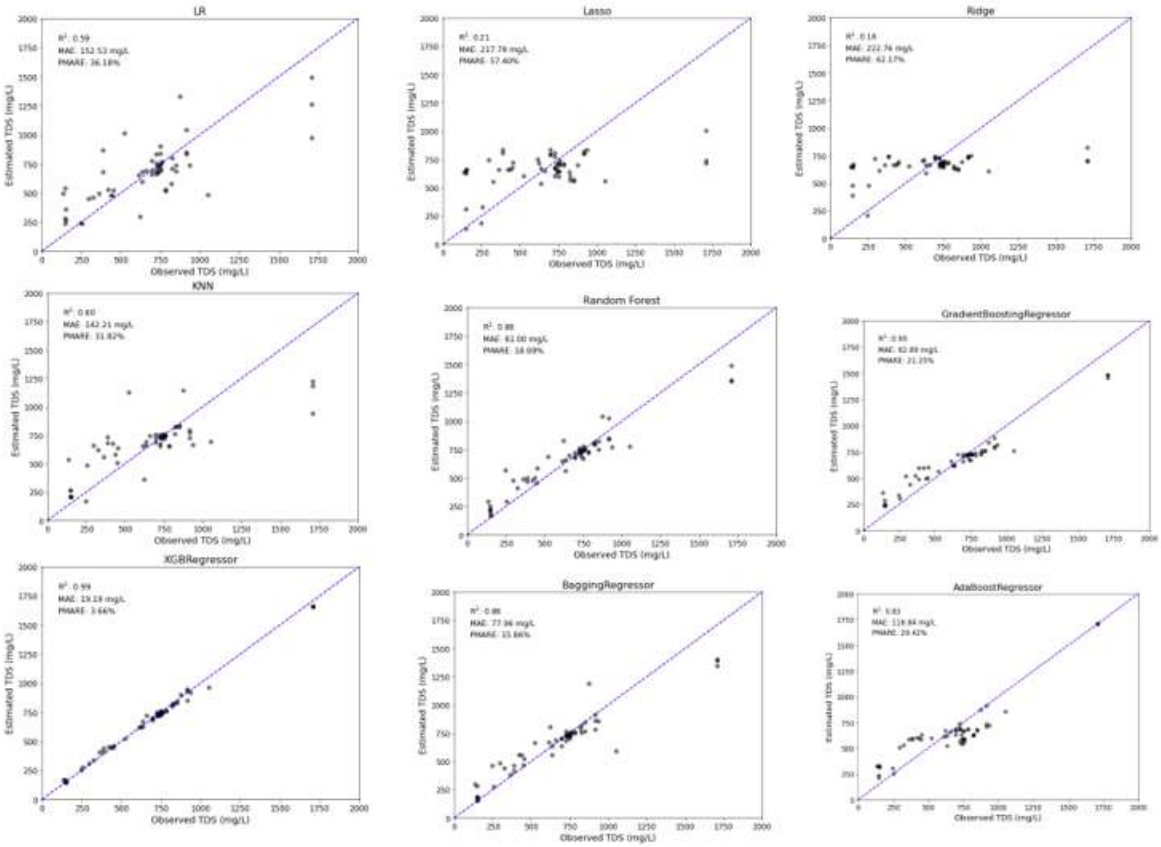


Figure 5. TDS Estimations from Sentinel-2 Images for the Various ML Models with the 45° Bisector Line.

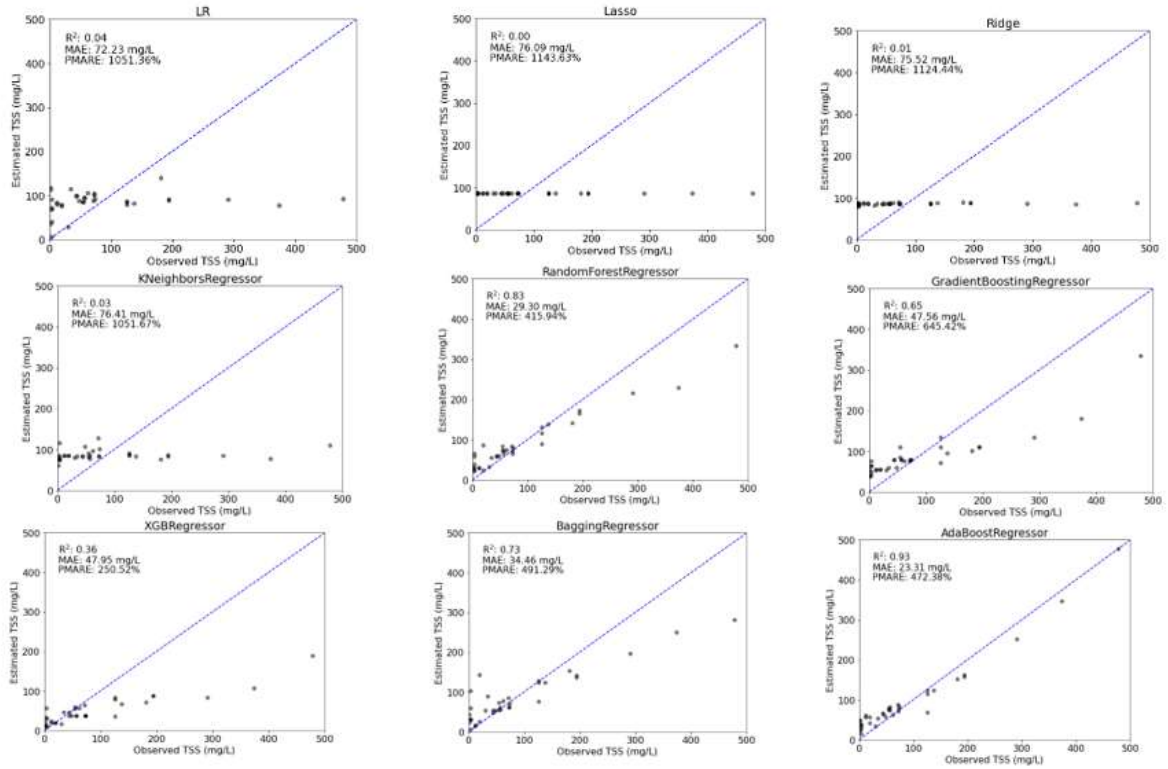


Figure 6. TSS Estimations from Sentinel-2 Images for the Various ML Models with the 45° Bisector Line.

 doi:10.20944/preprints202603.0403.v1

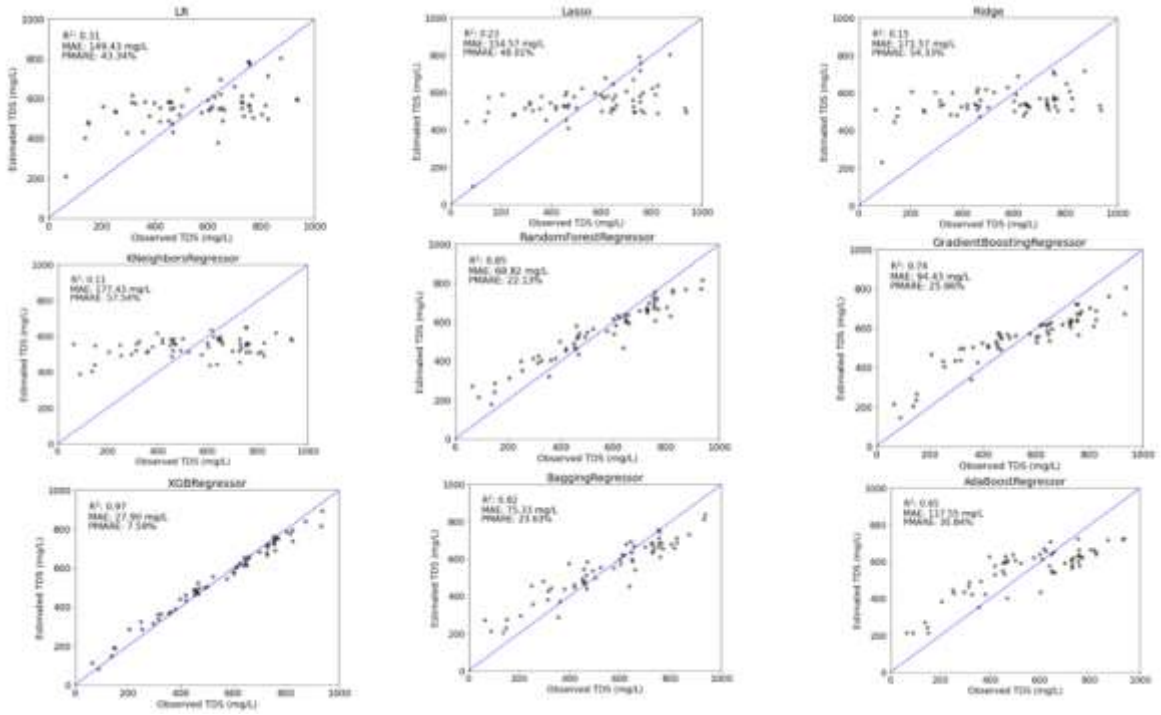


Figure 7. TDS Estimations from Landsat 8 OLI images for the Various ML Models with the 45° Bisector Line.

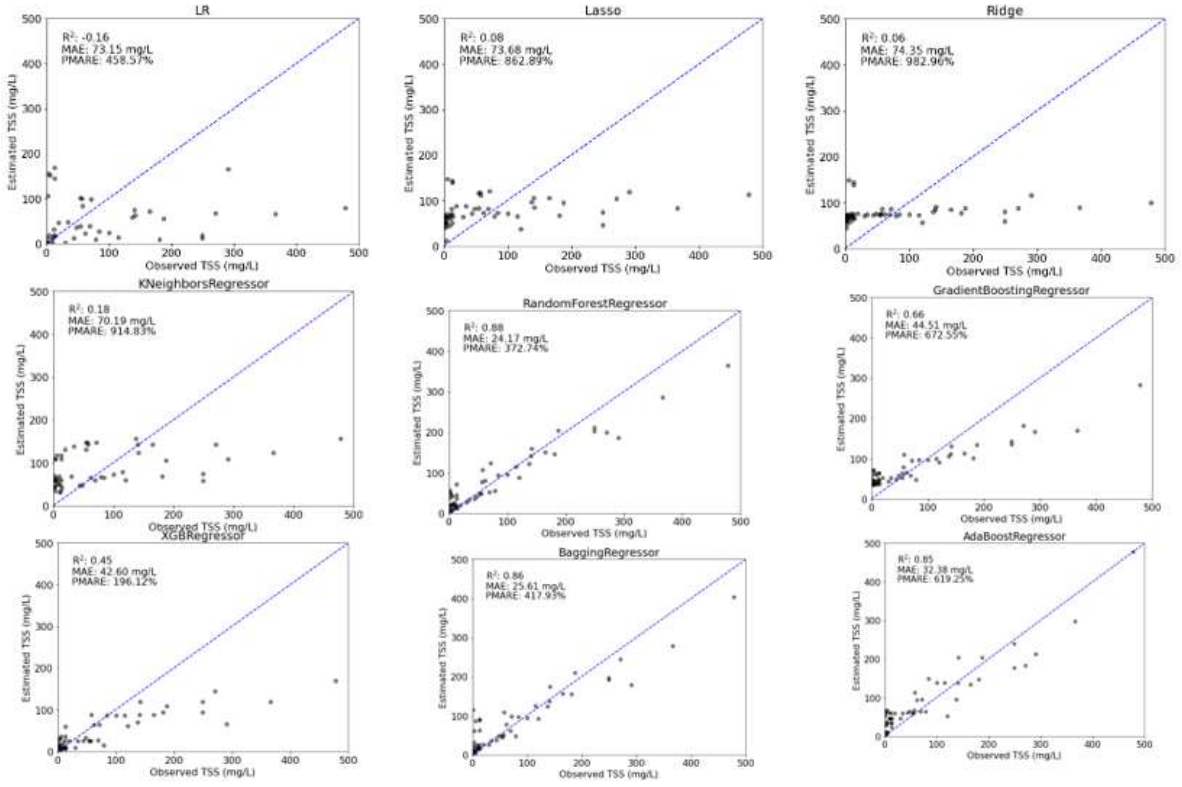


Figure 8. TSS Estimations from Landsat 8 OLI Images for the Various ML Models with the 45° Bisector Line.

The best-performing models exhibit smaller interquartile ranges and whiskers while poor-performing models exhibit wider interquartile ranges and long whiskers, an indication of greater variations and the possibility of underlying errors in the estimations. The results obtained for the TDS estimations demonstrate that the ensemble successfully captured extreme values, i.e., both low and

 doi:10.20944/preprints202603.0403.v1

high TDS levels, with the capture being more distinct in the best models as [140] as seen in the case of XGBoost for images from both Sentinel-2 and Landsat 8 OLI sensors shown in Figures 9 and 11

The box and whisker plots for TSS estimations are presented in Figures 10 and 12 for the Sentinel-2 and Landsat 8 OLI images, respectively, with AdaBoost and RF noted to show effective retrieving the TSS concentrations, respectively for the Sentinel-2 and Landsat 8 OLI images.

These findings provide a foundation for further research, the identification of areas for modeling technique improvement, and the selection of the best models for WQP estimation in the river system.

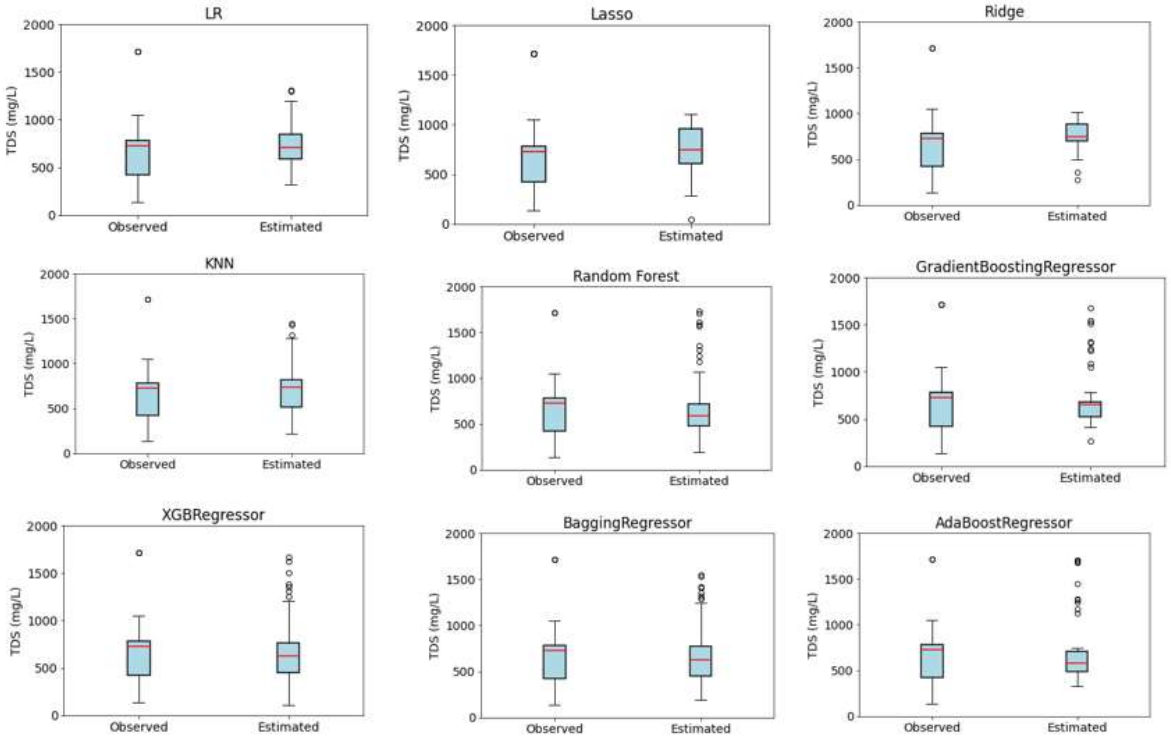
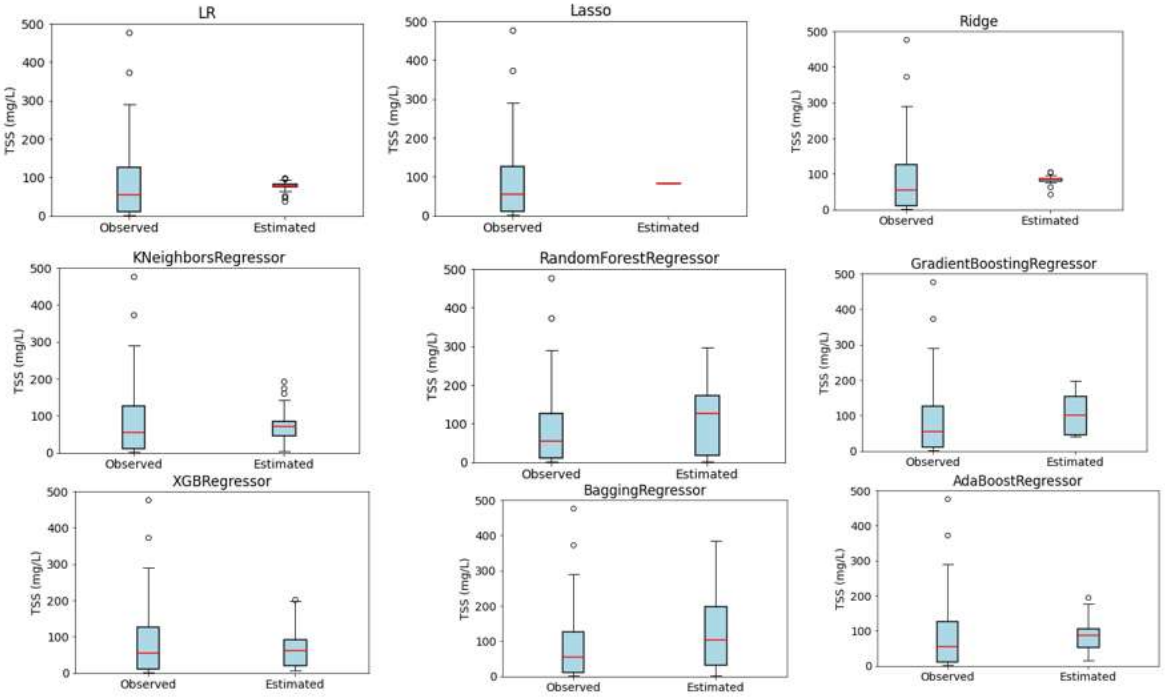


Figure 9. Box Plots of the Observed and Estimated TDS Concentration Retrieved by the Various ML Models Using Sentinel-2 Images.



 doi:10.20944/preprints202603.0403.v1

Figure 10. Box Plots of the Observed and Estimated TSS Concentration Retrieved by the Various ML Models Using Sentinel-2 Images.

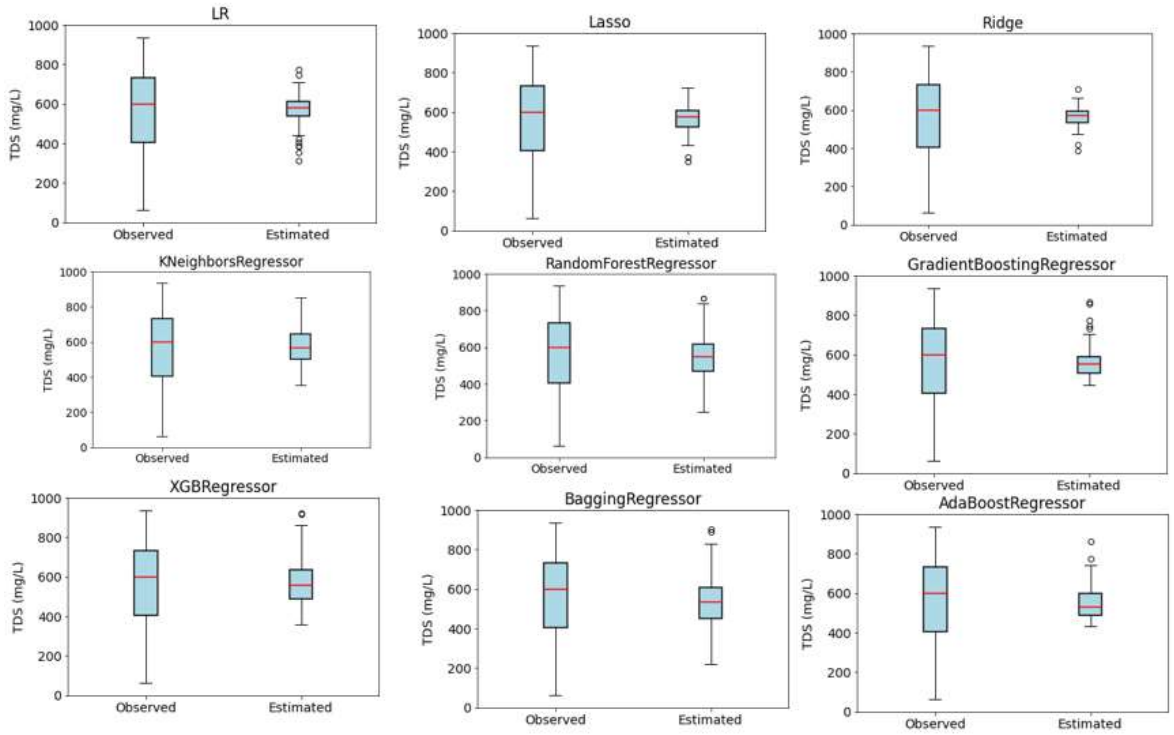


Figure 11. Box Plots of the Observed and Estimated TDS Concentration Retrieved by the Various ML Models Using Landsat 8 OLI Images.

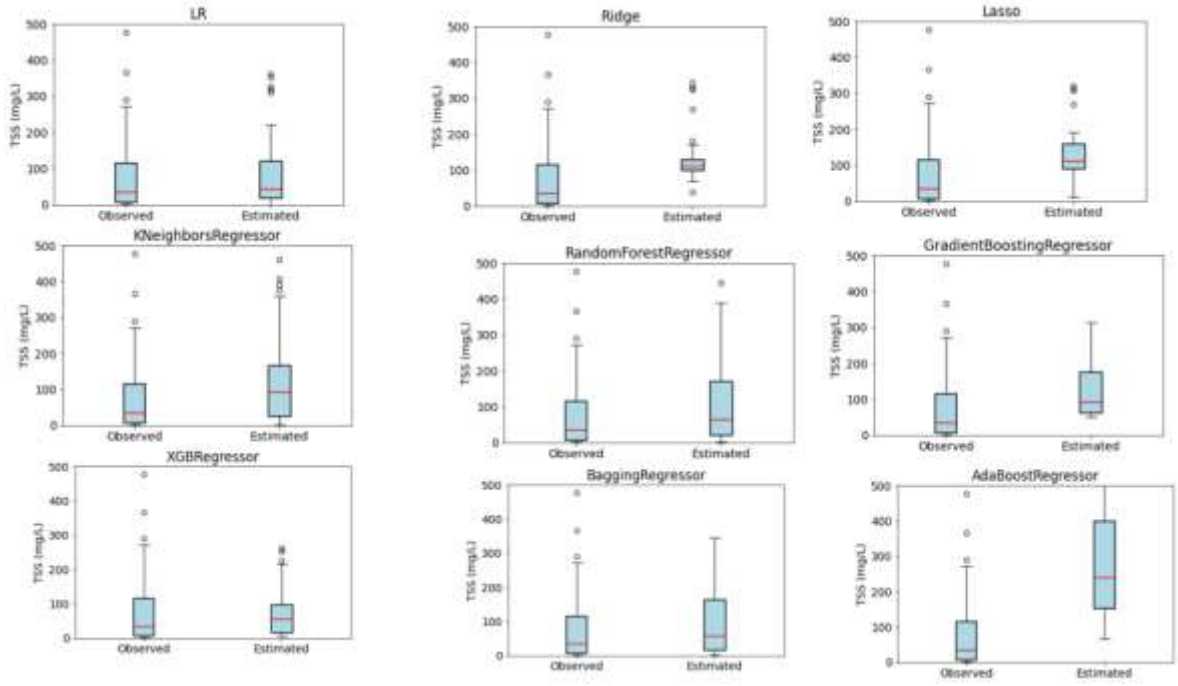


Figure 12. Box Plots of the Observed and Estimated TSS Concentration Retrieved by the Various ML Models Using Landsat 8 OLI Images.

 doi:10.20944/preprints202603.0403.v1

4. Conclusions

This research utilized satellite images from Sentinel-2 MSI and Landsat 8 OLI sensors and ML models to estimate TDS and TSS concentrations in the Colorado River System which supplies water to about 40 million people in the US and the Republic of Mexico. ML models used in the estimations ae categorized as standalone and ensemble models as described in Section 2.5.1 and 2.5.2. The river ,which is about 2,318 km long, faces water quality impairment worsened by the anthropogenic activities and the prolonged drought as it runs from rocky mountainous areas to the Gulf of California [1,2,64] hence the need to carry out continuous monitoring to ensure protection and sustenance of the limited water resources. Monitoring of the water quality has been conventionally performed using field sampling and laboratory experiments. Conducting field and laboratory analyses on river systems like the Colorado River is cost, labor, and time-intensive owing to length and the extent of the river. This research therefore applied a cost and labor efficient approach utilizing images capable to discriminating the spectral properties of the WQPs and ML algorithms techniques for the estimations of TDS and TSS concentrations in the Colorado River. Optimal features of the Landsat 8 OLI and Sentinel-2 MSI images and nine models consisting of four standalone models, namely LR, Lasso, Ridge, and KNN, and five ensemble models namely RF, GBM, XGBoost, Bagging, and Adaboost were used to retrieve optically active TSS and inactive WQPs [18,62,63,175].

Results produced for the TDS estimations show the superiority of ensemble models compared to standalone models for both datasets with the ensemble models being able to capture about the 81 to 100% variance in TDS across the three datasets (training, testing, and unseen datasets) with Sentinel-2 MSI images. The standalone models could only explain approximately 7 to 63% of the variance in TDS with images from the same sensor. The ensemble models could also account for about 32 to 97% of the variance in TDS with images from the Landsat 8 OLI while the standalone models were only able to explain about 11 to 31% TDS variance, across all the datasets. The superior performance of the ensemble models could be attributed to their intricacy in minimizing overfitting and generalization [126]. The superiority of the Sentinel-2 images in the estimation of TDS could also be attributed to their finer spatial and temporal resolutions [18,39] which is corroborated by the findings of researchers in [168] which found images from Sentinel-2B MSI sensor to be a superior estimator of nearly all the optically active and inactive parameters studied.

Models show similar but reduced performance in the estimations of TSS. While the Sentinel-2 images were able to account for about (-) 2 to 14% variations across the standalone models, Landsat 8 OLI accounted for approximately (-) 16 to 35% variations in TSS. The ensemble models show an improved performance with the spectral signatures explaining about 36 to 99% in TSS for the Sentinel-2 images. Landsat 8 OLI on the other hand saw a reduced explanation for the variance in TSS of about 11 to 90% across all the three datasets.

In general, the models perform better in estimating TDS compared to TSS due to the nature and the variability in these datasets in the river system. Unlike TDS which shows largely consistent variability in the river, TSS data presents underlying issues including spikes in the level of sediments in the river system contributed by the tributaries, limited amount, and possible underrepresentation of variations of TSS. ML is data-driven, and a lack of adequate and representative data for model training could contribute to subpar model performance [120].

In general, the XGBoost was found to be the most optimal model estimating TDS using images from both Sentinel-2 MSI and Landsat 8 OLI with performance values derived as 0.99, 26.52 mg/L, and 19.19 mg/L, respectively for R^2, RMSE, and MAE for Sentinel-2 images on the external validation dataset. The XGBoost yielded R^2, RMSE, and MAE of 0.97, 35.82 mg/L, and 27.90 mg/L, respectively.

The AdaBoost was found to be the best model for TSS estimations with values of 0.92, 29.48 mg/L, and 24.64 mg/L, respectively, for R^2, RMSE, and MAE for the Sentinel-2 image on the external validation dataset. The RF model was found to be the optimal model for TSS estimations with the Landsat 8 OLI with reported R^2, RMSE, and MAE of 0.90, 32.80 mg/L, and 22.91 mg/L, respectively, on the external validation dataset.

 doi:10.20944/preprints202603.0403.v1

The study proposed XGBoost with its hyperparameters and optimal features for the estimation of TDS for images from both sensors. The AdaBoost and RF together with their hyperparameters and the optimal features, respectively, for the Sentinel-2 MSI and Landsat 8 OLI as described in Section 2.5.3. The superior performance of ensemble models such as XGBoost could be a result of the underpinning algorithms, which combine several decision trees with weaker learner decision trees to resolve errors produced from prior trees, yielding improved and reliable estimations [60,126,172].

Researchers, water resource managers, and decision-makers can rely on these findings to carry out efficient water quality monitoring and protection efforts in waterbodies like the Colorado River System.

Further research could explore the impact of incorporation of large and representative datasets particularly for TSS on the overall reliability and generalization of the ML models. Future studies should also explore the impact of using data fusion, i.e., combining images from both optical sensors (Sentinel-2 and Landsat 8 OLI images) or microwave and optical sensors which have been observed additional details beyond what is offered by a single sensor [77,176] and therefore have the potential to enhance the overall model performance. Furthermore, future research could investigate the potential of deep learning models like convolutional neural networks and long short-term memory which has the potential to provide real-time estimations of WQPs [62].

Future research could also apply the established models, particularly the proposed one for TDS in a comparable water body to examine its overall robustness to TDS estimations.

Author Contributions: G.E.A. contributed to the conceptualization, methodology, analysis, and original and final writing of the manuscript. H.S. and S.A. both contributed to the conceptualization and writing—review and editing, and supervision of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data and codes used for this study are available upon request.

Acknowledgments: Appreciation to Charles Adjovu for their help in understanding some principles underlying machine learning modeling.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Rahaman, M.M.; Thakur, B.; Kalra, A.; Ahmad, S. Modeling of GRACE-Derived Groundwater Information in the Colorado River Basin. Hydrology 2019, 6, 19, doi:10.3390/hydrology6010019.
2. U.S. Department of the Interior; U.S. Geological Survey River Basins of the United States: The Colorado Available online: https://pubs.usgs.gov/gip/70039371/report.pdf (accessed on 17 July 2023).
3. Adjovu, G.E. Evaluating the Performance of A GIS-Based Tool for Delineating Swales Along Two Highways in Tennessee, Published by ProQuest LLC, Ann Arbor, MI, USA, 2020.
4. Shaikh, T.A.; Adjovu, G.E.; Stephen, H.; Ahmad, S. Impacts of Urbanization on Watershed Hydrology and Runoff Water Quality of a Watershed: A Review. In Proceedings of the World Environmental and Water Resources Congress 2023; 2023; Vol. 1, pp. 1271–1283.
5. Adjovu, G.E.; Ahmad, S.; Stephen, H. Analysis of Suspended Material in Lake Mead Using Remote Sensing Indices. World Environmental and Water Resources Congress 2021 2021, doi:https://doi.org/10.1061/9780784483466.069.
6. Adjovu, G.E.; Stephen, H.; Ahmad, S. Monitoring of Total Dissolved Solids Using Remote Sensing Band Reflectance and Salinity Indices: A Case Study of the Imperial County Section, AZ-CA, of the Colorado River. World Environmental and Water Resources Congress 2022 2022, doi:https://doi.org/10.1061/9780784484258.106.
7. Venkatesan, A.K.; Ahmad, S.; Batista, J.R.; Johnson, W.S. Total Dissolved Solids Contribution to the Colorado River Associated with the Growth of Las Vegas Valley. World Environmental and Water

 doi:10.20944/preprints202603.0403.v1

Resources Congress 2010: Challenges of Change - Proceedings of the World Environmental and Water Resources Congress 2010 2010, 3376–3385, doi:10.1061/41114(371)348.

8. U.S. EPA National Primary Drinking Water Guidelines. Epa 816-F-09-004 2009, 1, United States Environmental Protection Agency. 7p.
9. U.S. EPA 2018 Edition of the Drinking Water Standards and Health Advisories Tables Available online: https://www.epa.gov/system/files/documents/2022-01/dwtable2018.pdf (accessed on 12 November 2023).
10. Adjovu, G.E.; Stephen, H.; Ahmad, S. Spatial and Temporal Dynamics of Key Water Quality Parameters in a Thermal Stratified Lake Ecosystem : The Case Study of Lake Mead. Earth 2023, 4, 461–502, doi:https://doi.org/10.3390/earth4030025.
11. US EPA Developing Water Quality Criteria for Suspended and Bedded Sediments (SABS). 2019.
12. Shope, C.L.; Gerner, S.J. Assessment of Dissolved-Solids Loading to the Colorado River in the Paradox Basin between the Dolores River and Gypsum Canyon, Utah. U.S. Geological Survey, Scientific Investigations Report 2014–5031 2016, 1–18, doi:http://dx.doi.org/10.3133/sir20135229.
13. Cederberg, J.R.; Paretti, N. V.; Coes, A.L.; Hermosillo, E.; Lucia, A. Estimation of Dissolved-Solids Concentrations Using Continuous Water-Quality Monitoring and Regression Models at Four Sites in the Yuma Area , Arizona and California , January 2017 through March 2019. U.S. Geological Survey Scientific Investigations Report 2021, 26.
14. Tillman, F.D.; Day, N.K.; Miller, M.P.; Miller, O.L.; Rumsey, C.A.; Wise, D.R.; Longley, P.C.; McDonnell, M.C. A Review of Current Capabilities and Science Gaps in Water Supply Data, Modeling, and Trends for Water Availability Assessments in the Upper Colorado River Basin. Water (Switzerland) 2022, 14, doi:10.3390/w14233813.
15. Adjovu, G.E.; Stephen, H.; James, D.; Ahmad, S. Measurement of Total Dissolved Solids and Total Suspended Solids in Water Systems : A Review of the Issues , Conventional , and Remote Sensing Techniques. 2023.
16. Bureau of Reclamation Quality of Water Progress Report No. 24; United States, 2013;
17. Colorado River Basin Salinity Control Forum 2020 Review Water Quality Standards for Salinity, Colorado River System Available online: https://coloradoriversalinity.org/docs/2020 REVIEW - Final w appendices.pdf (accessed on 8 June 2023).
18. Adjovu, G.E.; Ali Shaikh, T.; Stephen, H.; Ahmad, S. Utilization of Machine Learning Models and Satellite Data for the Estimation of Total Dissolved Solids in the Colorado River System. In Proceedings of the World Environmental and Water Resources Congress 2023; American Society of Civil Engineers: Henderson, NV, 2023; Vol. 1, pp. 1147–1160.
19. Keeler, R. Colorado River Basin Salinity Control Forum Available online: https://www.coloradoriversalinity.org/ (accessed on 13 November 2023).
20. Nauman, T.W.; Ely, C.P.; Miller, M.P.; Duniway, M.C. Salinity Yield Modeling of the Upper Colorado River Basin Using 30-m Resolution Soil Maps and Random Forests. Water Resour. Res. 2019, 55, 4954–4973, doi:10.1029/2018WR024054.
21. Adjovu, G.E.; Stephen, H.; Ahmad, S. A Machine Learning Approach for the Estimation of Total Dissolved Solids Concentration in Lake Mead Using Electrical Conductivity and Temperature. Water (Basel). 2023, 15, doi:10.3390/w15132439.
22. Venkatesan, A.K.; Ahmad, S.; Johnson, W.; Batista, J.R. Salinity Reduction and Energy Conservation in Direct and Indirect Potable Water Reuse. Desalination 2011, 272, 120–127, doi:10.1016/j.desal.2011.01.007.
23. Imen, S.; Chang, N. Bin; Yang, Y.J. Developing the Remote Sensing-Based Early Warning System for Monitoring TSS Concentrations in Lake Mead. J. Environ. Manage. 2015, 160, 73–89, doi:10.1016/j.jenvman.2015.06.003.
24. Hajigholizadeh, M.; Moncada, A.; Kent, S.; Melesse, A.M. Land–Lake Linkage and Remote Sensing Application in Water Quality Monitoring in Lake Okeechobee, Florida, Usa. Land (Basel). 2021, 10, 1–17, doi:10.3390/land10020147.
25. Butler, B.A.; Ford, R.G. Evaluating Relationships Between Total Dissolved Solids (TDS) and Total Suspended Solids (TSS) in a Mining-Influenced Watershed. Mine Water Environ. 2018, 37, 18–30, doi:10.1007/s10230-017-0484-y.

 doi:10.20944/preprints202603.0403.v1

26. Sanjoto, T.B.; Elwafa, A.H.; Tjahjono, H.; Sidiq, W.A.B.N. Study of Total Suspended Solid Concentration Based on Doxaran Algorithm Using Landsat 8 Image in Coastal Water between Bodri River Estuary up to East Flood Canal Semarang City. IOP Conf. Ser. Earth Environ. Sci. 2020, 561, doi:10.1088/1755-1315/561/1/012053.
27. Lee, S.; Ahn, K.H. Monitoring of COD as an Organic Indicator in Waste Water and Treated Effluent by Fluorescence Excitation-Emission (FEEM) Matrix Characterization. Water Science and Technology 2004, 50, 57–63, doi:10.2166/wst.2004.0488.
28. Serafy, G.Y.H. El; Schaeffer, B.A.; Neely, M.; Spinosa, A.; Odermatt, D.; Weathers, K.C.; Baracchini, T.; Bouffard, D.; Carvalho, L.; Conmy, R.N.; et al. Integrating Inland and Coastal Water Quality Data for Actionable Knowledge. 2021, 1–24.
29. Zhao, J.; Zhang, F.; Chen, S.; Wang, C.; Chen, J.; Zhou, H.; Xue, Y. Remote Sensing Evaluation of Total Suspended Solids Dynamic with Markov Model: A Case Study of Inland Reservoir across Administrative Boundary in South China. Sensors 2020, 20, 6911, doi:10.3390/s20236911.
30. Adjovu, G.E.; Stephen, H.; James, D.; Ahmad, S. Overview of the Application of Remote Sensing in Effective Monitoring of Water Quality Parameters. Remote Sens. (Basel). 2023, 15, 1938, doi:10.3390/rs15071938.
31. U.S. Environmental Protection Agency NPDES Compliance Inspection Manual - Chapter 5 - Sampling. 2017.
32. Gholizadeh, M.H.; Melesse, A.M.; Reddi, L. A Comprehensive Review on Water Quality Parameters Estimation Using Remote Sensing Techniques. Sensors (Switzerland) 2016, 16, doi:10.3390/s16081298.
33. Xu, H.; Xu, G.; Wen, X.; Hu, X.; Wang, Y. Lockdown Effects on Total Suspended Solids Concentrations in the Lower Min River (China) during COVID-19 Using Time-Series Remote Sensing Images. International Journal of Applied Earth Observation and Geoinformation 2021, 98, doi:10.1016/j.jag.2021.102301.
34. Sa'ad, F.N.A.; Tahir, M.S.; Jemily, N.H.B.; Ahmad, A.; Amin, A.R.M. Monitoring Total Suspended Sediment Concentration in Spatiotemporal Domain over Teluk Lipat Utilizing Landsat 8 (OLI). Applied Sciences (Switzerland) 2021, 11, doi:10.3390/app11157082.
35. Aljoborey, A.D.A.; Abdulhay, H.S. Estimating Total Dissolved Solids and Total Suspended Solids in Mosul Dam Lake in Situ and Using Remote Sensing Technique. Periodicals of Engineering and Natural Sciences 2019, 7, 1755–1767, doi:10.21533/pen.v7i4.832.
36. Gholizadeh, M.H.; Melesse, A.M.; Reddi, L. Spaceborne and Airborne Sensors in Water Quality Assessment. Int. J. Remote Sens. 2016, 37, 3143–3180, doi:10.1080/01431161.2016.1190477.
37. Malahlela, O.E. Spatio-Temporal Assessment of Inland Surface Water Quality Using Remote Sensing Data in the Wake of Changing Climate. IOP Conf. Ser. Earth Environ. Sci. 2019, 227, doi:10.1088/1755-1315/227/6/062012.
38. Maliki, A.A.A.; Chabuk, A.; Sultan, M.A.; Hashim, B.M.; Hussain, H.M.; Al-Ansari, N. Estimation of Total Dissolved Solids in Water Bodies by Spectral Indices Case Study: Shatt al-Arab River. Water Air Soil Pollut. 2020, 231, doi:10.1007/s11270-020-04844-z.
39. Guo, H.; Huang, J.J.; Chen, B.; Guo, X.; Singh, V.P. A Machine Learning-Based Strategy for Estimating Non-Optically Active Water Quality Parameters Using Sentinel-2 Imagery. Int. J. Remote Sens. 2021, 42, 1841–1866, doi:10.1080/01431161.2020.1846222.
40. Zhang, C.; Liu, Y.; Chen, X.; Gao, Y. Estimation of Suspended Sediment Concentration in the Yangtze Main Stream Based on Sentinel-2 MSI Data. Remote Sens. (Basel). 2022, 14, doi:10.3390/rs14184446.
41. Dekker, A.G.; Zamurovic, Ž.; Hoogenboom, H.J.; Peters, S.W.M. Remote Sensing, Ecological Water Quality Modelling and in Situ Measurements: A Case Study in Shallow Lakes. Hydrological Sciences Journal 1996, 41, 531–547, doi:10.1080/02626669609491524.
42. Yang, H.; Kong, J.; Hu, H.; Du, Y.; Gao, M.; Chen, F. A Review of Remote Sensing for Water Quality Retrieval: Progress and Challenges. Remote Sens. (Basel). 2022, 14, 1770, doi:10.3390/rs14081770.
43. Dube, T.; Mutanga, O.; Seutloali, K.; Adelabu, S.; Shoko, C. Water Quality Monitoring in Sub-Saharan African Lakes: A Review of Remote Sensing Applications. Afr. J. Aquat. Sci. 2015, 40, 1–7, doi:10.2989/16085914.2015.1014994.

 doi:10.20944/preprints202603.0403.v1

44. Wang, J.J.; Lu, X.X. Estimation of Suspended Sediment Concentrations Using Terra MODIS: An Example from the Lower Yangtze River, China. Science of the Total Environment 2010, 408, 1131–1138, doi:10.1016/j.scitotenv.2009.11.057.
45. Kong, J.L.; Sun, X.M.; Wong, D.W.; Chen, Y.; Yang, J.; Yan, Y.; Wang, L.X. A Semi-Analytical Model for Remote Sensing Retrieval of Suspended Sediment Concentration in the Gulf of Bohai, China. Remote Sens. (Basel). 2015, 7, 5373–5397, doi:10.3390/rs70505373.
46. Jally, S.K.; Mishra, A.K.; Balabantaray, S. Retrieval of Suspended Sediment Concentration of the Chilika Lake, India Using Landsat-8 OLI Satellite Data. Environ. Earth Sci. 2021, 80, 1–18, doi:10.1007/s12665-021-09581-y.
47. Marinho, R.R.; Harmel, T.; Martinez, J.M.; Junior, N.P.F. Spatiotemporal Dynamics of Suspended Sediments in the Negro River, Amazon Basin, from in Situ and Sentinel-2 Remote Sensing Data. ISPRS Int. J. Geoinf. 2021, 10, doi:10.3390/ijgi10020086.
48. Chawla, I.; Karthikeyan, L.; Mishra, A.K. A Review of Remote Sensing Applications for Water Security: Quantity, Quality, and Extremes. J. Hydrol. (Amst). 2020, 585, 124826, doi:10.1016/j.jhydrol.2020.124826.
49. Novoa, S.; Doxaran, D.; Ody, A.; Vanhellemont, Q.; Lafon, V.; Lubac, B.; Gernez, P. Atmospheric Corrections and Multi-Conditional Algorithm for Multi-Sensor Remote Sensing of Suspended Particulate Matter in Low-to-High Turbidity Levels Coastal Waters. Remote Sens. (Basel). 2017, 9, doi:10.3390/rs9010061.
50. Mao, Z.; Chen, J.; Pan, D.; Tao, B.; Zhu, Q. A Regional Remote Sensing Algorithm for Total Suspended Matter in the East China Sea. Remote Sens. Environ. 2012, 124, 819–831, doi:10.1016/j.rse.2012.06.014.
51. Petus, C.; Chust, G.; Gohin, F.; Doxaran, D.; Froidefond, J.M.; Sagarminaga, Y. Estimating Turbidity and Total Suspended Matter in the Adour River Plume (South Bay of Biscay) Using MODIS 250-m Imagery. Cont. Shelf Res. 2010, 30, 379–392, doi:10.1016/j.csr.2009.12.007.
52. Kupssinskü, L.S.; Guimarães, T.T.; De Souza, E.M.; Zanotta, D.C.; Veronez, M.R.; Gonzaga, L.; Mauad, F.F. A Method for Chlorophyll-a and Suspended Solids Prediction through Remote Sensing and Machine Learning. Sensors (Switzerland) 2020, 20, doi:10.3390/s20072125.
53. Peterson, K.T.; Sagan, V.; Sidike, P.; Cox, A.L.; Martinez, M. Suspended Sediment Concentration Estimation from Landsat Imagery along the Lower Missouri and Middle Mississippi Rivers Using an Extreme Learning Machine. Remote Sens. (Basel). 2018, 10, doi:10.3390/rs10101503.
54. Di Napoli, M.; Carotenuto, F.; Cevasco, A.; Confuorto, P.; Di Martire, D.; Firpo, M.; Pepe, G.; Raso, E.; Calcaterra, D. Machine Learning Ensemble Modelling as a Tool to Improve Landslide Susceptibility Mapping Reliability. Landslides 2020, 17, 1897–1914, doi:10.1007/s10346-020-01392-9.
55. Wakjira, T.G.; Rahmzadeh, A.; Alam, M.S.; Tremblay, R. Explainable Machine Learning Based Efficient Prediction Tool for Lateral Cyclic Response of Post-Tensioned Base Rocking Steel Bridge Piers. Structures 2022, 44, 947–964, doi:10.1016/j.istruc.2022.08.023.
56. Saberioon, M.; Brom, J.; Nedbal, V.; Souček, P.; Císař, P. Chlorophyll-a and Total Suspended Solids Retrieval and Mapping Using Sentinel-2A and Machine Learning for Inland Waters. Ecol. Indic. 2020, 113, doi:10.1016/j.ecolind.2020.106236.
57. Sagan, V.; Peterson, K.T.; Maimaitijiang, M.; Sidike, P.; Sloan, J.; Greeling, B.A.; Maalouf, S.; 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, doi:10.1016/j.earscirev.2020.103187.
58. Hope, T.M.H. Linear Regression. In Machine Learning; Elsevier, 2020; pp. 67–81 ISBN 9780128157398.
59. Li, S.; Song, K.; Wang, S.; Liu, G.; Wen, Z.; Shang, Y.; Lyu, L.; Chen, F.; Xu, S.; Tao, H.; et al. Quantification of Chlorophyll-a in Typical Lakes across China Using Sentinel-2 MSI Imagery with Machine Learning Algorithm. Science of the Total Environment 2021, 778, 146271, doi:10.1016/j.scitotenv.2021.146271.
60. Singh, A.K. Impact of the Coronavirus Pandemic on Las Vegas Strip Gaming Revenue. The Journal of Gambling Business and Economics 2021, 14, 49–62, doi:10.5750/jgbe.v14i2.1965.
61. Wang, X.; Zhang, F.; Ding, J. Evaluation of Water Quality Based on a Machine Learning Algorithm and Water Quality Index for the Ebinur Lake Watershed, China. Sci. Rep. 2017, 7, 1–18, doi:10.1038/s41598-017-12853-y.

 doi:10.20944/preprints202603.0403.v1

62. Leggesse, E.S.; Zimale, F.A.; Sultan, D.; Enku, T.; Srinivasan, R.; Tilahun, S.A. Predicting Optical Water Quality Indicators from Remote Sensing Using Machine Learning Algorithms in Tropical Highlands of Ethiopia. Hydrology 2023, 10, 110, doi:10.3390/hydrology10050110.
63. Mejía Ávila, D.; Torres-Bejarano, F.; Martínez Lara, Z. Spectral Indices for Estimating Total Dissolved Solids in Freshwater Wetlands Using Semi-Empirical Models. A Case Study of Guartinaja and Momil Wetlands. Int. J. Remote Sens. 2022, 43, 2156–2184, doi:10.1080/01431161.2022.2057205.
64. Adjovu, G.E.; Stephen, H.; Ahmad, S. Spatiotemporal Variability in Total Dissolved Solids and Total Suspended Solids along the Colorado River. Hydrology 2023, 10, 1–27, doi:10.3390/hydrology10060125.
65. Rumsey, C.A.; Miller, O.; Hirsch, R.M.; Marston, T.M.; Susong, D.D. Substantial Declines in Salinity Observed Across the Upper Colorado River Basin During the 20th Century, 1929–2019. Water Resour. Res. 2021, 57, 1–21, doi:10.1029/2020WR028581.
66. Tillman, F.D.; Anning, D.W. A Data Reconnaissance on the Effect of Suspended-Sediment Concentrations on Dissolved-Solids Concentrations in Rivers and Tributaries in the Upper Colorado River Basin. J. Hydrol. (Amst). 2014, 519, 1020–1030, doi:10.1016/j.jhydrol.2014.08.020.
67. Lukas, J.; Payton, E. Colorado River Basin Climate and Hydrology: State of the Science. Western Water Assessment; Western Water Assessment, University of Colorado Boulder., 2020; ISBN 1219950920.
68. Hoerling, M.; Barsugli, J.; Livneh, B.; Eischeid, J.; Quan, X.; Badger, A. Causes for the Century-Long Decline in Colorado River Flow. J. Clim. 2019, 32, 8181–8203, doi:10.1175/JCLI-D-19-0207.1.
69. Anning, D.W.; Coes, A.L.; Mason, J.P. Conceptual and Numerical Models of Dissolved Solids in the Colorado River , Hoover Dam to Imperial Dam , and Parker Dam to Imperial Dam , Arizona , California , and Nevada. 2018.
70. Earp, K.J.; Moreo, M.T. Evaporation from Lake Mead and Lake Mohave , Nevada and Arizona , 2010 – 2019. 2021, 2010–2019.
71. Morfín, O. Effects of System Conservation on Salinity in Lake Mead Available online: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://www.multi-statesalinitycoalition.com/wp-content/uploads/2017-Morfin.pdf.
72. Venkatesan, A.K.; Ahmad, S.; Johnson, W.; Batista, J.R. Systems Dynamic Model to Forecast Salinity Load to the Colorado River Due to Urbanization within the Las Vegas Valley. Science of the Total Environment 2011, 409, 2616–2625, doi:10.1016/j.scitotenv.2011.03.018.
73. U.S. Department of the Interior; U.S. Geological Survey National Water Information System: Web Interface Available online: https://waterdata.usgs.gov/nwis/qw (accessed on 22 April 2023).
74. Chris Neigh; Taylor, M.P. LANDSAT 8 Available online: https://landsat.gsfc.nasa.gov/satellites/landsat-8/.
75. Phiri, D.; Simwanda, M.; Salekin, S.; Nyirenda, V.R.; Murayama, Y.; Ranagalage, M. Sentinel-2 Data for Land Cover/Use Mapping: A Review. Remote Sens. (Basel). 2020, 12, doi:10.3390/rs12142291.
76. European Space Agency About Copernicus Sentinel-2. 2020, 1.
77. Shrestha, B.; Stephen, H.; Ahmad, S. Impervious Surfaces Mapping at City Scale by Fusion of Radar and Optical Data through a Random Forest Classifier. Remote Sens. (Basel). 2021, 13, doi:10.3390/rs13153040.
78. European Space Agency Sentinel-2 User Handbook Available online: https://sentinel.esa.int/documents/247904/685211/sentinel-2_user_handbook (accessed on 3 July 2023).
79. Ferdous, J.; Ur Rahman, M.T.; Ghosh, S.K. Detection of Total Dissolved Solids from Landsat 8 OLI Image in Coastal Bangladesh. 2019, 35–44, doi:10.17501/2513258x.2019.3103.
80. 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, doi:10.1007/s10661-018-6506-9.
81. United States Geology Survey (USGS) What Are the Band Designations for the Landsat Satellites? Available online: https://www.usgs.gov/faqs/what-are-band-designations-landsat-satellites (accessed on 3 July 2023).
82. Elmahdy, S.I.; Mohamed, M.M. Land Use/Land Cover Change Impact on Groundwater Quantity and Quality: A Case Study of Ajman Emirate, the United Arab Emirates, Using Remote Sensing and GIS. Arabian Journal of Geosciences 2016, 9, doi:10.1007/s12517-016-2725-y.

 doi:10.20944/preprints202603.0403.v1

83. Page, B.P.; Olmanson, L.G.; Mishra, D.R. A Harmonized Image Processing Workflow Using Sentinel-2/MSI and Landsat-8/OLI for Mapping Water Clarity in Optically Variable Lake Systems. Remote Sens. Environ. 2019, 231, 111284, doi:10.1016/j.rse.2019.111284.
84. Pahlevan, N.; Sarkar, S.; Franz, B.A.; Balasubramanian, S. V.; He, J. Sentinel-2 MultiSpectral Instrument (MSI) Data Processing for Aquatic Science Applications: Demonstrations and Validations. Remote Sens. Environ. 2017, 201, 47–56, doi:10.1016/j.rse.2017.08.033.
85. Li, J.; Roy, D.P. A Global Analysis of Sentinel-2a, Sentinel-2b and Landsat-8 Data Revisit Intervals and Implications for Terrestrial Monitoring. Remote Sens. (Basel). 2017, 9, doi:10.3390/rs9090902.
86. Niroumand-Jadidi, M.; Bovolo, F.; Bresciani, M.; Gege, P.; Giardino, C. Water Quality Retrieval from Landsat-9 (OLI-2) Imagery and Comparison to Sentinel-2. Remote Sens. (Basel). 2022, 14, 1–18, doi:10.3390/rs14184596.
87. Amani, M.; Ghorbanian, A.; Ahmadi, S.A.; Kakooei, M.; Moghimi, A.; Mirmazloumi, S.M.; Moghaddam, S.H.A.; Mahdavi, S.; Ghahremanloo, M.; Parsian, S.; et al. Google Earth Engine Cloud Computing Platform for Remote Sensing Big Data Applications: A Comprehensive Review. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2020, 13, 5326–5350, doi:10.1109/JSTARS.2020.3021052.
88. Mohseni, F.; Saba, F.; Mirmazloumi, S.M.; Amani, M.; Mokhtarzade, M.; Jamali, S.; Mahdavi, S. Ocean Water Quality Monitoring Using Remote Sensing Techniques: A Review. Mar. Environ. Res. 2022, 180, 105701, doi:10.1016/j.marenvres.2022.105701.
89. Owusu, C.; Snigdha, N.J.; Martin, M.T.; Kalyanapu, A.J. PyGEE-SWToolbox: A Python Jupyter Notebook Toolbox for Interactive Surface Water Mapping and Analysis Using Google Earth Engine. Sustainability (Switzerland) 2022, 14, doi:10.3390/su14052557.
90. Pereira, O.J.R.; Merino, E.R.; Montes, C.R.; Barbiero, L.; Rezende-Filho, A.T.; Lucas, Y.; Melfi, A.J. Estimating Water PH Using Cloud-Based Landsat Images for a New Classification of the Nhecolândia Lakes (Brazilian Pantanal). Remote Sens. (Basel). 2020, 12, doi:10.3390/rs12071090.
91. Topp, S.N.; Pavelsky, T.M.; Jensen, D.; Simard, M.; Ross, M.R.V. Research Trends in the Use of Remote Sensing for Inland Water Quality Science: Moving towards Multidisciplinary Applications. Water (Switzerland) 2020, 12, 1–34, doi:10.3390/w12010169.
92. Kontopoulou, E.; Kolokoussis, P.; Karantzalos, K. Water Quality Estimation in Greek Lakes from Landsat 8 Multispectral Satellite Data. European Water 2017, 58, 191–196.
93. Waxter, M.T. Analysis of Landsat Satellite Data to Monitor Water Quality Parameters in Tenmile Lake , Oregon. Civil and Environmental Engineering Master's Project Reports 2014.
94. Hossain, A.K.M.A.; Mathias, C.; Blanton, R. Remote Sensing of Turbidity in the Tennessee River Using Landsat 8 Satellite. Remote Sens. (Basel). 2021, 13, 3785, doi:10.3390/rs13183785.
95. Japitana, M. V.; Burce, M.E.C. A Satellite-Based Remote Sensing Technique for Surface Water Quality Estimation. Engineering, Technology & Applied Science Research 2019, 9, 3965–3970, doi:10.48084/etasr.2664.
96. Barrett, D.C.; Frazier, A.E. Automated Method for Monitoring Water Quality Using Landsat Imagery. Water (Switzerland) 2016, 8, 1–14, doi:10.3390/W8060257.
97. Kapalanga, T.S.; Hoko, Z.; Gumindoga, W.; Chikwiramakomo, L. Remote-Sensing-Based Algorithms for Water Quality Monitoring in Olushandja Dam, North-Central Namibia. Water Supply 2021, 21, 1878–1894, doi:10.2166/ws.2020.290.
98. Robinson, A.P.; Hamann, J.D. Forest Analytics with R; 2011; ISBN 9781441977618.
99. Chen, F. Missing No More: Using the MCMC Procedure to Model Missing Data; 2013;
100. Poudel, U.; Stephen, H.; Ahmad, S. Evaluating Irrigation Performance and Water Productivity Using EEFlux ET and NDVI. Sustainability (Switzerland) 2021, 13, doi:10.3390/su13147967.
101. Bigaignon, L.; Fieuzal, R.; Delon, C.; Tallec, T. Combination of Two Methodologies, Artificial Neural Network and Linear Interpolation, to Gap-Fill Daily Nitrous Oxide Flux Measurements. Agric. For. Meteorol. 2020, 291, 108037, doi:10.1016/j.agrformet.2020.108037.
102. United States Geology Survey (USGS) Why Are Negative Values Observed over Water in Some Landsat Surface Reflectance Products?

 doi:10.20944/preprints202603.0403.v1

103. Lubbe, S.; Filzmoser, P.; Templ, M. Comparison of Zero Replacement Strategies for Compositional Data with Large Numbers of Zeros. Chemometrics and Intelligent Laboratory Systems 2021, 210, doi:10.1016/j.chemolab.2021.104248.
104. Claytex Services Ltd How to Avoid Divide by Zero Errors Dymola Refactor the Problem Available online: https://www.claytex.com/tech-blog/how-to-avoid-divide-by-zero-errors/ (accessed on 15 June 2023).
105. Ferdous, J.; Rahman, M.T.U. Developing an Empirical Model from Landsat Data Series for Monitoring Water Salinity in Coastal Bangladesh. J. Environ. Manage. 2020, 255, 109861, doi:10.1016/j.jenvman.2019.109861.
106. Abdelmalik, K.W. Role of Statistical Remote Sensing for Inland Water Quality Parameters Prediction. Egyptian Journal of Remote Sensing and Space Science 2018, 21, 193–200, doi:10.1016/j.ejrs.2016.12.002.
107. Tran1, P.H.; Nguyen, A.K.; Liou, Y.-A.; Hoang, P.P.; Thanh, H.N. Estimation of Salinity Intrusion by Using Landsat 8 OLI Data in The Mekong Delta, Vietnam. Prog. Earth Planet. Sci. 2020, 7, 1, doi:10.1186/s40645-019-0311-0.
108. Montalvo, L.G. Spectral Analysis of Suspended Material in Coastal Waters : A Comparison between Band Math Equations. 2010, 1–6.
109. Arisanty, D.; Nur Saputra, A. Remote Sensing Studies of Suspended Sediment Concentration Variation in Barito Delta. IOP Conf. Ser. Earth Environ. Sci. 2017, 98, 1–6, doi:10.1088/1755-1315/98/1/012058.
110. Hossain, A.K.M.A.; Chao, X.; Jia, Y. Development of Remote Sensing Based Index for Estimating / Mapping Suspended Sediment Concentration in River and Lake Environments. 8th International Symposium on Ecohydraulics (ISE) 2010, 578–585.
111. Sudheer, K.P.; Chaubey, I.; Garg, V. Lake Water Quality Assessment from Landsat Thematic Mapper Data Using Neural Network: An Approach to Optimal Band Combination Selection. J. Am. Water Resour. Assoc. 2006, 42, 1683–1695, doi:10.1111/j.1752-1688.2006.tb06029.x.
112. Sahoo, D.P.; Sahoo, B.; Tiwari, M.K. MODIS-Landsat Fusion-Based Single-Band Algorithms for TSS and Turbidity Estimation in an Urban-Waste-Dominated River Reach. Water Res. 2022, 224, 119082, doi:10.1016/j.watres.2022.119082.
113. Soudani, K.; François, C.; le Maire, G.; Le Dantec, V.; Dufrêne, E. Comparative Analysis of IKONOS, SPOT, and ETM+ Data for Leaf Area Index Estimation in Temperate Coniferous and Deciduous Forest Stands. Remote Sens. Environ. 2006, 102, 161–175, doi:10.1016/j.rse.2006.02.004.
114. Skiena, S. Lecture 14: Correlation and Autocorrelation Available online: www3.cs.stonybrook.edu/~skiena/691/lectures/lecture14.pdf (accessed on 15 June 2023).
115. Philipson, P.; Kratzer, S.; Ben Mustapha, S.; Strömbeck, N.; Stelzer, K. Satellite-Based Water Quality Monitoring in Lake Vänern, Sweden. Int. J. Remote Sens. 2016, 37, 3938–3960, doi:10.1080/01431161.2016.1204480.
116. Ghasemi, A.; Zahediasl, S. Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. Int. J. Endocrinol. Metab. 2012, 10, 486–489, doi:10.5812/ijem.3505.
117. Lantz, B. Machine Learning with R Learn; Packt Publishing Ltd: Birmingham, UK, 2013; ISBN 978-1-78216-214-8.
118. Moeini, M.; Shojaeizadeh, A.; Geza, M. Supervised Machine Learning for Estimation of Total Suspended Solids in Urban Watersheds. Water (Switzerland) 2021, 13, doi:10.3390/w13020147.
119. Kaur, H.; Malhi, A.K.; Pannu, H.S. Machine Learning Ensemble for Neurological Disorders. Neural Comput. Appl. 2020, 32, 12697–12714, doi:10.1007/s00521-020-04720-1.
120. Hafeez, S.; Wong, M.S.; Ho, H.C.; Nazeer, M.; Nichol, J.; Abbas, S.; Tang, D.; Lee, K.H.; Pun, L. Comparison of Machine Learning Algorithms for Retrieval of Water Quality Indicators in Case-II Waters: A Case Study of Hong Kong. Remote Sens. (Basel). 2019, 11, 617, doi:10.3390/rs11060617.
121. Kavitha, S.; Varuna, S.; Ramya, R. A Comparative Analysis on Linear Regression and Support Vector Regression. Proceedings of 2016 Online International Conference on Green Engineering and Technologies, IC-GET 2016 2017, doi:10.1109/GET.2016.7916627.
122. Rong, S.; Bao-Wen, Z. The Research of Regression Model in Machine Learning Field. MATEC Web of Conferences 2018, 176, 8–11, doi:10.1051/matecconf/201817601033.
123. Grus, J. Data Science from Scratch; O'Reilly Media, 2019; Vol. 1542; ISBN 9781491901427.

 doi:10.20944/preprints202603.0403.v1

124. Phyo, P.P.; Byun, Y.C.; Park, N. Short-Term Energy Forecasting Using Machine-Learning-Based Ensemble Voting Regression. Symmetry (Basel). 2022, 14, 1–13, doi:10.3390/sym14010160.
125. Alexei Botchkarev Performance Metrics (Error Measures) in Machine Learning Regression, Forecasting and Prognostics: Properties and Typology. J. Chem. Inf. Model. 1981, 53, 1689–1699.
126. Zounemat-Kermani, M.; Batelaan, O.; Fadaee, M.; Hinkelmann, R. Ensemble Machine Learning Paradigms in Hydrology: A Review. J. Hydrol. (Amst). 2021, 598, 126266, doi:10.1016/j.jhydrol.2021.126266.
127. Scikit Learn Hyperparameter Tuning Available online: https://inria.github.io/scikit-learn-mooc/python_scripts/ensemble_hyperparameters.html.
128. Erdebilli, B.; Devrim-İçtenbaş, B. Ensemble Voting Regression Based on Machine Learning for Predicting Medical Waste: A Case from Turkey. Mathematics 2022, 10, 7–9, doi:10.3390/math10142466.
129. Freund, Y.; Schapire, R.E. Experiments with a New Boosting Algorithm. Proceedings of the 13th International Conference on Machine Learning 1996, 148–156, doi:10.1.1.133.1040.
130. Livingston, F. Implementation of Breiman's Random Forest Machine Learning Algorithm. Machine Learning Journal Paper 2005, 1–13.
131. Rodriguez-Galiano, V.; Sanchez-Castillo, M.; Chica-Olmo, M.; Chica-Rivas, M. Machine Learning Predictive Models for Mineral Prospectivity: An Evaluation of Neural Networks, Random Forest, Regression Trees and Support Vector Machines. Ore Geol. Rev. 2015, 71, 804–818, doi:10.1016/j.oregeorev.2015.01.001.
132. Tillman, F.D.; Anning, D.W.; Heilman, J.A.; Buto, S.G.; Miller, M.P. Managing Salinity in Upper Colorado River Basin Streams: Selecting Catchments for Sediment Control Efforts Using Watershed Characteristics and Random Forests Models. Water (Switzerland) 2018, 10, 1–17, doi:10.3390/w10060676.
133. Rocca, J. Ensemble Methods: Bagging, Boosting and Stacking Available online: https://towardsdatascience.com/ensemble-methods-bagging-boosting-and-stacking-c9214a10a205.
134. Wolff, S.; O'Donncha, F.; Chen, B. Statistical and Machine Learning Ensemble Modelling to Forecast Sea Surface Temperature. Journal of Marine Systems 2020, 208, 103347, doi:10.1016/j.jmarsys.2020.103347.
135. Bedi, S.; Samal, A.; Ray, C.; Snow, D. Comparative Evaluation of Machine Learning Models for Groundwater Quality Assessment; 2020; Vol. 192; ISBN 1066102008695.
136. Chen, J.; de Hoogh, K.; Gulliver, J.; Hoffmann, B.; Hertel, O.; Ketzel, M.; Bauwelinck, M.; van Donkelaar, A.; Hvidtfeldt, U.A.; Katsouyanni, K.; et al. A Comparison of Linear Regression, Regularization, and Machine Learning Algorithms to Develop Europe-Wide Spatial Models of Fine Particles and Nitrogen Dioxide. Environ. Int. 2019, 130, doi:10.1016/j.envint.2019.104934.
137. Zhang, C.; Ma, Y.; Eds. Ensemble Machine Learning; Zhang, C., Ma, Y., Eds.; Springer US: Boston, MA, 2012; ISBN 978-1-4419-9325-0.
138. Avijeet Biswal Bagging in Machine Learning: Step to Perform and Its Advantages Available online: https://www.simplilearn.com/tutorials/machine-learning-tutorial/bagging-in-machine-learning.
139. Brajer, N.; Cozzi, B.; Gao, M.; Nichols, M.; Revoir, M.; Balu, S.; Futoma, J.; Bae, J.; Setji, N.; Hernandez, A.; et al. Prospective and External Evaluation of a Machine Learning Model to Predict In-Hospital Mortality of Adults at Time of Admission. JAMA Netw. Open 2020, 3, doi:10.1001/jamanetworkopen.2019.20733.
140. Ahmad, S.; Kalra, A.; Stephen, H. Estimating Soil Moisture Using Remote Sensing Data: A Machine Learning Approach. Adv. Water Resour. 2010, 33, 69–80, doi:10.1016/j.advwatres.2009.10.008.
141. Banadkooki, F.B.; Ehteram, M.; Panahi, F.; Sh. Sammen, S.; Othman, F.B.; EL-Shafie, A. Estimation of Total Dissolved Solids (TDS) Using New Hybrid Machine Learning Models. J. Hydrol. (Amst). 2020, 587, 124989, doi:10.1016/j.jhydrol.2020.124989.
142. Eertink, J.J.; Heymans, M.W.; Zwezerijnen, G.J.C.; Zijlstra, J.M.; de Vet, H.C.W.; Boellaard, R. External Validation: A Simulation Study to Compare Cross-Validation versus Holdout or External Testing to Assess the Performance of Clinical Prediction Models Using PET Data from DLBCL Patients. EJNMMI Res. 2022, 12, 4–11, doi:10.1186/s13550-022-00931-w.
143. Wakjira, T.G.; Abushanab, A.; Ebead, U.; Alnahhal, W. FAI: Fast, Accurate, and Intelligent Approach and Prediction Tool for Flexural Capacity of FRP-RC Beams Based on Super-Learner Machine Learning Model. Mater. Today Commun. 2022, 33, 104461, doi:10.1016/j.mtcomm.2022.104461.

 doi:10.20944/preprints202603.0403.v1

144. Kutty, A.A.; Wakjira, T.G.; Kucukvar, M.; Abdella, G.M.; Onat, N.C. Urban Resilience and Livability Performance of European Smart Cities: A Novel Machine Learning Approach. J. Clean. Prod. 2022, 378, 134203, doi:10.1016/j.jclepro.2022.134203.
145. Leigh, C.; Kandanaarachchi, S.; McGree, J.M.; Hyndman, R.J.; Alsibai, O.; Mengersen, K.; Peterson, E.E. Predicting Sediment and Nutrient Concentrations from High-Frequency Water-Quality Data. PLoS One 2019, 14, 1–22, doi:10.1371/journal.pone.0215503.
146. Mahanty, B.; Lhamo, P.; Sahoo, N.K. Inconsistency of PCA-Based Water Quality Index – Does It Reflect the Quality? Science of the Total Environment 2023, 866, 161353, doi:10.1016/j.scitotenv.2022.161353.
147. Jung, K.; Bae, D.H.; Um, M.J.; Kim, S.; Jeon, S.; Park, D. Evaluation of Nitrate Load Estimations Using Neural Networks and Canonical Correlation Analysis with K-Fold Cross-Validation. Sustainability (Switzerland) 2020, 12, doi:10.3390/SU12010400.
148. Mamat, N.; Hamzah, M.F.; Jaafar, 1Othman Hybrid Support Vector Regression Model and K-Fold Cross Validation for Water Quality Index Prediction in Langat River, Malaysia. 2021.
149. Normawati, D.; Ismi, D.P. K-Fold Cross Validation for Selection of Cardiovascular Disease Diagnosis Features by Applying Rule-Based Datamining. Signal and Image Processing Letters 2019, 1, 23–35, doi:10.31763/simple.v1i2.3.
150. Wakjira, T.G.; Ibrahim, M.; Ebead, U.; Alam, M.S. Explainable Machine Learning Model and Reliability Analysis for Flexural Capacity Prediction of RC Beams Strengthened in Flexure with FRCM. Eng. Struct. 2022, 255, 113903, doi:10.1016/j.engstruct.2022.113903.
151. Abba, S.I.; Linh, N.T.T.; Abdullahi, J.; Ali, S.I.A.; Pham, Q.B.; Abdulkadir, R.A.; Costache, R.; Nam, V.T.; Anh, D.T. Hybrid Machine Learning Ensemble Techniques for Modeling Dissolved Oxygen Concentration. IEEE Access 2020, 8, 157218–157237, doi:10.1109/ACCESS.2020.3017743.
152. Sciikit Learn Sklearn.Model_selection.KFold Available online: https://scikit-learn.org/stable/modules/generated/sklearn.model_selection.KFold.html.
153. Wang, Z.; Lei, Y.; Cui, H.; Miao, H.; Zhang, D.; Wu, Z.; Liu, G. Enhanced RBF Neural Network Metamodelling Approach Assisted by Sliced Splitting-Based K-Fold Cross-Validation and Its Application for the Stiffened Cylindrical Shells. Aerosp. Sci. Technol. 2022, 124, 107534, doi:10.1016/j.ast.2022.107534.
154. VanderPlas, J. Python Data Science Handbook; 2019; Vol. 53; ISBN 9788578110796.
155. da Silva, M.G.; de Aguiar Netto, A. de O.; de Jesus Neves, R.J.; do Vasco, A.N.; Almeida, C.; Faccioli, G.G. Sensitivity Analysis and Calibration of Hydrological Modeling of the Watershed Northeast Brazil. J. Environ. Prot. (Irvine, Calif). 2015, 06, 837–850, doi:10.4236/jep.2015.68076.
156. Moriasi, D.N.; Arnold, J.G.; Van Liew, M.W.; Bingner, R.L.; Harmel, R.D.; Veith, T.L. Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. Trans. ASABE 2007, 50, 885–900, doi:10.13031/2013.23153.
157. Sun, K.; Rajabtabar, M.; Samadi, S.Z.; Rezaie-Balf, M.; Ghaemi, A.; Band, S.S.; Mosavi, A. An Integrated Machine Learning, Noise Suppression, and Population-Based Algorithm to Improve Total Dissolved Solids Prediction. Engineering Applications of Computational Fluid Mechanics 2021, 15, 251–271, doi:10.1080/19942060.2020.1861987.
158. Saberi-Movahed, F.; Najafzadeh, M.; Mehrpooya, A. Receiving More Accurate Predictions for Longitudinal Dispersion Coefficients in Water Pipelines: Training Group Method of Data Handling Using Extreme Learning Machine Conceptions. Water Resources Management 2020, 34, 529–561, doi:10.1007/s11269-019-02463-w.
159. Adjovu, G.E.; Gamble, R. Development of HEC-HMS Model for the Cane Creek Watershed. In Proceedings of the 28th Tennessee Water Resources Symposium Tennessee Section of the American Water Resources Association; Tennessee Section of the American Water Resources Association: Montgomery Bell State Park Burns, Tennessee, 2019; pp. 1C-2-1C – 6.
160. Chicco, D.; Warrens, M.J.; Jurman, G. The Coefficient of Determination R-Squared Is More Informative than SMAPE, MAE, MAPE, MSE and RMSE in Regression Analysis Evaluation. PeerJ Comput. Sci. 2021, 7, 1–24, doi:10.7717/PEERJ-CS.623.
161. Goodness-of-Fit Statistics Sum of Squares Due to Error Available online: https://web.maths.unsw.edu.au/~adelle/Garvan/Assays/GoodnessOfFit.html (accessed on 15 June 2023).

 doi:10.20944/preprints202603.0403.v1

162. Uncover Colorado Dolores River – near Dolores Dolores River Route Available online: https://www.uncovercolorado.com/activities/dolores-river/ (accessed on 13 July 2023).
163. Bureau of Reclamation Water Quality Impacts in the Animas and San Juan River Basins Available online: https://www.usbr.gov/research/projects/download_product.cfm?id=2957 (accessed on 13 July 2023).
164. Rhoades, J.D.; Corwin, D.L.; Lesch, S.M. Geospatial Measurements of Soil Electrical Conductivity to Assess Soil Salinity and Diffuse Salt Loading from Irrigation. Geophysical Monograph Series 1998, 108, 197–215, doi:10.1029/GM108p0197.
165. Kurra, S.S.; Naidu, S.G.; Chowdala, S.; Yellanki, S.C.; Sunanda, E. Water Quality Prediction Using Machine Learning. International Research Journal of Modernization in Engineering Technology and Science 2022.
166. Lin, S.; Zheng, H.; Han, B.; Li, Y.; Han, C.; Li, W. Comparative Performance of Eight Ensemble Learning Approaches for the Development of Models of Slope Stability Prediction. Acta Geotech. 2022, 17, 1477–1502, doi:10.1007/s11440-021-01440-1.
167. Ewusi, A.; Ahenkorah, I.; Aikins, D. Modelling of Total Dissolved Solids in Water Supply Systems Using Regression and Supervised Machine Learning Approaches. Appl. Water Sci. 2021, 11, 1–16, doi:10.1007/s13201-020-01352-7.
168. Pizani, F.M.C.; Maillard, P.; Ferreira, A.F.F.; De Amorim, C.C. Estimation of Water Quality in a Reservoir from Sentinel-2 Msi and Landsat-8 Oli Sensors. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences 2020, 5, 401–408, doi:10.5194/isprs-Annals-V-3-2020-401-2020.
169. Abdelal, Q.; Assaf, M.N.; Al-Rawabdeh, A.; Arabasi, S.; Rawashdeh, N.A. Assessment of Sentinel-2 and Landsat-8 OLI for Small-Scale Inland Water Quality Modeling and Monitoring Based on Handheld Hyperspectral Ground Truthing. J. Sens. 2022, 2022, doi:10.1155/2022/4643924.
170. Löwe, R.; Böhm, J.; Jensen, D.G.; Leandro, J.; Rasmussen, S.H. U-FLOOD – Topographic Deep Learning for Predicting Urban Pluvial Flood Water Depth. J. Hydrol. (Amst). 2021, 603, doi:10.1016/j.jhydrol.2021.126898.
171. Chen, P.; Wang, B.; Wu, Y.; Wang, Q.; Huang, Z.; Wang, C. Urban River Water Quality Monitoring Based on Self-Optimizing Machine Learning Method Using Multi-Source Remote Sensing Data. Ecol. Indic. 2023, 146, 109750, doi:10.1016/j.ecolind.2022.109750.
172. Rumora, L.; Miler, M.; Medak, D. Impact of Various Atmospheric Corrections on Sentinel-2 Land Cover Classification Accuracy Using Machine Learning Classifiers. ISPRS Int. J. Geoinf. 2020, 9, doi:10.3390/ijgi9040277.
173. Nguyen, P.T.B.; Koedsin, W.; McNeil, D.; Van, T.P.D. Remote Sensing Techniques to Predict Salinity Intrusion: Application for a Data-Poor Area of the Coastal Mekong Delta, Vietnam. Int. J. Remote Sens. 2018, 39, 6676–6691, doi:10.1080/01431161.2018.1466071.
174. Nanglia, S.; Ahmad, M.; Ali Khan, F.; Jhanjhi, N.Z. An Enhanced Predictive Heterogeneous Ensemble Model for Breast Cancer Prediction. Biomed. Signal Process. Control 2022, 72, 103279, doi:10.1016/j.bspc.2021.103279.
175. Sharaf El Din, E.; Zhang, Y.; Suliman, A. Mapping Concentrations of Surface Water Quality Parameters Using a Novel Remote Sensing and Artificial Intelligence Framework. Int. J. Remote Sens. 2017, 38, 1023–1042, doi:10.1080/01431161.2016.1275056.
176. Slonecker, E.T.; Jennings, D.B.; Garofalo, D. Remote Sensing of Impervious Surfaces: A Review. Remote Sensing Reviews 2001, 20, 227–255, doi:10.1080/02757250109532436.