

**Testing Water Quality from Space:
A Comparison between Remote Sensing and *in situ* Measurements**

A directed research thesis

Presented in partial fulfillment of the requirements for the

Degree of Master of Applied Geography

with a

Major in GIScience

in the

Graduate College

Texas State University

By:

Hannah Sturm

May 2019

Advising Professor: Jennifer L.R. Jensen, Ph.D.

Table of Contents

List of Tables.....	iii
List of Figures	iv
List of Equations.....	v
Abbreviations.....	vi
1.0 Problem Statement	1
2.0 Background.....	3
2.1 Study Region: Guadalupe River Basin	3
2.2 Population Pressure Increase on Water Quality	5
2.3 Water quality policies	6
3.0 Literature Review	9
3.1 Remote Sensing Compared to <i>In Situ</i> Measurements	9
3.2 Remote Sensing to Assess Water Quality	10
3.3 Measuring Turbidity and Total Suspended Solids of Water through Remote Sensing	11
4.0 Research Methods	13
4.1 Data Collection and Preparation	13
4.2 Reflectance Value Measurements	14
4.3 Comparison in <i>In Situ</i> Data vs Reflectance Data	15
4.4 Current Extent of Turbidity and Total Suspended Solids.....	16
5.0 Results.....	17
5.1 Regression modeling of turbidity and TSS.....	17
6.0 Discussion.....	27
7.0 Conclusion	36
8.0 Bibliography	38
9.0 Definitions.....	41

List of Tables

1. Data sources for satellite imagery and *in situ* turbidity and TSS measurements 14
2. Linear regression equations for predicting turbidity and TSS from band reflectance 25

List of Figures

1. Map of the Guadalupe River basin.....	3
2. Actual vs predicted plot for sqrt(turbidity) using B8 and B3/B8.....	17
3. Actual vs predicted plot for sqrt(turbidity) for Dec. 21, 2016 using B3/B8.....	18
4. Actual vs predicted plot for sqrt(turbidity) for Jan. 30, 2017 using B4/B8.....	19
5. Actual vs predicted plot for sqrt(turbidity) for Apr. 30, 2017 using B3.....	20
6. Actual vs predicted plot for log(TSS) using B8 and B8/B2.....	21
7. Actual vs predicted plot for log(TSS) for Dec. 21, 2016 using B3/B8.....	22
8. Actual vs predicted plot for log(TSS) for Jan. 30, 2017 using B3.....	23
9. Actual vs predicted plot for log(TSS) for Apr. 30, 2017 using B2.....	24
10. Map of Guadalupe and Comal River showing estimated turbidity using the all date's model.....	28
11. Map of Guadalupe and Comal River showing estimated TSS using the all date's model.....	29
12. Map of Guadalupe and Comal River showing estimated TSS using the Dec 2016 model.....	31
13. Map of Guadalupe and Comal River showing estimated turbidity using the Jan 2017 model.....	32
14. Map of Guadalupe and Comal River showing estimated TSS using the Apr 2017 model.....	33

List of Equations

1. Conversion to Top-of-Atmosphere reflectance..... 14

Abbreviations

ANOVA – Analysis of Variance

B2 – Sentinel-2A band 2, blue

B3 – Sentinel-2A band 3, green

B4 – Sentinel-2A band 4, red

B8 – Sentinel-2A band 8, near infrared

FWPC – Federal Water Pollution Control

GBRA – Guadalupe Blanco River Authority

Landsat ETM+ - Landsat Enhanced Thematic Mapper Plus

Landsat MSS – Landsat Multispectral Sensor

Landsat TM – Landsat Thematic Mapper

MODIS – Moderate Resolution Imaging Spectroradiometer

NIR – near infrared

R² – correlation of determination

TCEQ – Texas Commission of Environmental Quality

TOA – Top of Atmosphere

TSS – Total Suspended Solids

VNIR – visible and near infrared

1.0 Problem Statement

On any given weekend during the summer, over 2,000 people gather in New Braunfels, Texas in preparation to float either the Guadalupe or Comal Rivers. With so many people in the river, water quality can decrease quickly because of debris thrown in the river, human urination, and sediment disturbance along the river bottom. Decreases in water quality can disrupt species that use the river as a source of water, habitat, and/or their food supply. Water quality can also affect the health of the people that use the water for recreation, resulting in increases in water treatment plant prices because it is more difficult to revert the water back to the quality where it can be used for human consumption. Ensuring water quality requires water testing. There is great potential to monitor water quality by using remote sensing data, but there are few studies that report how image spectral information compares to *in situ* testing despite remote sensing data being more cost efficient (Ritchie et al 2003). This study provides one of the few comparisons how remote sensing can test the utility of remote sensing data to estimate water quality.

In situ water quality testing currently takes place both in the field as well as in a laboratory. Tests conducted in the field include pH, conductivity, chlorine concentration, water temperature, and turbidity. Laboratory tests include dissolved oxygen, total phosphorous, nitrate, sulfate, ammonia, and chlorophyll – a. A challenge to measuring and monitoring water quality *in situ* is that it can be prohibitively expensive and time consuming. A person has to go to the collection site, collect the water, do on-site testing, and then bring the water to the laboratory so a lab technician can test the water and validate the results that were collected in the field. An alternative method to monitor water quality is by using satellite imagery. With the development and launching of satellites such as Landsat and Sentinel-2, people in areas where water quality data are limited or missing can use remote sensing techniques to estimate water quality.

By using satellite imagery available from the European Space Agency and water quality data from the Guadalupe Blanco River Authority (GBRA), this research will analyze images at different testing sites along the Guadalupe and Comal Rivers to determine if spectral responses within the water column have a statistical relationship with turbidity and total suspended solids (TSS) measurements collected by GBRA on site. Specifically, my research question is “How effective are Sentinel-2 satellite data for estimating turbidity and total suspended solids in the Guadalupe and Comal Rivers, near New Braunfels, TX, compared to *in situ* measurements collected by the Guadalupe Blanco River Authority?” I hypothesize that reflectance values from the satellite imagery can be used to estimate quantities that GBRA collected at their testing sites.

To answer the proposed question, this study will involve the use of quantitative data that were acquired from the Sentinel-2 satellite sensor and measurements that GBRA collected and posted on their website. The GBRA website has water quality data and associated GPS testing locations, dates, and the specific data acquired. Five testing sites were chosen along the Guadalupe and Comal Rivers based on a date buffer of five days before or after Sentinel-2 satellite overpass. This analysis will be useful for local water authorities because it can be a way to mitigate costs associated with field data collection and monitoring costs.

2.0 Background

2.1 Study Region: Guadalupe River Basin

The Guadalupe River headwaters are located in Kerr County, Texas, approximately 145 km northwest of San Antonio. The river flows out to the San Antonio Bay roughly 56 km southeast of Victoria and empties into the Gulf of Mexico. Total river length is 370 km (Smyrl 2010). Major cities along the Guadalupe River include Kerrville, New Braunfels, Seguin, Cuero and Victoria. The Guadalupe River is divided into three parts, the Upper Guadalupe River, the Middle Guadalupe River, and the Lower Guadalupe River. The upper and middle sections of the river bring in tourists for activities including fly fishing, rafting, canoeing, hiking and tubing, while the lower part of the river is more suitable for camping (Smyrl 2010).

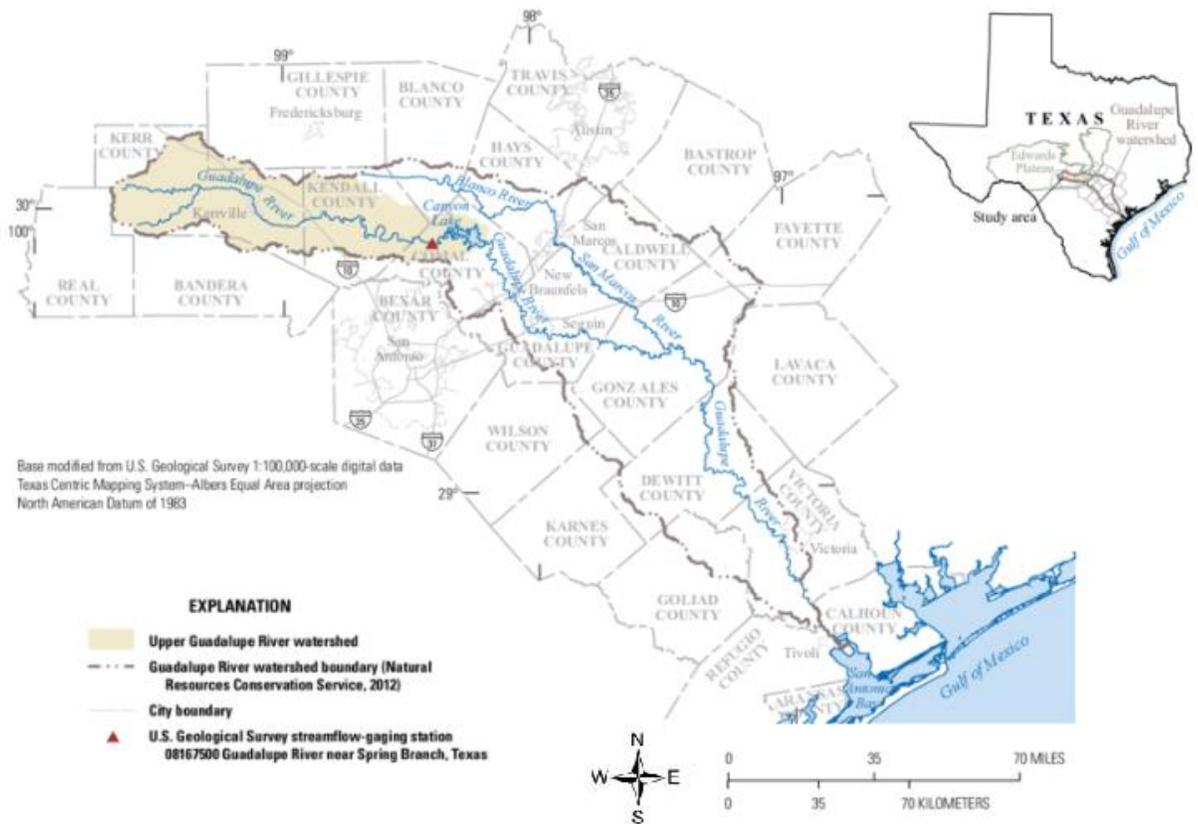


Figure 1: Map of the Guadalupe River basin (Asquith and Bumgarner, 2014).

The Middle Guadalupe River drains from Canyon Lake Dam in Comal County, Texas to Cuero, Texas in DeWitt County, Texas. Water quality testing for this section is important due to the large number of people using the river for recreation as well as species' habitats. This section of the river has "the finest white-water" stretch in the state because it cuts through the Balcones Fault Zone (GBRA 2013). It also contains seven water power plants used by the Guadalupe Blanco River Authority (GBRA) whose dam structures have created lakes used for recreational activities. Up until the confluence with the San Marcos River in Gonzales, Texas, the Guadalupe River is a narrow fast moving river. However, when it combines with the San Marcos River, it becomes large and slow moving. This middle section of the river is very scenic and has limestone bluffs, bald cypress trees (*Taxodium distichum*), pecan trees (*Carya illinoensis*), and elm trees (*Ulmus americana*) (GBRA 2013). The Middle Guadalupe River is also stocked with rainbow trout (*Oncorhynchus mykiss*) between December and January by the Trout Unlimited organization and the Texas Parks and Wildlife Department.

The Upper Guadalupe River runs from the headwaters in Kerr County, Texas to Canyon Lake in Comal County, Texas. Water quality is important in the Upper Guadalupe River because of the four major aquifers and plant and animal biodiversity. This portion of the river flows over four major aquifers which are defined by the Texas Water Development Board as areas "that supply large quantity of water in large areas of the state" (Ashworth and Hopkins 1995). These aquifers are layers of limestone which were formed during the Lower Cretaceous period (65 – 145 million years ago) with the beds of the canyons that were formed from the erosion of the water flowing over the rocks (Stricklin et al. 1971). The landscape around the Upper Guadalupe River includes riparian forests that contain species such as sycamore (*Platanus occidentalis*), cedar elm (*Ulmus crassifolia*), bald cypress (*Taxodium distichum*), mountain laurel (*Kalmia*

latifolia) and Ashe juniper (*Juniperus ashei*). Additionally, this section contains grasslands that include native grasses such as sideoats grama mix (*Bouteloua curtipendula*), indian grass (*Sorghastrum nutans*) and little bluestem (*Schizachyrium scoparium*) (TNC 2007). The Golden Cheek Warbler (*Setophaga chrysoparia*) and Black-capped vireo (*Vireo atricapilla*) are two federally-listed endangered bird species that have critical habitats along the Upper Guadalupe River (Flatten 2017) as well as other endangered species including Guadalupe bass (*Micropterus treculii*), Honey Creek Cave salamander (*Eurycea tridentifera*), and Cagle's map turtle (*Graptemys caglei*) (TPWD 2000).

The Lower Guadalupe River is classified as the section that runs through Cuero, Texas in DeWitt County, Texas to the Gulf of Mexico through the San Antonio Bay in Calhoun County, Texas. Water quality is pertinent to this section because this is where the fresh river water meets the salty ocean water. This section of the river is surrounded by grasslands and contains marshes which are created by the tide of the river mixing with the Gulf of Mexico (GBRA 2013). This section includes habitat used by the endangered whooping crane (*Grus americana*), whose diet consist of blue crabs; a food source dependent on brackish water (GBRA 2013).

2.2 Population Pressure Increase on Water Quality

The Middle Guadalupe River flows through the City of New Braunfels, Texas which is located in Comal County between San Antonio, Texas to the south and San Marcos, Texas to the north. It is situated along Interstate 35 and listed as one of the fastest growing cities in the United States with a total population of 73, 959 in 2016 (US Census Bureau 2017). With urban development, there are many water quality issues that can occur such as the redistribution of water resources to new developments, the movement and settlement of suspended sediments due to development taking place upstream, an increase in nitrogen and phosphorus runoff from crop

fertilization and amendments, and sewage overflows due to clogged pipes or the breakdown of pumping stations (USGS 2016).

The redistribution of water causes more pipes to be placed for the pumping of water and also leads to areas that will no longer get adequate amounts of water necessary for their population. Increased sediment loads can cause the lifetime of reservoir to decrease because as the sediment flows in at a quick rate, the water flow within the reservoir slows and keeps the sediment suspended causing the sediment to settle to the bottom and build up over time. This build up then causes the reservoir to become useable (USGS 2016). Excess nitrogen and phosphorus results in eutrophication, the process of rapid plant growth and oxygen depletion, with detrimental effects on a water body. In lakes and reservoirs, eutrophication can occur which produces algae on the water's surface and can kill fish due to the lack of sunlight being able to penetrate through the water (USGS 2016). Excess nitrogen and phosphorus, in drinking water, can be harmful to young infants and livestock, and it can also lead to the restriction of oxygen in the bloodstream (USGS 2016).

2.3 Water quality policies

Water quality testing standards have not always been in place and to the degree with which they are enforced today. To control the pollution that was occurring in United States waterways, the United States Legislature enacted the Federal Water Pollution Control Act (FWPCA) of 1948. This act initiated programs that would eliminate and reduce pollution occurring in waterways and tributaries, along with improving the sanitary conditions of surface and underground waters (US Fish and Wildlife Service 2013). This statute also allowed for establishment of water treatment plants to prevent the discharge of untreated sewage and other wastes into the water. Since its implementation, this act has been amended multiple times to

allow for additional water quality programs, updates to procedures and standards that are allowed for discharge, and funding for grants and programs (US Fish and Wildlife Service 2013). Before the implementation of the FWPC Act, the Federal government did not address the pollution issues that occurred in the waterways that stayed in one state; they were only interested in the waterways that were crossing multiple states (Copeland 2016). The 1965 amendment was the first reference to water quality standards that required the states to set standards to determine the amount of pollution in the water that crossed multiple state boundaries and also requirements to control pollution (Copeland 2016). In 1972, the Federal Water Pollution Control Act of 1948 was revised and is now known as the Clean Water Act. This revamp led to goals that all wastewater had to be treated, increased federal government funding for construction of water treatment plants, and the increased role the federal government in implementing the law (Copeland 2016).

In 1991, the Texas Clean River Program was formed through a partnership with the Texas Commission of Environmental Quality (TCEQ) and regional water authorities, such as GBRA and Canyon Regional Water Authority, etc., to coordinate water quality monitoring and conduct assessments to improve the surface water qualities in Texas river basins (TCEQ 2018). There are six main objectives of the Clean River Program including: 1) providing the TCEQ with quality assured data that can be used in water quality decision making, 2) identifying and evaluating issues related to the water quality, 3) involve federal, state and local agencies on the issues that are related to planning and cleaning up the water, 4) inform and involve those that have an interest in the program so their ideas can be heard, 5) use the funds that the TCEQ provides without going over, and 6) come up with a plan and committee to adapt to the changes that occur in the program (TCEQ 2018). These objectives have led to increased sampling of the rivers around the state to get a better understanding of the river water quality, built relationships

among regional water authorities and TCEQ, increased communication between TCEQ and regional water authorities on issues related to poor water quality, and increased the public participation for issues that are plaguing their local river basin and coming up with ideas on how to make their area cleaner and healthier (TCEQ 2018).

3.0 Literature Review

3.1 Remote Sensing Compared to In Situ Measurements

According to House Document 110-91 (2008), the advantages of remote sensing include the ability to incorporate data with other information that can aid in decision making, characterize physical objects and natural features on the ground, collect data over large spatial areas, and to observe areas and monitor their changes over time. The location of this committee hearing took place in Colorado and the State of Colorado is currently using infrared remote sensing to estimate aquifer depletion, water rights and climate change (US H.Doc. 110-91 2008). The use of remote sensing techniques can replace on the ground investigation because satellite based information is beneficial for proficient water management, proficient use of limited resources and improve the decision making processes (US H.Doc. 110-91 2008). The use of satellite imagery is a way to measure large areas when rivers and crops lands can vary in size (US H.Doc. 110-91 2008).

Baumgartner and Apfl (1996) and Santamaria-del-Angel et al. (2011) state that remote sensing measurements are taken at a much larger scale, depending on the satellite that is chosen, ranging from 1 m² to 1 km², while *in situ* measurements are based on one liter of water. Remote sensing measurements integrate data through optic approximations from greater volume which yield averages in the water quality values (Santamaria-del-Angel et al 2011). Remote sensing offers a cost-effective way for scientists to measure water quality by collecting measurements over large areas and provides an overall summary of the surface (Spath et al 2008). Additionally, remote sensing offers the advantage of having datasets that are large in both temporal and spectral areas (Baumgartner & Apfl 1996). This benefit gives scientists the added capability to re-measure a location due to lost data or an initial bad reading. Spath et al (2008) also describes

that a limitation of remote sensing can be the spatial and temporal resolution of the data as well as the environmental parameters, such as current and sediment.

In situ measurements have the ability to monitor water quality with a high resolution but are limited to the area in which they are fixed (Spath et al 2008). When measuring a large area, this requires a lot of testing stations and with the increased number of testing stations come the increase operational maintenance cost and the initial deployment which also costs a lot of money (Spath et al 2008). When the results are taken back to the lab, they are based off of information from a spectrophotometer, a fluorometer, or a high performance liquid chromatograph which uses the one liter of water (Santamaria-del-Angel et al 2011).

3.2 Remote Sensing to Assess Water Quality

Approaches to estimate or monitor water quality with remotely sensed data have taken multiple approaches. Ritchie et al. (2003) discuss how remote sensing can be used to calculate reflectance values for chlorophyll, total suspended solids, and temperature. Curran and Novo (1988) describe that the best wavelengths to use to measure total suspended solids are the blue/green wavelengths but these wavelengths are most affected by atmospheric scattering making it difficult for the values to be calculated. For chlorophyll, the Landsat and SPOT sensors have bandwidths that are too large to be able to differentiate the difference between chlorophyll and total suspended solids (Ritchie et al., 2003). It was determined that the narrow wavelength difference between the red band and NIR band, “red edge” is the best area to depict chlorophyll because it shows the relationship between chlorophyll and total suspended solids. The larger the “red edge” the more photosynthesis is occurring. To measure the temperature of water, the thermal bands are the most effective. Monitoring the thermal changes of water is useful because biological productivity can be detected (Ritchie et al 2003).

Kabbara et al. (2008) conducted a study of the coastal area of Tripoli (Lebanon) to map the distribution of water quality parameters that included turbidity and chlorophyll-a concentration. They used two sensors, Landsat 7 Enhanced Thematic Mapper Plus (ETM+) as their high resolution sensor and the SeaWiFS as their low resolution sensor to compare to the data that they gather via *in situ* using the Secchi Disk depth method for turbidity. Kabbara et al. (2008) determined that the use of the Landsat 7 ETM+ was not a viable option to measure water quality because high concentrations of chlorophyll-a results in high reflectance values making it difficult to determine whether the energy is being reflected from the water or chlorophyll-a. Ogashawara and Moreno-Madrinan (2014) also studied the reflectance of chlorophyll-a. They used the Moderate Resolution Imaging Spectroradiometer (MODIS) which has a pixel size the same as SeaWiFS (1 km), much larger than the Landsat 7 ETM+ at 30m spatial resolution. They found that the NIR band was not sufficient because the band width was too large to detect the small spectral variations that were required to estimate chlorophyll-a concentrations.

3.3 Measuring Turbidity and Total Suspended Solids of Water through Remote Sensing

Kabbara et al. (2008) states that measuring turbidity is difficult because it can be affected by many factors such as the presence of zooplankton and phytoplankton. Chen et al. (2007), Petus et al. (2010) and Robert et al. (2016) used MODIS to measure turbidity. The areas that were studied were highly turbid, but the conclusions were the same, that there is a positive correlation between reflectance values of the red/NIR bands and turbidity. Petus et al (2010) also concluded that MODIS band 2 (NIR), which detects land/clouds/aerosols/boundaries, was not sensitive enough to detect turbidity due to the location of the bands wavelengths (841 nm – 876 nm).

Hellweger et al. (2004) chose to study the New York Harbor with two different sensors, Landsat Thematic Mapper (TM) and MODIS. To study turbidity, the red band of the TM was used. They noted that as the Secchi Disk got deeper in the harbor, the reflectance values also decreased, suggesting there is a strong negative correlation between red reflectance values and turbidity. Harrington et al. (1992) studied Lake Chicot, Arkansas with the Landsat Multispectral Sensor (MSS) and determined that the best MSS bands were the visible red and near infrared bands due to the ability to eliminate the detection of photosynthesis. MODIS can also be used to measure totals suspended solids (TSS) concentration and the 250 meter bands are the optimal wavelengths (Chen et al, 2015). Liu et al. (2017) used Sentinel 2 data and compared those results to those from MODIS to estimate the TSS concentration of Poyang Lake, China. They determined that Sentinel-2 bands 4 (red) and 7 (vegetation red edge) were the best at estimating TSS concentration.

Lim et al. (2010) studies TSS concentration in Malaysia using THEOS satellite imagery. They found that the visible bands had the strongest correlation with TSS concentration. Mabwoga et al. (2010) studied the concentration TSS in a wetland in India using imagery from the Indian Remote Sensing satellite, LISS IV. The band that showed the strongest correlation with TSS was the NIR band. Wang et al (2007) used the Hyperion sensor to measure the TSS concentration of a bay in China. After testing, they determined that the VNIR wavelength would be suited to continue with future work but of the visible bands that were tested, the red wavelength showed to have the highest correlation.

4.0 Research Methods

As defined above, this study seeks to determine how effective a satellite sensor data, particularly Sentinel-2, is at estimating *in situ* turbidity and total suspended solids (TSS) measurements of the Guadalupe and Comal Rivers, near New Braunfels, Texas. I plan to use a positivist epistemology to measure the pixel values and show statistically by date areas that are turbid and have a high TSS value.

To conduct a thorough understanding of the water quality of the Guadalupe and Comal Rivers study area, I planned my analysis in four steps: data collection and preparation, calculating reflectance values of pixels, comparison of reflectance values to *in situ* measurements, and mapping what each testing location looks like during the day the image was acquired. The study's methodology includes a linear regression model and visualizations, which include GIS-generated maps and remote sensing data analysis.

4.1 Data Collection and Preparation

Five; Texas Commission on Environmental Quality (TCEQ) testing locations were identified along the Guadalupe and Comal Rivers that have dates after the launching of the Sentinel-2 satellites, June 23, 2015 for Sentinel-2A and March 2, 2017 for Sentinel-2B. These testing location are: TCEQ station 12598, Canyon Reservoir (Canyon Park Marina), station 12653, Comal River at Hinman Island, station 12658, Guadalupe River at River Road (Second Crossing), station 13700, Guadalupe River near Spring Branch, and station 15082, Comal River at Landa Park Area 16. The GBRA collection dates that were chosen for the TCEQ testing locations were December 19, 2016, January 29, 2017 and April 29, 2017.

Sentinel-2A Level-1C imagery were downloaded from the European Space Agency (Copernicus) website for the following dates: December 21, 2016, January 30, 2017 and April

30, 2017. Level-1C processing includes radiometric and geometric corrections and top of atmosphere reflectance (ESA 2018). These images were chosen because they have minimal cloud coverage and are based on a temporal buffer of 5 days before or 5 days after the collection of the data GBRA posted on their website.

Agency	Data
European Space Agency	Sentinel-2 Imagery
Guadalupe Blanco River Authority	Water quality measurements and testing site locations

Table 1: Data sources for satellite imagery and *in situ* turbidity and TSS measurements.

This temporal buffer ensured that the water quality is similar to when the collection occurred and also facilitated for the possibility of using alternative additional images if there was cloud cover over the location. After the images were acquired, the testing locations were geolocated and the turbidity and total suspended solids values were added into a GIS. These locations used the NAD_1983_2011_StatePlane_Texas_South_Central_FIPS_4204 projection.

4.2 Reflectance Value Measurements

Second, measuring reflectance values of pixels occurred by measuring the spectral response of the pixel for each band. The bands chosen included: visible blue (band 2), visible green (band 3), visible red (band 4), and the near infrared (band 8). Equation 1 takes the digital number, which is the pixel value the sensor stores, and converts it to top-of-atmosphere (TOA) reflectance (ESA 2018).

$$\rho_{\kappa}(i, j) = \frac{\pi \times CN_{\kappa,NTDI}(i, j)}{A_{\kappa,NTDI} \times E_s \times d(t) \times \cos(\theta_s(i, j))}$$

Equation 1: Conversion to Top-of-Atmosphere reflectance

where:

$$\rho_{\kappa} = \text{TOA reflectance}$$

CN = Digital Number of each pixel in the image

i, j = Pixel image value

k = Spectral band

E_s = Extra-terrestrial solar spectrum

θ_s = Incoming solar direction, zenith angle

A_k = Absolute calibration of the instrument

$d(t)$ = Correction for the sun-Earth distance variation

The TOA reflectance was used in the Sen2Cor program which converts TOA reflectance to scaled surface reflectance values. Surface reflectance is different from TOA reflectance because it accounts for scattering and absorption in the atmosphere and provides file pixel values that are spectrally similar to what an observer would have measured with a spectroradiometer *in situ* during satellite overpass. The Sen2Cor program converted the Level-1C imagery to Level-2A imagery, which resampled the bands to their respective 10m spectral resolutions.

To measure the reflectance values of each band at the testing locations, ESRI's ArcMap tool, Extract Multi Point Values to Points, was used. This tool extracts cell values that coincide spatially to a specified point feature class from one or more rasters and records that value as an attribute in the point feature class. These values were then entered in an Excel spreadsheet along with observation data, TCEQ station number, and the *in situ* turbidity and TSS data. Additional band ratios were included as well, for example B2/B3, B2/B4, B2/B8, etc.

4.3 Comparison of In Situ Data vs Reflectance Data

To meet the assumptions required for linear regression modeling, transformation of the *in situ* data collected by taking the square root of the *in situ* turbidity samples, and the logarithm of the *in situ* TSS samples so that the dependent variables were normally distributed. To determine which band or band ratios explained the most variability in field-measured values, the highest R^2

(correlation of determination) value was selected based on an all possible model approach executed in JMP Pro.

The overall models for turbidity and TSS implemented those variables as dependent variables (Y value) and Sentinel-2 reflectance values or band ratios as independent variables (X values). From the output, all possible models were created using all possible combinations of the independent variables. The top four models with up to two terms per model were reported. From those results the model that produced the highest R^2 values and lowest Root Mean Square Error (RMSE) values were chosen. Taking the band and band combinations from the all possible models, the fit model tool was ran again using the standard least squares personality, producing an actual by predicted plot. The predicted values were then saved and the fit Y by X tool was ran creating a predicted by actual values plot and regression outputs. This process was carried out for all dates and for individual dates.

4.4 Current Extent of Turbidity and Total Suspended Solids

Five maps were generated using the regression equations and the raster calculator function in ArcMap to show the estimated turbidity and TSS throughout the study area.

5.0 Results

5.1 Regression modeling of turbidity and TSS

The all dates model for turbidity (n=14) produced an R^2 value of 0.567 and a root mean square error (RMSE) of 0.405 and included two terms; band 8 (near infrared) and ratio band 3/band 8 (green/near infrared). Figure 2 provides an actual vs. predicted plot for this model output.

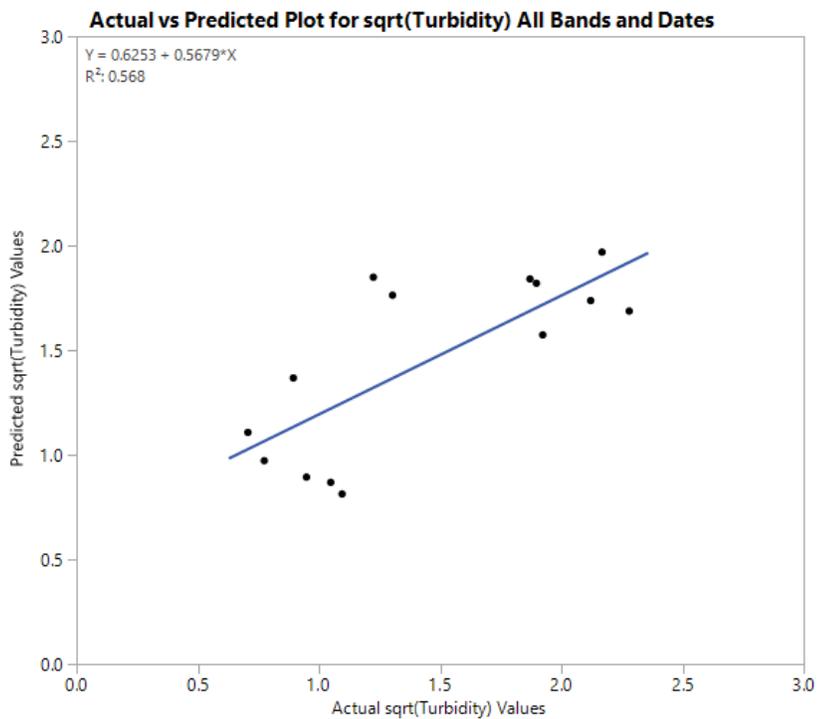


Figure 2: Actual vs predicted plot for sqrt(turbidity) using B8 and B3/B8.

Figure 3 shows the actual versus predicted values for December 21, 2016 based on a one term model that included ratio band 3/band 8 (green/near infrared) as predictor variables and produced an $R^2 = 0.758$ and $RMSE = 0.30$. This model used five observation and only included the *in situ* measurements that were collected on December 21, 2016.

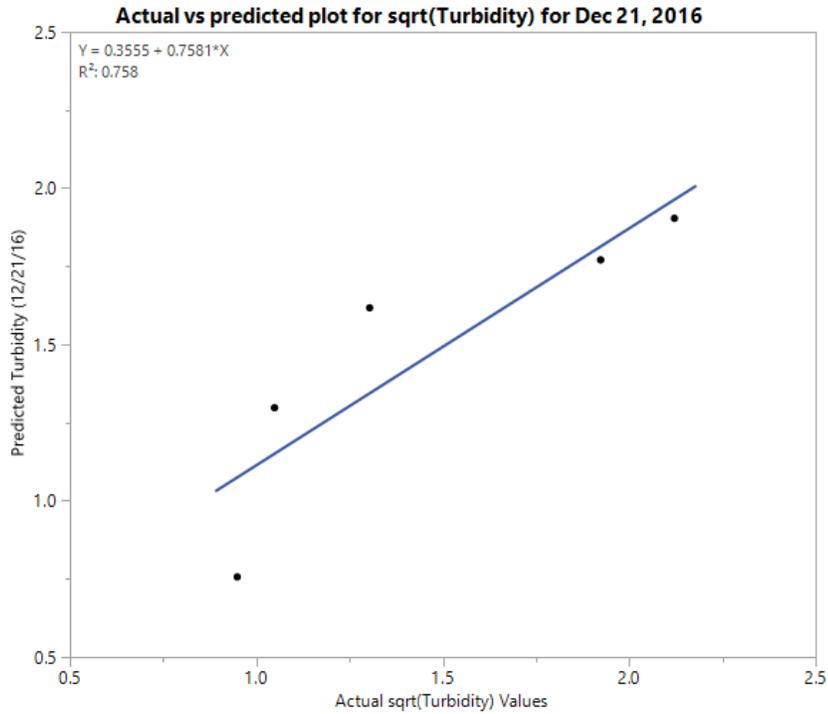


Figure 3: Actual vs predicted plot for sqrt(turbidity) for Dec. 21, 2016 using B3/B8.

A one term model produce an R^2 value of 0.86 with a RMSE of 0.26 ($n = 5$) for the reflectance values and *in situ* data collected on January 30, 2017, as shown in Figure 4. The term used in this model was the ratio of band 4/band 8 (red/near infrared).

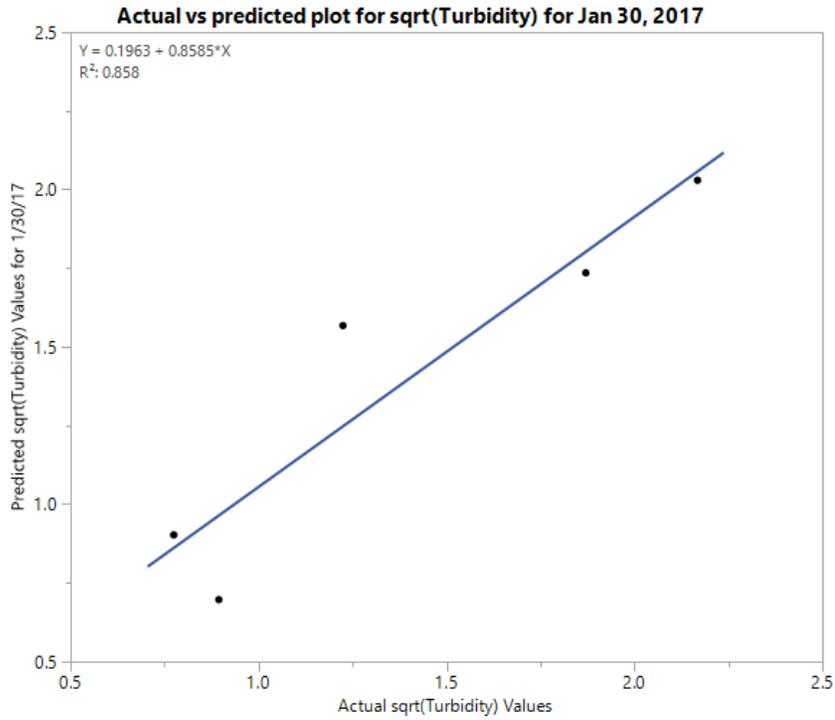


Figure 4: Actual vs predicted plot for sqrt(turbidity) for Jan. 30, 2017 using B4/B8.

To best predict the turbidity for April 30, 2017, a one term model included band 3 (green) was used. To generate the $R^2 = 0.73$, $RMSE = 0.46$, with $n = 4$. Refer to Figure 5 for the actual vs. predicted plot.

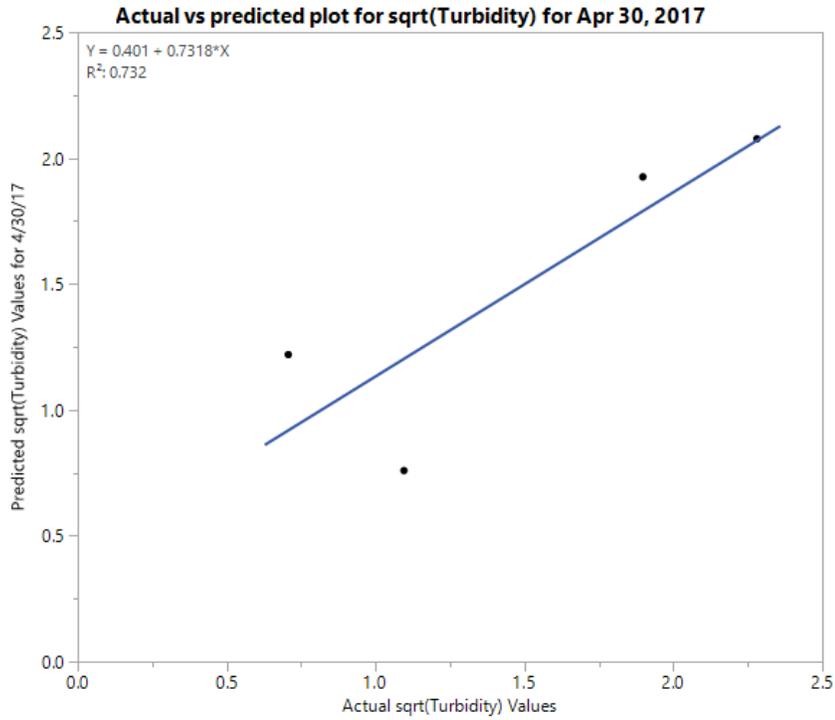


Figure 5: Actual vs predicted plot for sqrt(turbidity) for Apr. 30, 2017 using B3.

The linear regression for TSS produced for all observed dates, also included two model terms: band 8 (near infrared) and band ratio band 8/band 2 (near infrared/blue). This model generated an $R^2 = 0.6396$ (RMSE = 0.23; n = 14). Refer to Figure 6 for the actual vs predicted TSS plot.

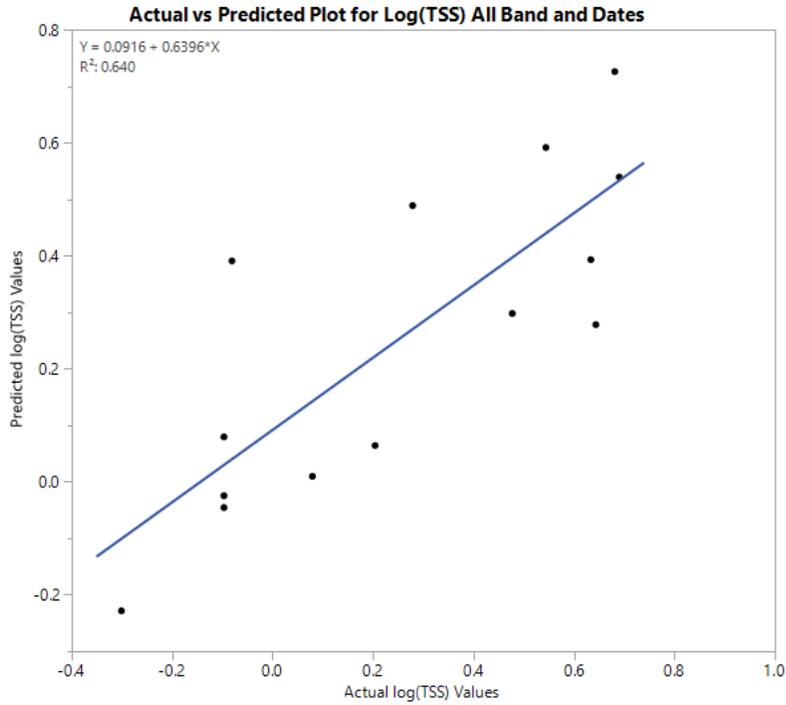


Figure 6: Actual vs predicted plot for log(TSS) using B8 and B8/B2.

Figure 7 shows the actual TSS values versus the predicted reflectance values for December 21, 2016 based on a one term model for ratio band 3/band 8 (green/near infrared). The model (n=5) produced an $R^2 = 0.79$ and an RMSE of 0.14.

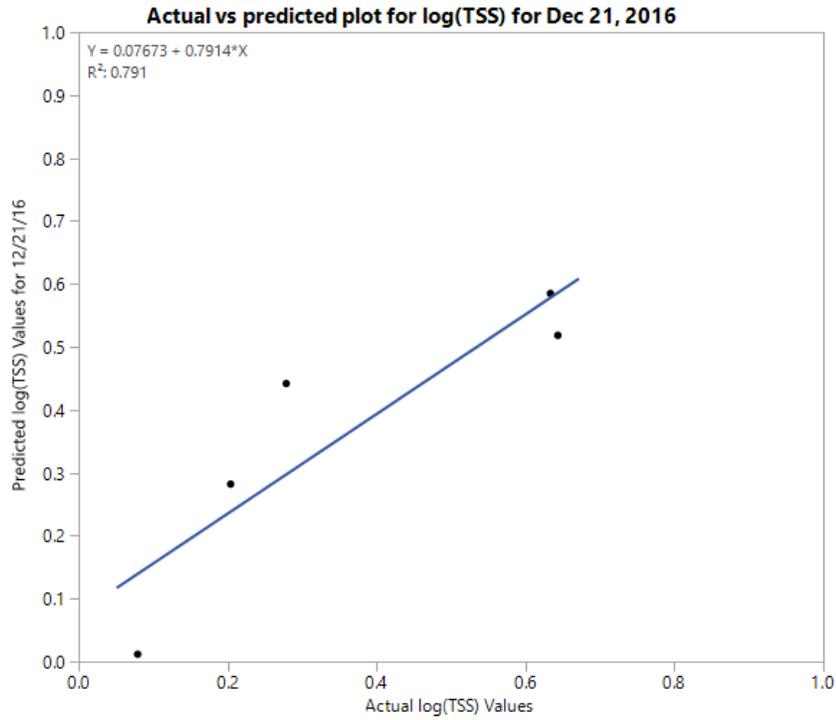


Figure 7: Actual vs predicted plot for log(TSS) for Dec. 21, 2016 using B3/B8.

A one term model resulted in the highest R^2 value ($n=5$), $R^2 = 0.51$ (RMSE = 0.27), for the reflectance values and data collected via *in situ* on January 30, 2017, as shown in Figure 8. The term used in this model was band 3 (green).

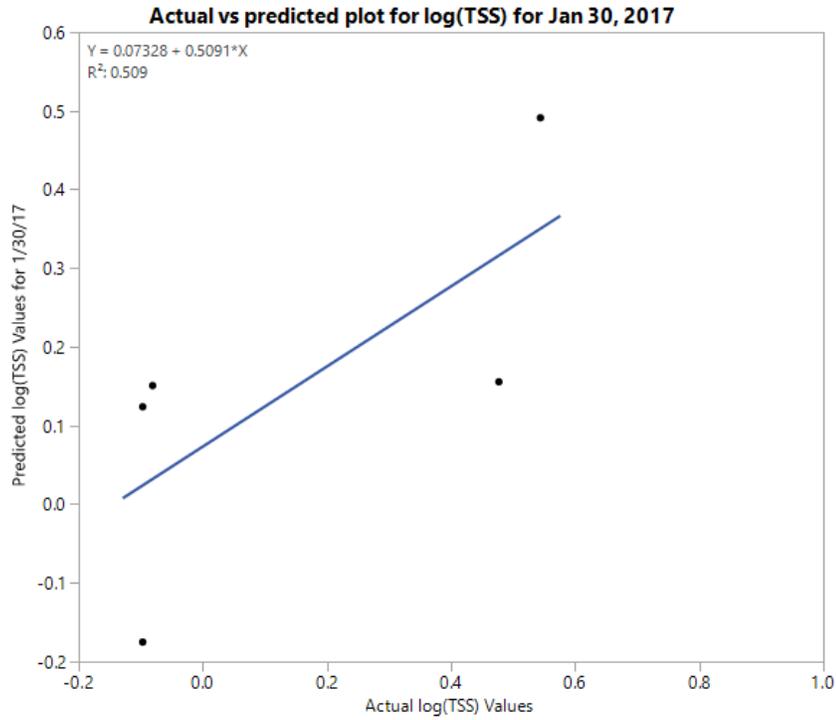


Figure 8: Actual vs predicted plot for log(TSS) for Jan. 30, 2017 using B3.

A one term TSS model (n=4) was selected for April 30, 2017 and included, band 2 (blue) with an R² of 0.97 and RMSE of 0.11. Refer to Figure 9 for the actual vs. predicted plot.

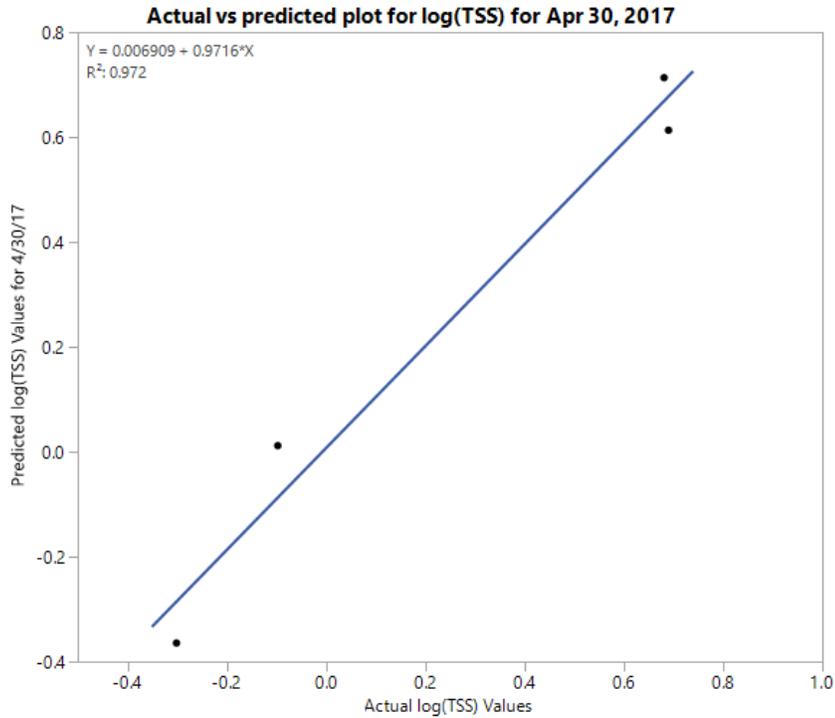


Figure 9: Actual vs predicted plot for log(TSS) for Apr. 30, 2017 using B2.

In describing each model by the bands that produced the highest R^2 values, linear regression equations were produced based on each model's parameter estimates. These parameters include the intercept and the coefficients for each model term. These equations are used to communicate the values that will be used when determining the actual turbidity and TSS values from the predicted reflectance values. Table 2 summarizes the model equations.

Dataset	n	R ²	RMSE	Model
All Data	14	0.57	0.41	Turbidity = -1.408 + 0.00058(B8) + 2.7568(B3/B8)
Dec. 21, 2016	5	0.76	0.30	No significant model
Jan. 30, 2017	5	0.86	0.26	Turbidity = -1.745 + 3.72886(B4/B8)
Apr. 30, 2017	4	0.73	0.46	No significant model

Dataset	n	R ²	RMSE	Model
All Data	14	0.64	0.23	TSS = 1.01141 + 0.00045(B8) – 0.71419(B8/B2)
Dec. 21, 2016	5	0.79	0.14	TSS = -1.103 + 1.79261(B3/B8)
Jan. 30, 2017	5	0.51	0.27	No significant model
Apr. 30, 2017	4	0.97	0.11	TSS = -1.321 + 0.00176(B2)

Table 2: Linear regression equations for predicting turbidity and TSS from band reflectance.

When deciding if the overall model best represents the ability to calculate either turbidity or TSS, the value of the *F*-test in the Analysis of Variance (ANOVA) was verified in determining that both models as a whole are statistically significant. The *F*-test values for turbidity and TSS were 0.0099 and 0.0036 respectively, with a significance level of < 0.01.

In breaking down the equations by observation date, the regression equations for turbidity and TSS on December 21, 2016, showed an ANOVA *F*-test significance value of 0.0547 with no significance and 0.0433 at significance level of 0.05, respectively. Even though turbidity cannot be predicted in December, having a 99% confidence level for TSS, the hypothesis can be accepted for TSS with the reflectance values for the model to best explain the variability from *in situ* measurements. Since the parameter estimates for turbidity were not statistically significant, the model produced was not used. The regressions equations produced for turbidity and TSS on January 1, 2017, the ANOVA *F*-test value showed 0.0236 at significance level of 0.05 for turbidity and 0.1759 with no significance for TSS. For TSS, the parameter estimates were not statistically significant, so the model produced from the regression equation was not used. Lastly, the regression equations that were produced for April 30, 2017, resulted in an ANOVA *F*-test significance of 0.1446 with no significance for turbidity and an ANOVA *F*-test significance of

0.0143 with a 0.05 significance level. With a confidence level of 95% confidence for TSS, the hypothesis of reflectance values may be used to estimate TSS for this study area. For turbidity, the model produced from the regression equation was not used, due to the statistically insignificant parameter estimates.

6.0 Discussion

The hypothesis for the ability to use reflectance values from satellite imagery to estimate water quantities that GBRA collected at their testing sites, can be accepted as true.

Figure 10 shows what the turbidity looks like for all observations (n=14). The darker the brown the higher the turbidity and the darker the blue the lower the turbidity. From Station 13700, Guadalupe River near Spring Branch, it can be noticed that there is a mixture of turbidity, some areas are higher than other. When the value of the pixel is measured and compared to the normalized turbidity measurement, the values are similar resulting in the all date's model prediction as a valid option for predicting turbidity. Station 12598, Canyon Reservoir (Canyon Park Marina), has a mixture of turbidities, ranging from high turbidity farther in the lake and lower turbidity the closer to land. As for where the station is located, the turbidity is on the higher side which relates to what the normalized *in situ* turbidity measurement collected read. Moving down the river to Station 12658, Guadalupe River at River Road (Second Crossing), from the map, it is noticeable that the turbidity is low. This station is not accurately predicted as these predicted values are lower than the measurements collected by GBRA. This could be caused from the shallowness of the Guadalupe River at this location, with more light being reflected back to the sensor, than being absorbed by the water. As the stations become closer to the City of New Braunfels, the more detection of light there is to the sensor due to the large amount of asphalt that surrounds Station 12653, Comal River at Hinman Island, and Station 15082, Comal River at Landa Park. Station 12653 shows to have a low turbidity value and when compared to the *in situ* measurement, the *in situ* measurement is similar leading to the all date's model accurately predicting this location. The values of the pixels surrounding Station 15082, were accurately predicted from the all date's model. The blue pixel turbidity value is similar to

the measurement collected *in situ*. Overall, the all date's model accurately predicted 80% of the stations that were selected for turbidity measurement. The one station that the model did not accurately predict could have resulted from a bad measurement GBRA which lead to the model being skewed or not having enough observation points which resulted in a model that was no longer used for generalized points but just for those particular observations.

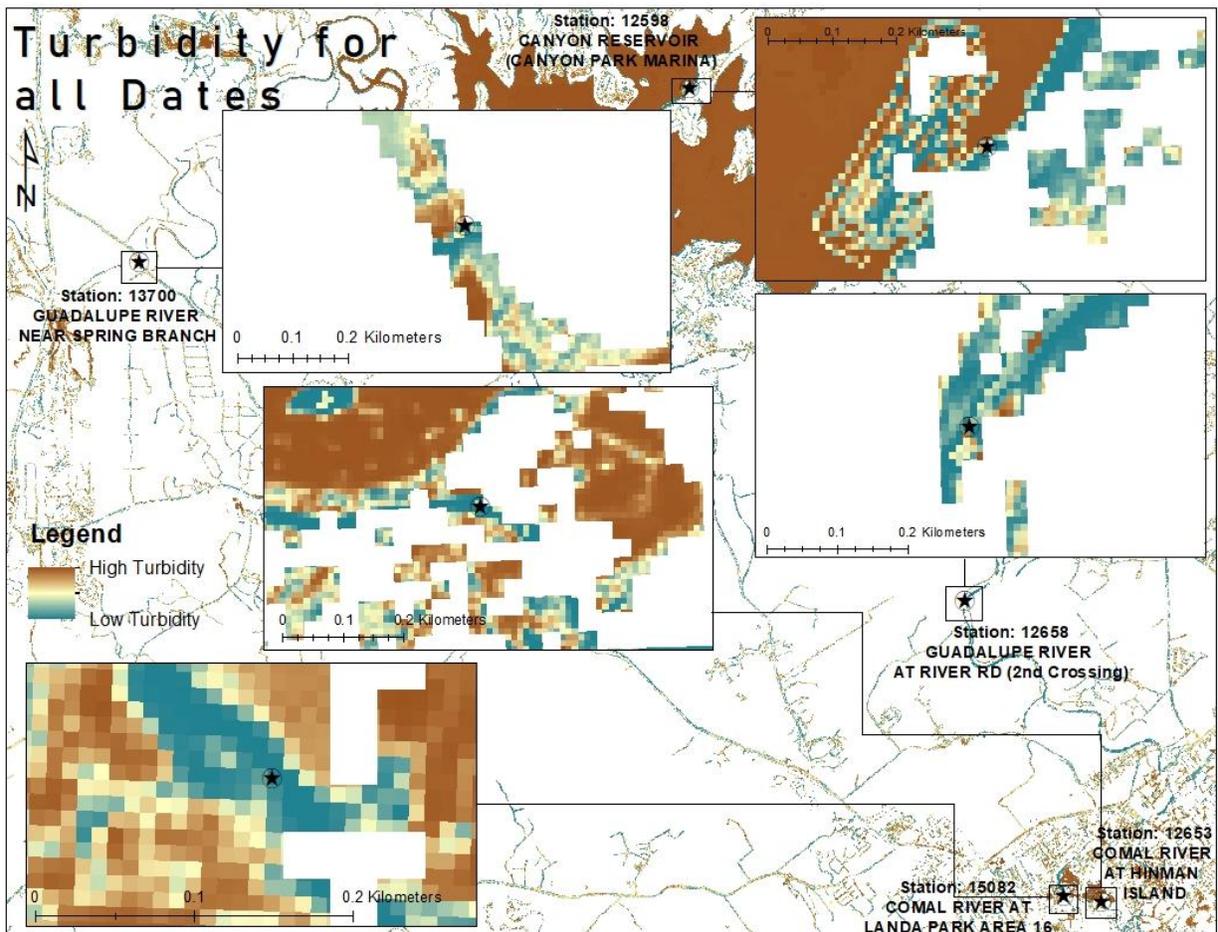


Figure 10: Map of Guadalupe and Comal River showing estimated turbidity using the all date's model.

Figure 11 shows the predictions of TSS using the all date's model. The darker the red the higher the TSS and the darker the blue the lower the TSS. Station 13700, shows to have a

mixture of yellow and red surrounding its location. When comparing the predicted values to the *in situ* collected measurements, the all date's model prediction does not accurately represents the *in situ* collected measurements. The prediction model is higher than the collected measurement. When looking at Station 12598, it is noticeable that the area surrounding the station is a yellow green, this results in a TSS value that is mid value and approaching high. When looking at the predicted value versus the actual value, the all date's model does not accurately predict TSS. The all date's model predicted value is larger than what was collected in the field. Station 12658, shows to have a mixture of TSS with values ranging from high to low. The predicted values for this station are also higher than the measurements that GBRA collected. Station 12653, shows to have a range of TSS of the predicted values. In comparing the predicted values to the actual values, the predicted values are slightly higher than the actual values, but they are very similar, resulting in an accurate prediction. Predicting TSS for Station 15082, resulted in TSS values ranging from medium to high. In the comparison, the predicted values again where higher than the values that were *in situ* collected. Overall, TSS could not be accurately predicted.

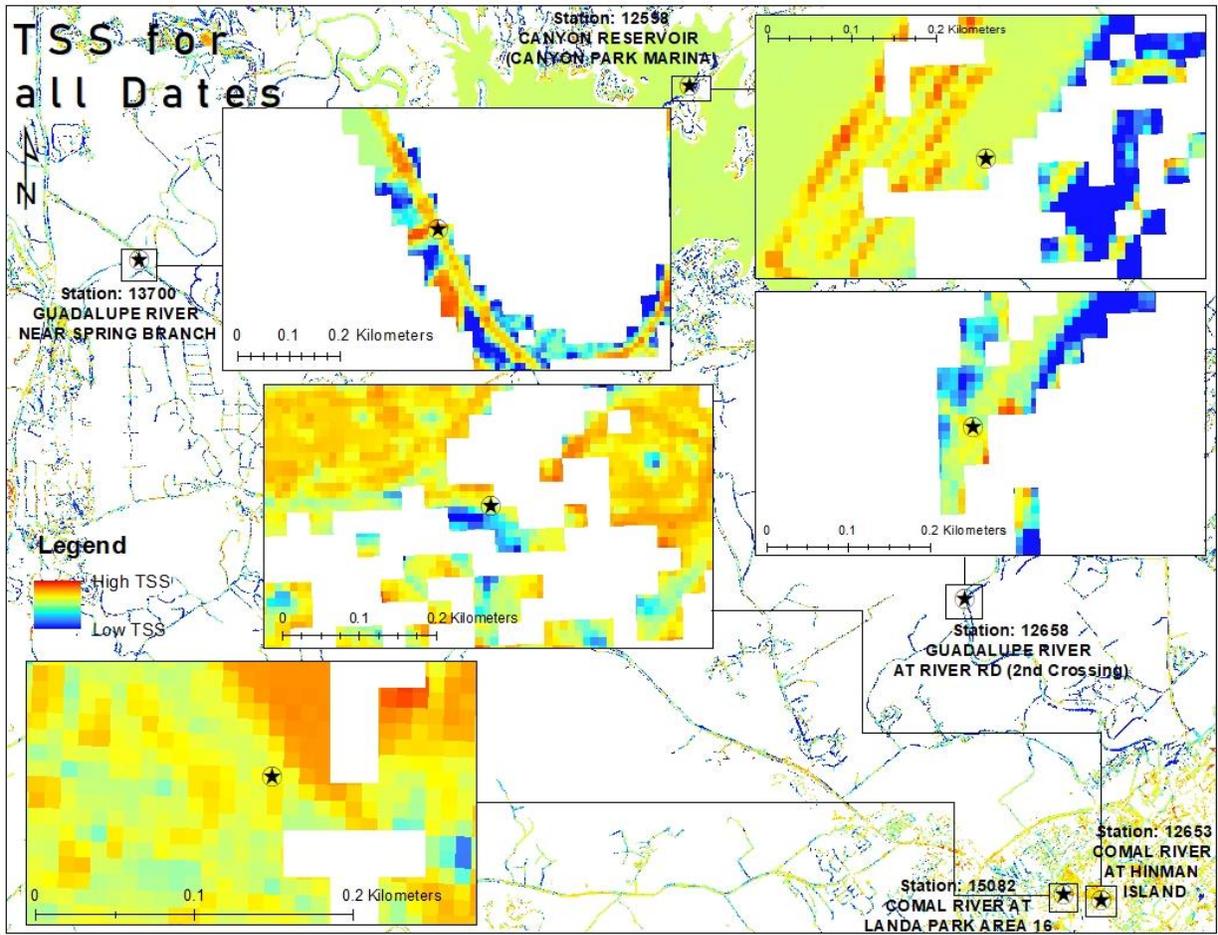


Figure 11: Map of Guadalupe and Comal River showing TSS using the all date's model.

When breaking the data down by observation date, starting with December 2016, the amount of solids within the water was low as can be seen in Figure 12. Station 13700 near Spring Branch, shows to have predicted low TSS. When comparing that value to the value GBRA collected they are very similar, meaning that the model accurately predicted TSS for that station. Station 12598, Canyon Reservoir, shows to have low to mid-range TSS. With the comparison between the predicted December model and the measurements collected via *in situ*, the model accurately predicted TSS at this station. At Station 12658, Guadalupe River second crossing, the model predicted that the TSS is low, which was accurate when looking at the comparison between the model and actual measurements. Station 12653, Comal River at Hinman

Island, shows to have low TSS for the month of December. In seeing how that relates to the actual TSS data, the model accurately predicted TSS for that station. Station 15082, Comal River at Landa Park, is predicted to have low TSS and when comparing the predicted values to the actual measurements, the predicted values were lower than the actual measurement. The December model did not accurately predict the TSS as this location. Overall, the December model accurately predicted 80% TSS at those station locations. As there was no significant model created for turbidity in December that data is not available.

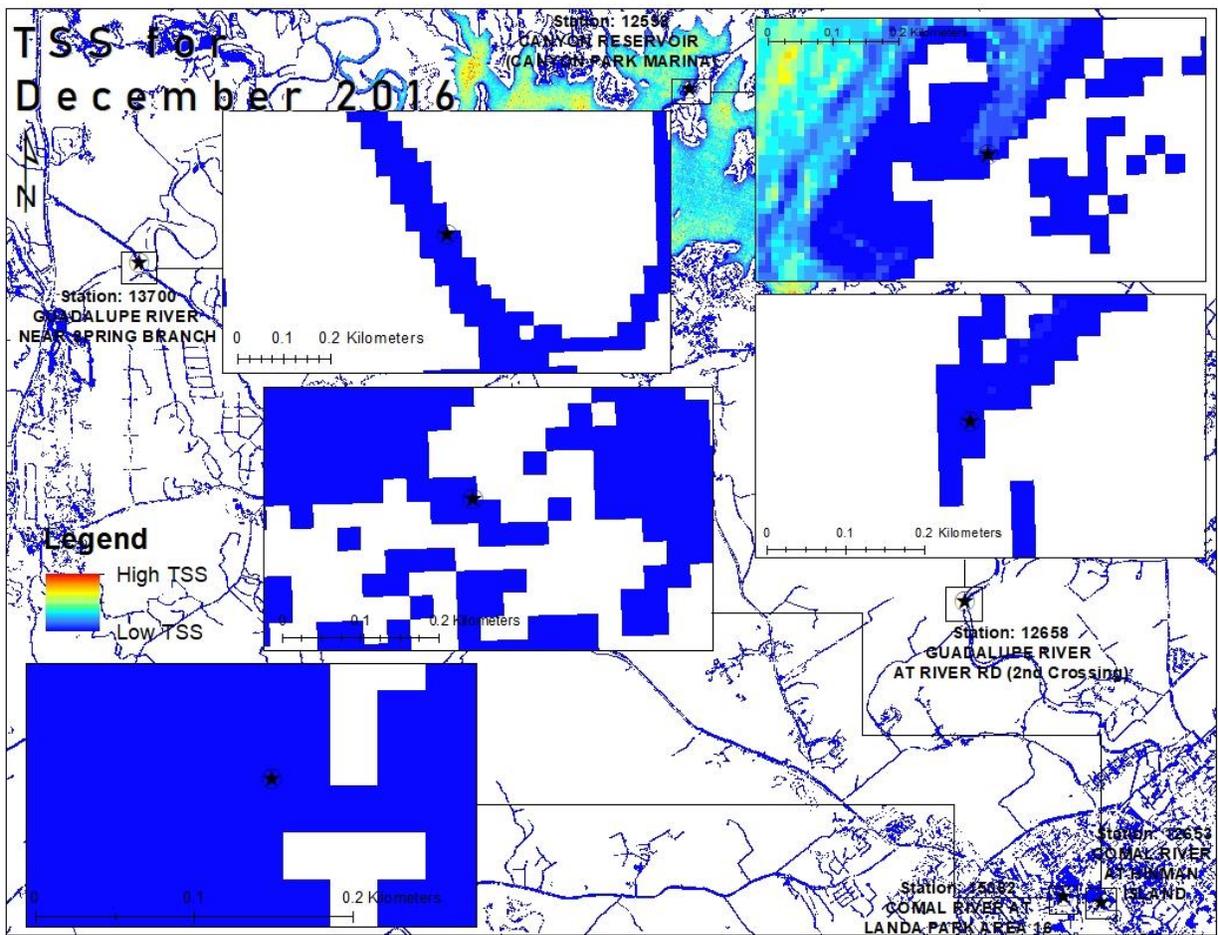


Figure 12: Map of Guadalupe and Comal River showing TSS using the Dec 2016 model.

When analyzing the turbidity image from January 2017, it is noticeable that there is a lot of low turbidity areas, but upon further inspection of each station, it can be seen that this is not

necessarily true. At Station 13700, there is low turbidity with one section of mid-range turbidity. In comparing the predicted January model to the actual *in situ* collected data, the predicted model accurately predicts the turbidity at that location. Station 12598, at Canyon Lake Marina, show to have high turbidity and low turbidity, but when comparing the model to the measurement, the model's prediction was higher than what was collected by GBRA. Station 12658, Guadalupe River at the second crossing, shows to have section of high turbidity running through the middle of the river. When comparing the model the GBRA collected measurements, the model is lower than the actual collected data, leading to an inaccurate prediction. Station 12653, Hinman Island, shows to have low turbidity all around the location, in comparison, the model accurately predicted the turbidity. Station 15082, Landa Park, was predicted to have low turbidity, when comparing the January model to the actual data, lead to an accurately predicted model. Overall, turbidity was accurately predicted at 60% of the station. Since there was no significant model created for TSS in January that data is not available.

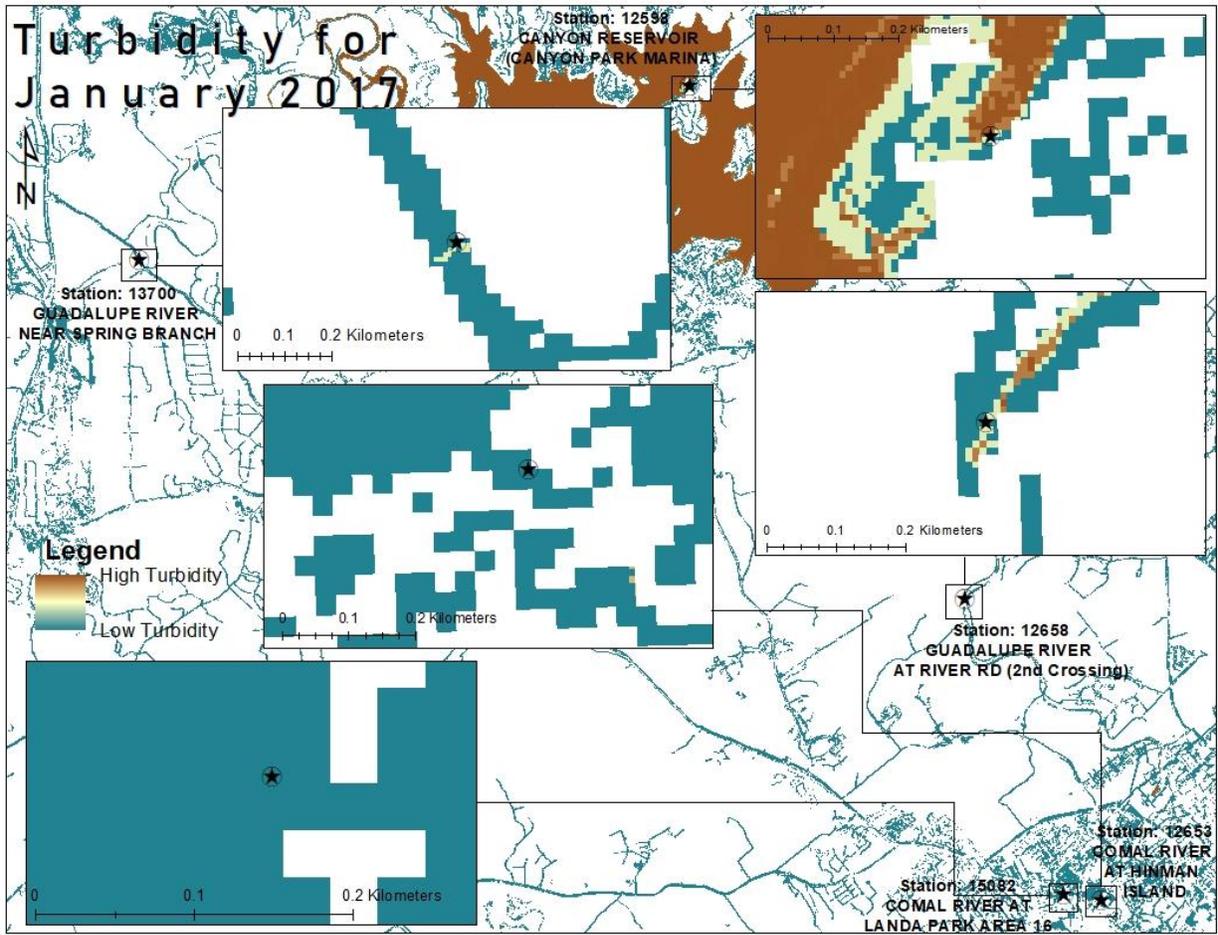


Figure 13: Map of Guadalupe and Comal River showing turbidity using the Jan 2017 model.

In Figure 14, there is an overall trend of having a high TSS prediction values at all stations. Upon further investigation, it can be seen that at Station 12598, Canyon Marina, that the model predicted high TSS, which is an accurate prediction based on the data collected by GBRA. At Station 12658, Guadalupe River at the second crossing, the model predicted high TSS, which is accurately represented in comparison to *in situ* collected measurements. Station 12653, Hinman Island, shows a wide range of high TSS and low TSS. In the comparison, the April model accurately predicted the low TSS. Station 15082, show to also have a mixture of high TSS and low TSS, and where the station is located the model accurately predicted the low TSS. Since Station 13700 was excluded from the April model, it cannot be verified that the model accurately

predicted that station. Overall, the model accurately predicted the TSS at these four stations. Since there was no significant model created for TSS in April that data is not available.

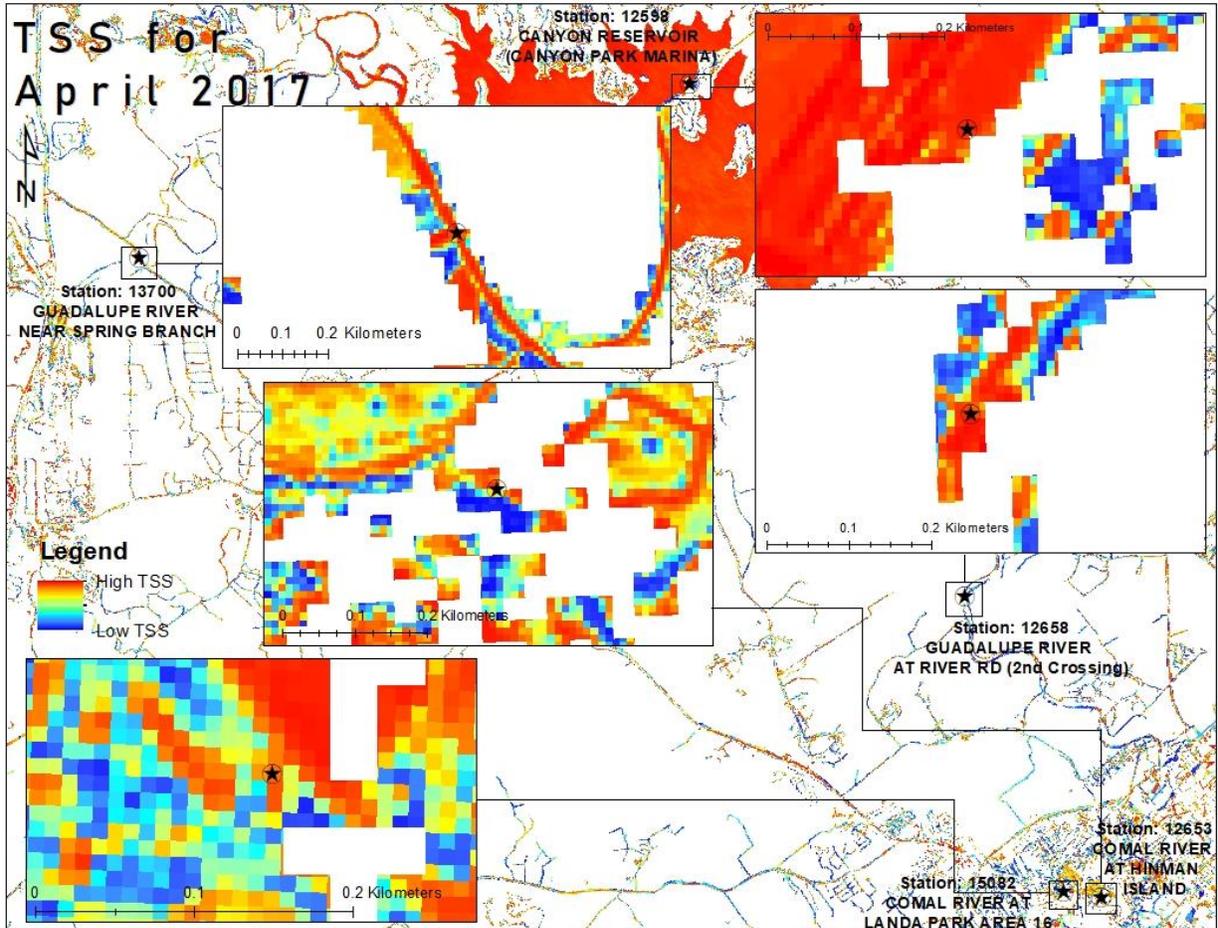


Figure 14: Map of Guadalupe and Comal River showing TSS using the Apr 2017 model.

Initially, there were fifteen total samples that were collected, but due to the high turbidity and TSS values at Station 13700, Guadalupe River near Spring Branch, on April 30, 2017, those values were excluded from the model. The reason this station was excluded was due to late night thunderstorms that passed over the Spring Branch, Texas causing the sediment to stir. Since this part of the river is shallow, it does not take much rain to for the particles to become suspended (WeatherSpark, 2019). According to GBRA, Station 13700 experienced runoff from a rainfall

that occurred on the date of collection. This rainfall washed debris into the water column from the surrounding watershed which increased the turbidity and TSS readings.

A reason as to why the model only accurately predicted a portion of the stations in all the model could be due to overfitting each model. To have a successful model, statisticians recommend having at least 10-15 observation per term. If the model has less than that recommended number, then the model can become overfit. Overfitting the model means that the model no longer is valid if additional observation are added. That overfit model then starts predicting random error which leads to high R^2 values (Frost 2019). For example, in the all date's model, there were 14 observation and two terms, which equates to having seven observations per term. When the individual date's model was ran, there was only 4 or 5 observation points and if a two term model was used, then there would be 2-2.5 observations per term causing the model to be overfit.

Another reason for the models only accurately predicting a portion of the stations could be due to the shallowness of the Guadalupe and Comal Rivers. The shallowness can cause light to be reflected off the bottom and returned back to the Sentinel-2 sensor leads to the false prediction values.

Conclusion

Although research has been done with comparing satellite reflectance values for turbidity and total suspended solids with *in situ* data measurement, these waterways include deep water, such as oceans and lakes, and are also located in highly populated, polluted areas outside the United States. If there are studies done in the United States they are up North in the Great Lakes region or the New York Harbor. There is very little research on the water quality of rivers in the Southern United States. Through the comparison of reflectance values and linear regression, I have produced a water quality assessment of the five testing locations along the Guadalupe and Comal Rivers that flows through New Braunfels, Texas.

Using remote sensing techniques, water authorities, such as the Guadalupe Blanco River Authority, can get the data quicker, cheaper and more often. With these factors, they are able to update the stakeholders of their findings more frequently so if there is trouble with the water quality then preventative measures can be taken quickly.

Future studies will include taking testing locations in the upper and lower portions of the Guadalupe River, and comparing those reflectance values and *in situ* measurement to get a better understanding of what the turbidity and total suspended solids look like throughout the whole waterway. After looking at the Guadalupe River, testing this method on another shallow river waterway such as the Baro River in Ethiopia.

Furthermore, the study of water quality is an important aspect of everyday life. Developed countries, such as the United States and Australia, are accustomed to turning on the tap and receiving fresh, clean, clear water that can be used for everything. This water has gone through a water treatment plant and will not contain any of the pesticides or parasites that will make the human population ill. In undeveloped and developing areas of the world, such as

Mexico and Argentina, the water coming out of the tap is not always reliable. People may have to walk great distances to get water for their family usage from the local watering hole which could contain parasites and pesticides. With the use of remote sensing, these developing countries can use this technology to determine the quality of their water and enlist help on how to make it better for their communities and people.

Bibliography

1. Asquith, William & R. Bumgarner, Johnathan. (2014). Linkage of the Soil and Water Assessment Tool and the Texas Water Availability Model to simulate the effects of brush management on monthly storage of Canyon Lake, south-central Texas, 1995–2010.
2. Baumgartner, M. F., & Apfl, G. M. (1996). Remote sensing and geographic information systems. *Hydrological Sciences Journal*, 41(4), 593–607.
<https://doi.org/10.1080/02626669609491527>
3. Chen, Z., Hu, C., & Muller-Karger, F. (2007). Monitoring turbidity in Tampa Bay using MODIS/Aqua 250-m imagery. *Remote Sensing of Environment*, 109(2), 207–220.
<https://doi.org/10.1016/j.rse.2006.12.019>
4. Chen, S., Han, L., Chen, X., Li, D., Sun, L., & Li, Y. (2015). Estimating wide range Total Suspended Solids concentrations from MODIS 250-m imageries: An improved method. *ISPRS Journal of Photogrammetry and Remote Sensing*, 99, 58–69.
<https://doi.org/10.1016/j.isprsjprs.2014.10.006>
5. Copeland, C. (2016). CRS Report for Congress Clean Water Act: A Summary of the Law, 1–10.
6. Curran, P., & Novo, E. M. (1988). The relationship between suspended sediment concentration and remotely sensed spectral radiance: a review. *Journal of Coastal Research*, 4(3), 351–368.
7. Esa. "ESA." European Space Agency. Accessed April 29, 2018. <https://www.esa.int/ESA>.
8. ESRI. 2018. "Extract Multi Values to Points." Extract Multi Values to Points | ArcGIS Desktop. ESRI. <http://desktop.arcgis.com/en/arcmap/latest/tools/spatial-analyst-toolbox/extract-multi-values-to-points.htm>.
9. Flatten, C. (2017). The Upper Guadalupe River : Stewarding a Hill Country Icon. Retrieved February 11, 2018, from http://www.hillcountryalliance.org/wp-content/uploads/2017/11/GRA_HCA_UpperGuadalupe_LitReview.pdf
10. Frost, Jim. 2019. "Overfitting Regression Models: Problems, Detection, and Avoidance." Statistics By Jim. March 30. <https://statisticsbyjim.com/regression/overfitting-regression-models/>.
11. Guadalupe Blanco River Authority (GBRA). (2013). *2013 Clean Rivers Program Basin Summary Report: Guadalupe River and Lavaca-Guadalupe Coastal Basins*. Seguin, TX. Retrieved from <http://www.gbra.org/documents/publications/basinsummary/2013.pdf>
12. *Handbook of Texas Online*, Vivian Elizabeth Smyrl, "Guadalupe River," accessed February 11, 2018, <http://www.tshaonline.org/handbook/online/articles/rng01>. Uploaded on June 15, 2010. Published by the Texas State Historical Association.
13. Harrington, J. A., Schiebe, F. R., & Nix, J. F. (1992). Remote sensing of Lake Chicot, Arkansas: Monitoring suspended sediments, turbidity, and Secchi depth with Landsat MSS data. *Remote Sensing of Environment*, 39(1), 15–27. [https://doi.org/10.1016/0034-4257\(92\)90137-9](https://doi.org/10.1016/0034-4257(92)90137-9)

14. Hellweger, F. L., Schlosser, P., Lall, U., & Weissel, J. K. (2004). Use of satellite imagery for water quality studies in New York Harbor. *Estuarine, Coastal and Shelf Science*, 61(3), 437–448. <https://doi.org/10.1016/j.ecss.2004.06.019>
15. Kabbara, N., Benkhelil, J., Awad, M., & Barale, V. (2008). Monitoring water quality in the coastal area of Tripoli (Lebanon) using high-resolution satellite data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 63(5), 488–495.
16. Lim, H. S., MatJafri, M. Z., & Abdullah, K. (2010). TSS monitoring by using THEOS satellite imagery over Penang Island Malaysia. *Proceedings of SPIE - The International Society for Optical Engineering*, 7858(November 2010). <https://doi.org/10.1117/12.869488>
17. Liu, H., Li, Q., Shi, T., Hu, S., Wu, G., & Zhou, Q. (2017). Application of Sentinel 2 MSI Images to Retrieve Suspended Particulate Matter Concentrations in Poyang Lake. *Remote Sensing*, 9(7), 761. <https://doi.org/10.3390/rs9070761>
18. Mabwoga, S. O., Chawla, A., & Thukral, A. K. (2010). Assessment of water quality parameters of the Harike wetland in India, a Ramsar site, using IRS LISS IV satellite data. *Environmental Monitoring and Assessment*, 170(1–4), 117–128. <https://doi.org/10.1007/s10661-009-1220-2>
19. Ogashawara, I., & Moreno-Madriñn, M. (2014). Improving Inland Water Quality Monitoring through Remote Sensing Techniques. *ISPRS International Journal of Geo-Information*, 3(4), 1234–1255. <https://doi.org/10.3390/ijgi3041234>
20. Petus, C., Chust, G., Gohin, F., Doxaran, D., Froidefond, J. M., & Sagarminaga, Y. (2010). Estimating turbidity and total suspended matter in the Adour River plume (South Bay of Biscay) using MODIS 250-m imagery. *Continental Shelf Research*, 30(5), 379–392. <https://doi.org/10.1016/j.csr.2009.12.007>
21. Ritchie, J. C. J., Zimba, P. P. V, & Everitt, J. H. J. (2003). Remote sensing techniques to assess water quality. *Photogrammetric Engineering & Remote Sensing*, 69(6), 695–704. <https://doi.org/10.14358/PERS.69.6.695>
22. Robert, E., Grippa, M., Kergoat, L., Pinet, S., Gal, L., Cochonneau, G., & Martinez, J. M. (2016). Monitoring water turbidity and surface suspended sediment concentration of the Bagre Reservoir (Burkina Faso) using MODIS and field reflectance data. *International Journal of Applied Earth Observation and Geoinformation*, 52, 243–251.
23. Santamaria-del-Angel, E., Millan-Nunez, R., Gonzalez-Silvera, A., & Cajal-Medrano, R. (2011). Comparison of In Situ and Remotely-Sensed Chl-a Concentrations: A Statistical Examination of the Match-up Approach. *Handbook of Satellite Remote Sensing Image Interpretation: Marine Applications*, 293.
24. Spath, B. H. C., Faust, O., Pettersson, L. H., Hamre, T., Vitulli, D., Allen, A. R., & Spracklen, T. (2008). Integration of In -situ and remote sensing Data for Water Risk Management, 486–493. Retrieved from <http://www.iemss.org/iemss2008/index.php?n=Main>.
25. Stricklin, F. L., Jr., Smith, C. I., and Lozo, F. E., (1971). Stratigraphy of Lower Cretaceous Trinity Deposits of Central Texas: Bureau Of Economic Geology Report Of Investigations No. 71, 63p

26. "The Texas Clean Rivers Program." TCEQ. Accessed February 17, 2018. .
27. TNC (The Nature Conservancy). 2007. Conservation Plan for the Pedernales River Watershed. TNRCC (Texas Natural Resource Conservation Commission). 1995. Texas Surface Water Quality Standards. Texas Natural Resource Conservation Commission, Austin, Texas.
28. TPWD (Texas Parks and Wildlife Department). 2000. Annotated County Lists of Rare Species. Wildlife Diversity Program, Wildlife Division, Austin, Texas
29. United States Census Bureau. "The South Is Home to 10 of the 15 Fastest-Growing Large Cities." News release, May 25, 2017. Newsroom. Accessed February 17, 2018. <https://www.census.gov/newsroom/press-releases/2017/cb17-81-population-estimates-subcounty.html>.
30. U.S. Congress. House. Committee on Science and Technology. *Remote Sensing Data: Applications and Benefits*. 110th Cong., 2nd sess., 2008. H. Doc. 110-91
31. U.S. Fish and Wildlife Service. (2013). Federal Water Pollution Control Act (Clean Water Act). Retrieved from <https://www.fws.gov/laws/lawsdigest/FWATRPO.HTML>
32. USGS. "The Effects of Urbanization on Water Quality." Urbanization and Water Quality. December 02, 2016. Accessed April 22, 2018. <https://water.usgs.gov/edu/urbanquality.html>.
33. Wang, J. J., Tian, Q. J., Wang, Y., Li, J., Lei, B., & Yang, J. (2007). Total suspended solids concentration retrieval by hyperspectral remote sensing in Liaodong Bay - art. no. {67901Y}. *Remote Sensing and Gis Data Processing and Applications; and Innovative Multispectral Technology and Applications, Pts 1 and 2*, 6790(November 2007), Y7901--Y7901.
34. "WeatherSpark.com." Weather Spark. Accessed April 07, 2019. <https://weatherspark.com/>.

Definitions

Analysis of Variance – a statistical method used to measure the difference between two or more means.

Atmospheric correction – the process of removing the effect of the atmosphere on the reflectance values of images take by satellites and remote sensors

Atmospheric Scattering – the redirection of electromagnetic energy by suspended particles in the air. There are three types of scattering: Rayleigh scattering: light interacts with molecules/particles that are smaller in diameter than the wavelength of incoming light, scatters blue light; Mie scattering: light interacts with molecule/particles that are similar in diameter to the wavelength of incoming light, scatters near ultraviolet to mid infrared light; and Non-selective scattering: diameter of particles is much larger than the wavelength of incoming light; scatter all visible light.

Bottom of atmosphere reflectance or surface reflectance – accounts for scattering and absorption in the atmosphere and provides pixel values that are spectrally similar to what an observer would measure on the ground during satellite overpass

Correlation of determination (R^2) – assesses how well a statistical model explains and predicts future outcomes. If $R^2 = 0$, then the dependent variable cannot be predicted using the independent variable. If $R^2 = 1$, then the dependent variable is always predicted by the independent variable.

Digital Number (DN) – the value of the pixel the satellite sensor stores.

European Space Agency – an intergovernmental organization of 22 member states dedicated to the exploration of space.

F-test – a test within ANOVA, the measures the dispersal of data points around the mean.

Geolocation – the identification/estimation of the real world geographic location of an object.

Geometric corrections – the process of digitally manipulating an image's data such as the image's projection to precisely match what is seen on the ground.

Hyperion satellite – a sensor launched by NASA on November 21, 2000, has 220 bands, a spatial resolution of 30 meters, was decommissioned on March 30, 2017.

In situ – taken/left in original location.

Landsat 7 Enhanced Thematic Mapper Plus – a sensor launched by NASA on April 15, 1999, has 8 bands, a spatial resolution of 30 meters. Used to study agriculture, forestry, geology, and regional planning.

Landsat Thematic Mapper – a sensor aboard two satellites launched by NASA, the first Landsat 4 launched on July 16, 1982, had 7 bands, a spatial resolution of 30 meters, and decommissioned on June 15, 2001. The second Landsat 5 launched on March 1, 1984, had 7 bands, a spatial resolution of 30 meters, and decommissioned in June 5, 2013. Used to study global warming and climate change.

Landsat Multispectral Sensor – a aboard five satellites (Landsat 1 – 5), had 4 bands, a spatial resolution of 60 meters, and was decommissioned between 1978 and 2013. Used to study agriculture, forestry, geology, and regional planning.

Level-1C imagery – composed of 100 km² tiles, contain radiometric and geometric corrections and top of atmosphere reflectance.

Level-2A imagery – composed of 100 km² tiles, includes atmospheric correction applied to top of atmosphere Level-1C imagery, main output is bottom of atmosphere corrected reflectance product.

LISS IV – an Indian satellite, has 3 bands, and a spatial resolution of 5.8 m. Used to monitor land and water resources.

MODIS – a sensor launched by NASA in 1998, has 36 bands and spatial resolutions ranging from 250 m to 1 km. Use to monitor cloud cover, process occurring the ocean, on land and in the atmosphere.

Positive epistemology – uses the scientific method to answer the research question.

Radiometric corrections – the process of fixing digital numbers when the observed energy does not match with the energy emitted or reflected from the same object.

Red edge – the region of rapid change in reflectance of vegetation in the near infrared band, used to monitor plant health.

Reflectance values – the ratio of the amount of light leaving an object to the amount of light hitting the object.

Remote sensing – the art and science of obtaining information about an object without being in direct contact with that object.

Root Mean Square Error – the difference between values (sample or population) predicted by the model and the values observed.

Secchi Disk depth method – used to gauge the transparency of water by measuring the depth at which the disk ceases to be visible from the surface.

Sen2Cor – a processor for Sentinel-2 Level-2A product generation and formatting.

Sentinel-2 sensor – two sensors owned by the European Space Agency where 2A was launched June 23, 2015 and 2B was launched March 7, 2017. Used to monitor agriculture, emergency management, water quality, and land cover classification.

Spatial – occupying space.

Spectral – a collection of energy that is reflected or emitted from features on Earth's surface.

Spectroradiometer – an instrument used to measure the spectral response of a source.

SPOT sensor – a satellite owned by the French space agency, with launch dates starting in February 22, 1986 through June 30, 2014. Used to detect and forecast phenomena involving climatology and oceanography, along with monitoring human activity and natural phenomena.

Temporal – time.

THEOS sensor – satellite owned by the French space agency, launched October 1, 2008, has 4 bands and a spatial resolution of 15 m. Used to monitor Thailand.

Top of Atmosphere reflectance – a unit less measurement which provides the ratio of incoming light to the emitted light on the surface of an object.

Total Suspended Solids – the dry weight of suspended particles that can be trapped using a filtration system.

Turbidity – the cloudiness or haziness of a fluid.

Water column – the area from the surface of a sea, river or lake to the bottom sediment.

Water Quality – the physical, chemical, and biological aspects of water.