Modeling Suspended Solids in the San Francisco Bay Estuary with Remote Sensing Data

Johanna Press

Advisor: Ron Smith
Second Reader: Xuhui Lee
April 29th, 2015

A Senior Thesis presented to the faculty of the Department of Geology and Geophysics, Yale University, in partial fulfillment of the Bachelor's Degree

Abstract

Suspended solids concentration (SSC) is a key determinant of water quality in coastal and estuarine waters. High SSC in the San Francisco Bay estuary necessitates consistent and thorough monitoring to mitigate adverse effects on ecosystems and water resources; however, the current *in situ* monitoring program collects only monthly samples from fixed locations. This paper presents remote sensing tools for optical suspended solids measurement, working towards establishing a method for continuous, synoptic water quality monitoring in the Bay. Section I tests semi-analytical and empirical SSC algorithms using data from the Hyperspectral Imager for the Coastal Ocean (HICO); the greatest predictive power is found in a linear model derived from the 634nm band. Modeled SSC is applied to predict the spatial distribution of trace element pollutants. To improve on the spatial and temporal limits of HICO, Section II derives a statistical SSC model from Landsat 8 data that accounts for the optical effects of environmental factors including bottom depth, wind speed, tide phase, and temporal offset between *in situ* and satellite measurement. The model verifies the significance of near infrared reflectance, bottom depth, and temporal offset in predicting *in situ* SSC from optical measurement, but it is not generally applicable beyond the calibration dataset. Section III investigates the optical behavior of the most disruptive environmental factor—seafloor reflectance—using spectrometry. An optical model is developed to calculate the backscatter properties of suspended solids over low- and high-reflectivity backgrounds. The model is unsuccessful where the reflectance of the suspended solids and of the background are of a similar magnitude, suggesting particular environmental conditions that are problematic for optical algorithms.

Contents

Introduction

Suspended solids are an important component of the San Francisco Bay estuarine ecosystem. The silts, clays, and sands that form the benthic substrate are often transported in suspension, carrying nutrients necessary for biological productivity as well as potentially toxic trace metal and organic pollutants adsorbed to suspended particles (Domagalski and Kuivila 1993; A.R. Flegal et al. 1991; Hammond et al. 1985). Annual suspended sediment flux to the San Francisco Bay has been estimated at 1.2 million metric tons, with over 85% entering as fluvial discharge from the Central Valley watershed (McKee, Ganju, and Schoellhamer 2006). Suspended solids distribution throughout the estuary depends on salinity gradient, bottom topography, tidal currents, and wind-wave resuspension (Schoellhamer 1996).

Suspended solids concentration (SSC) measures the quantity of particles larger than 2 microns in a volume of water, including inorganic and organic sediment as well as living microorganisms (U.S. Geological Survey 2014b). Despite the importance of sediment to estuary biology and geomorphology, high levels of SSC in the San Francisco Bay can be detrimental to both estuarine ecosystems and anthropogenic activities: suspended solids carry adsorbed contaminants, attenuate sunlight in the water column, and deposit on tidal marsh and intertidal mudflats as well as in ports and shipping channels (Buchanan et al. 2014; Schoellhamer 1996).

To monitor SSC in the San Francisco Bay, the U.S. Geological Survey (USGS) operates an *in situ* sampling program, acquiring measurements from 36 fixed locations in the Bay and surrounding estuary on a monthly scientific cruise (Figure 1). Optical measurements of turbidity are made at various heights in the water column as a proxy for SSC and results are projected to predict SSC for the full water column at that site (Buchanan et al. 2014; U.S. Geological Survey 2014a). Although the USGS SSC sampling program provides essential data for environmental management in the Bay, its data are limited temporally and spatially to fairly narrow windows; it is difficult to understand the full scheme of SSC transport from monthly measurements at discrete locations. In addition, the sampling cruises are a significant expense for this publicly-funded organization.

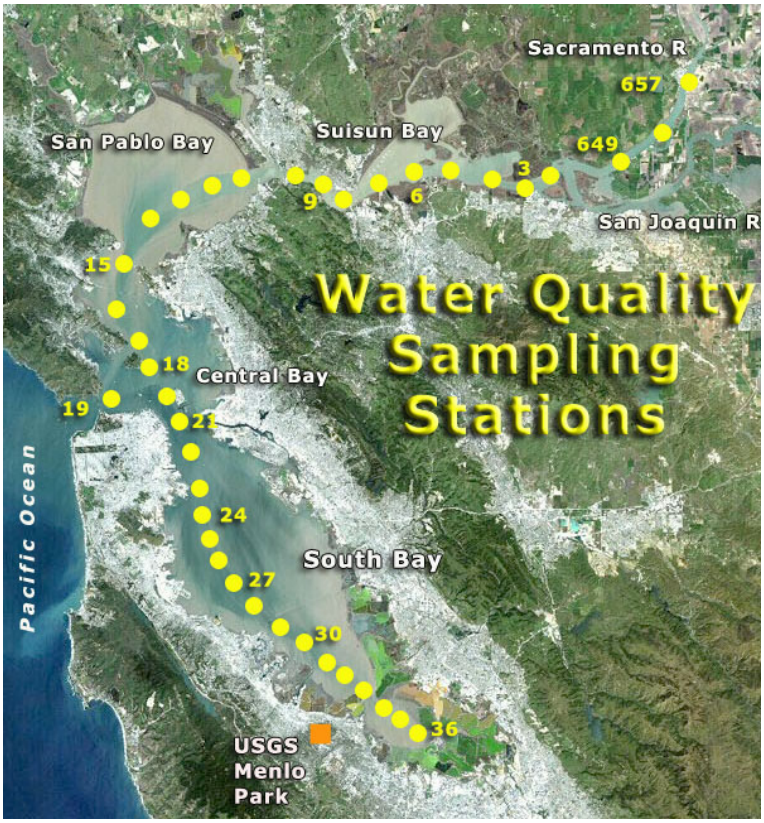


Figure 1: USGS SSC sampling sites. From San Francisco Estuary Institute (2014).

Autonomous optical measurement can serve as a tool to enhance SSC monitoring. In addition to the monthly water sampling cruise, USGS also operates seven *in situ* optical sensors that automatically measure turbidity at 15-minute intervals (U.S. Geological Survey 2014b). These sensors have very high temporal resolution but narrow spatial coverage. Satellite remote sensing, on the other hand, provides synoptic spatial coverage of the San Francisco Bay estuary, with spatial, temporal, and spectral resolution varying by sensor. Previous work has been done to supplement USGS *in situ* measurement of SSC with optical assessment using data from NOAA's Advanced Very High Resolution Radiometers (AVHRR). The 1.09km spatial resolution of these sensors is useful for detecting large features such as Bay-scale SSC gradients and sediment plumes, but large pixels prohibit detection in proximity to land features and obscure small-scale patterns integral to local ecosystem functions (Ruhl et al. 2001). Moreover, only two of AVHRR's broad spectral bands are within the wavelength range of ocean color sensing, limiting precision of SSC measurement (National Oceanic and Atmospheric Administration n.d.). While

useful for studying broad trends in the distribution of suspended solids, the AVHRR method was not able to replace monthly *in situ* sampling.

In the past five years new satellite sensors with higher spectral and spatial resolution have been developed, enhancing capacity for remote monitoring of SSC. This paper investigates remote sensing tools using the recently-launched Hyperspectal Imager for the Coastal Ocean (HICO) and Landsat 8 sensors to establish a method for continuous and synoptic water quality monitoring in the San Francisco Bay. Statistical and optical approaches are used to develop, assess, and apply methods of measuring SSC from satellite imagery.

Section I: Hyperspectral measurement of suspended solids for remote trace element monitoring in the San Francisco Bay estuary

Introduction

The concentrations of total and dissolved trace elements in the water column are key determinants of water quality for estuarine and coastal waters. High concentrations of trace elements in the San Francisco Bay adversely affect both biological and anthropogenic systems by contaminating organisms and ecosystems and rendering estuaries unsafe for human use (A.R. Flegal et al. 1991; Luoma and Phillips 1988). Toxic levels of trace metals discovered in phytoplankton, fish, waterfowl, and sediment within the Bay's drainage basin prompted the establishment in 1993 of the Regional Monitoring Program (RMP) by the San Francisco Estuary Institute to monitor trace element levels regionally (San Francisco Estuary Institute n.d.). Elevated levels of trace elements are attributed largely to anthropogenic activity, but episodic events such as floods, storms, and algal blooms may have significant and as-of-yet not understood impacts on biogeochemical cycles (A Russell Flegal et al. 2005; Nichols et al. 1986). Understanding the spatial and temporal distribution of trace elements is essential to identifying their sources, motions, and effects.

Benoit, Kudela, and Flegal (2010a) studied the concentrations of 13 trace elements (Ag, As, Cd, Co, Cr, Cu, Fe, Hg, Mn, Ni, Pb, Se, and Zn) and methylmercury (MeHg) sampling during the first 10 years of RMP monitoring and demonstrated correlation between *in situ* measurements of SSC and the measured concentrations of these trace elements. From these relationships, predictive models based on SSC were developed to estimate the distribution ratio *D*—a measure of degree of extraction in solid versus aqueous form—as well as water column total concentration (WCT) and water column dissolved concentrations (WCD) of each element. The influences of geographic location within the bay and season of sampling were also incorporated into the models.

Modeling dissolved trace element concentrations is valuable for monitoring biogeochemical cycling within the estuary on a regular basis, as well as predicting fluxes to adjacent coastal waters where few *in situ* measurements exist. The modeling method of Benoit, Kudela, and Flegal (2010a) is limited in that it relies on *in situ* samples of SSC collected monthly by the USGS at discrete, fixed monitoring stations in the San Francisco Bay. Remote sensing,

however, can provide synoptic coverage to better capture spatial and temporal patterns of SSC, and consequently, to better predict trace element concentrations.

Methods and Results

Satellite imagery of the San Francisco Bay was chosen to align with the dates of annual RMP trace metal sampling. Image 8054 from the Hyperspectral Imager for the Coastal Ocean (HICO) containing the area of interest was taken on September 17 2011, during the RMP sampling period from September 13-21 2011. HICO, mounted on the International Space Station, collected data from 2009 to 2014 in 87 bands from 400 – 900 nm at 5.7 nm spectral resolution with spatial resolution of approximately 100m in a 6.92° (42 x 192 km) field of view ("HICO - Sensor and Data Characteristics" n.d.; Lucke et al. 2011). L1B imagery was processed by Dr. Raphael Kudela at UC Santa Cruz using the Tafkaa algorithm for atmospheric correction. The image was georeferenced with a HICO rad_geom file and tuned using Google Earth as a reference image for ground control points. The area of interest was selected in a spatial subset.

Figure 2: 87-band image cube of HICO image 8054; 42-27-11-RGB in front.

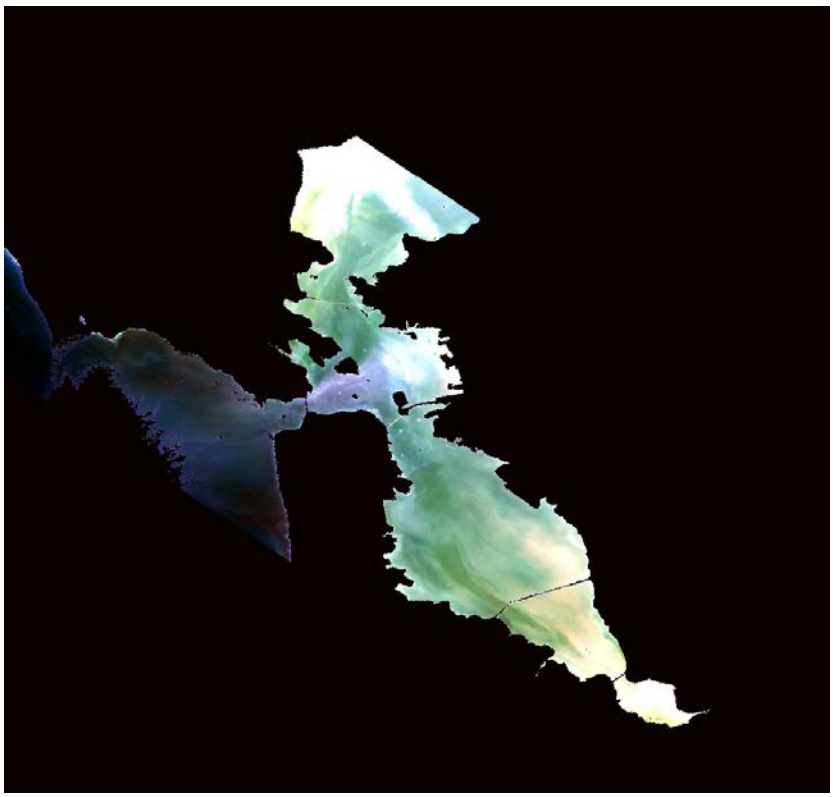


Figure 3: Area of interest with non-water pixels masked.

Coastal from NOAA's GSHHS shoreline database were subset in GEODAS and combined with the boundaries of the area of interest in ArcMap to create a polygon covering land areas, which was used to mask land pixels in ENVI. Maximum likelihood classification was used to mask non-water pixels containing clouds, surf, and man-made objects like bridges (Figure 3).

Based on the review of methods for calculating SSC from remote sensing data in Odermatt et al. (2012), several different algorithms appropriate for a range of approximately 5 – 30 g/m^3 total suspended material were tested using bandmath in ENVI. The results of these algorithms were compared to *in situ* measurement of SSC from September 20 2011 in the USGS dataset on Water Quality of the San Francisco Bay (U.S. Geololgical Survey n.d.). Matlab code was written to compare measured SSC values at geographic coordinates of sampling with modeled SSC of the corresponding pixel using ordinary least squares regression. The semi-analytical method of Nechad, Ruddick, and Park (2010) failed to tune to the area of interest, and no meaningful results were obtained. Empirical algorithms tested include a band ratio of near infrared and visible wavelengths suggested by Doxaran, Froidefond, and Castaing (2002)

(Equation 1), and single band proxies in the visible and NIR, suggested by Miller and McKee (2004) (Equation 2) and Binding et al. (2010) (Equation 3), respectively.

$$\frac{R_{rs}(852)}{R_{rs}(548)};\ R^2 = 0.02 \quad (1)$$

$$R_{rs}(749);\ R^2 = 0.50 \quad (2)$$

$$R_{rs}(668);\ R^2 = 0.74 \quad (3)$$

The relative success of single band methods, although originally designed for sensors with broader spectral resolution, prompted the testing of all narrow HICO bands in the visible and NIR range for correlation with SSC (Figure 4). Remote sensing reflectance (R_{rs}) in the 634nm band had the best predictive power ($R^2 = 0.794$) for *in situ* SSC.

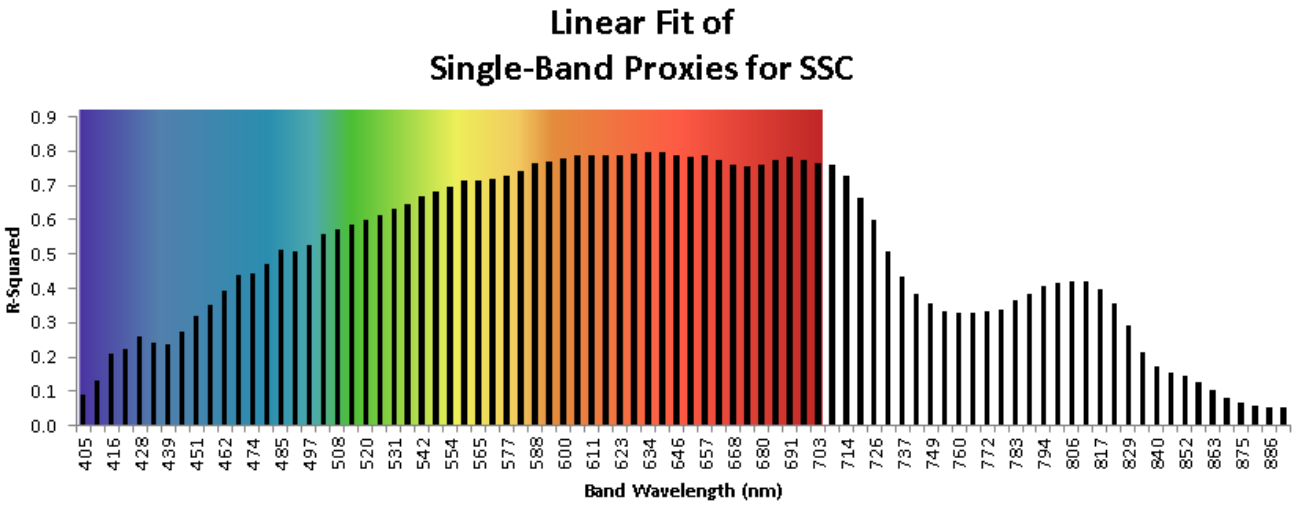


Figure 4: R-squared values of linear models created from R_{rs} in individual HICO bands, with the visible spectrum in the background. Models based on red wavelengths (600-700nm) exhibit highest predictive power.

The model calculated with ordinary least squares regression (Equation 4) was applied using bandmath to calculate an SSC band (Figure 5).

$$SSC = 0.10795 \times R_{rs}(634) + 0.514968 \quad (4)$$

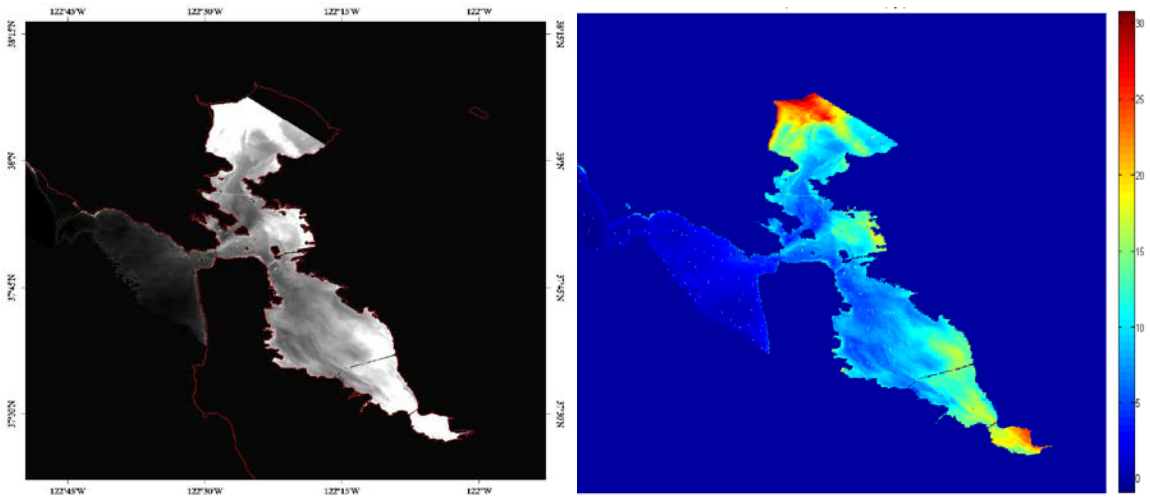


Figure 5: Map of R_{rs} at 634nm(left) and the calculation of SSC in mg/L (right) based on 4.

The models of Benoit, Kudela, and Flegal (2010a) calculate water column total concentration (WCT; µg/L) and distribution ratio (*D*; L/kg) from linear relationships with SSC (mg/L), and then water column dissolved concentration (WCD; µg/L) from these two values according to Equations 5-7.

$$\mathrm{WCT} = m \times \mathrm{SSC} + b + \alpha + \varepsilon + \delta \quad (5)$$

$$\log D = m \times \mathrm{SSC}_{\mathrm{root}} + b + \alpha + \varepsilon + \delta \quad (6)$$

$$\mathrm{WCD} = \frac{\mathrm{WCT} \times 10^6}{D \times \mathrm{SSC} + 10^6} \quad (7)$$

The models predict 12 trace elements measured annually by the RMP[1]: Ag, As, Cd, Co, Cu, Fe, Hg, Mn, Ni, Pb, Se, Zn. Equations were applied in Matlab with coefficients *m* and *b* for each element's equations as well as parameters to adjust for geographic region of the bay (α), season (ε), and interactions of the two (δ) taken from the supporting information (Benoit, Kudela, and Flegal 2010b). The models account for 7 geographic regions in the San Francisco Bay estuary as shown in Figure 6 and 3 seasons; the area of interest in this image, taken during the summer period, includes 4 geographic regions. Maps for the ten elements yielded meaningful non-negative results for WCT and the six for WCD are shown in Figure 7 and Figure 8, respectively.

[1] As noted above, Benoit, Kudela, and Flegal (2010a) also examines Chromium and methylmercury but are unable to create a statistically significant model for WCT or WCD of these pollutants.

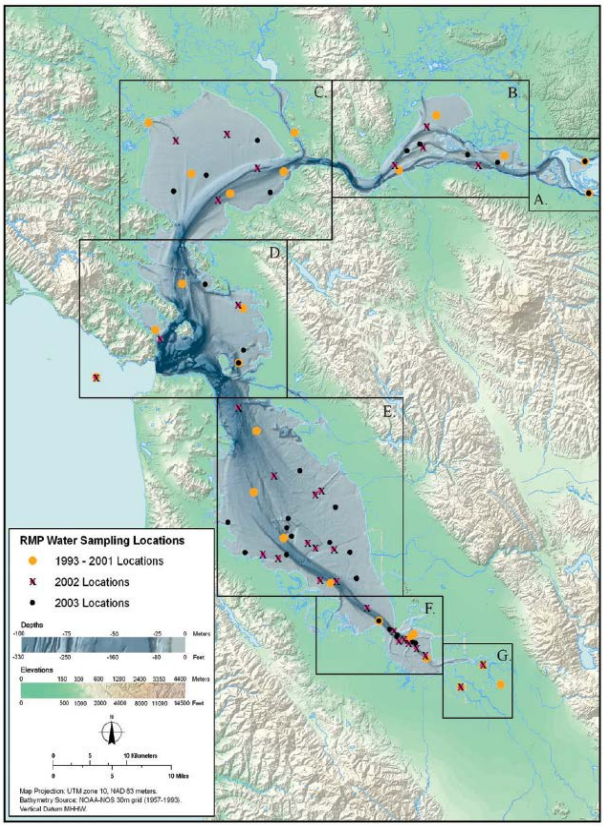


Figure 6: Geographic regions in the San Francisco Bay Estuary. From Benoit, Kudela, and Flegal (2010a).

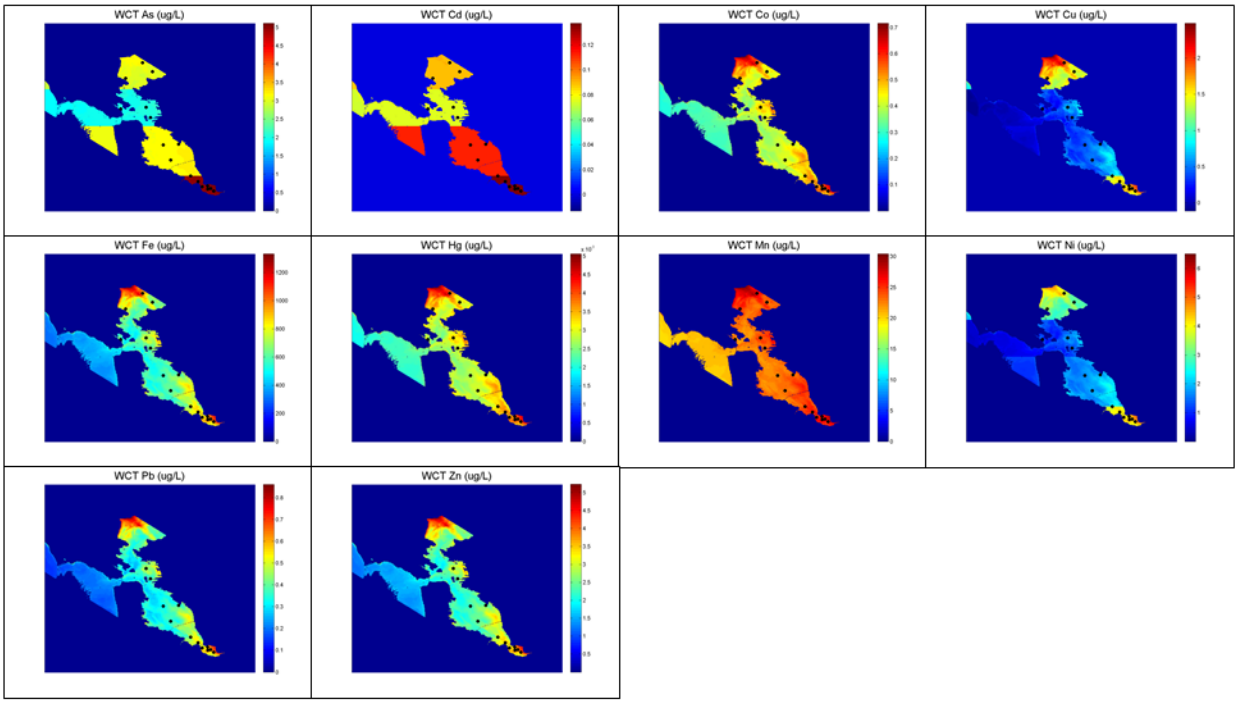


Figure 7: Maps of modeled WCT for As, Cd, Co, Cu, Fe, Hg, Mn, Ni, Pb, and Zn.

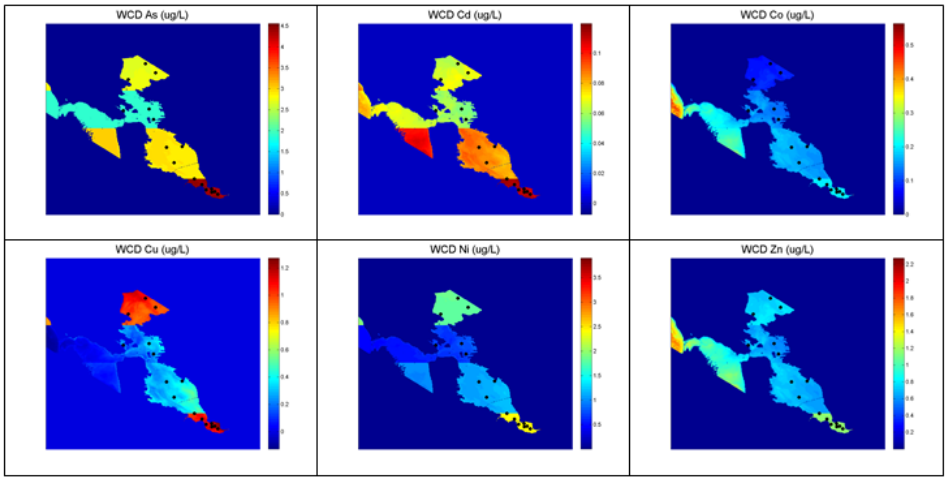


Figure 8: Maps of modeled WCD for As, Cd, Co, Cu, Ni, and Zn.

Matlab code was written to compare *in situ* WCT and WCD values from the 2011 RMP sampling period at geographic coordinates of sampling with modeled values of the corresponding pixel using robust regression[2]. Data were available from 6 – 16 sampling sites within the area of interest, varying by pollutant. Table 1shows the wide spread of R^2 values in the initial models, from 0.18 to 0.70 for WCT and 0.41 to 0.62 for WCD. Few models were significant to the $p < 0.01$ level.

[2] Supporting information describes the use of robust regression in developing and testing models to minimize the influence of outliers for data with a non-normal distribution (Benoit, Kudela, and Flegal 2010b).

Element	WCT		WCD	
	R^2	*p*	R^2	*p*
Ag	-	-	-	-
As	0.70510	0.00005	0.56354	0.00081
Cd	0.49784	0.00226	0.55339	0.00095
Co	0.29633	0.05857	0.50234	0.40502
Cu	0.36750	0.01280	0.62717	0.00240
Fe	0.24818	0.08830	-	-
Hg	0.18152	0.47276	-	-
Mn	0.22187	0.08992	-	-
Ni	0.36302	0.02578	0.41508	0.01473
Pb	-	-	-	-
Se	-	-	-	-
Zn	-	-	0.48018	0.15299

Table 1: Predictive power and significance of initial models for WCT and WCD.

To assess the success of the models, normal quantile plots were created in Matlab for sampled and modeled populations. Sample populations for Co, Cu, Fe, hg, Mn, Ni, and Pb (WCT and WCD), as well as Zn (WCD only) exhibited a logarithmic distribution, and were transformed before repeating the robust regression. This improved fit, but many models still exhibited discrete grouping in the data, as seen in the characteristic banding in the map on the left in Figure 9. This was determined to be an effect of geographic region. Numerous models retained geographic groups that introduced non-uniform variance into the modeled concentrations, acting as influential points and distorting the line of best fit.

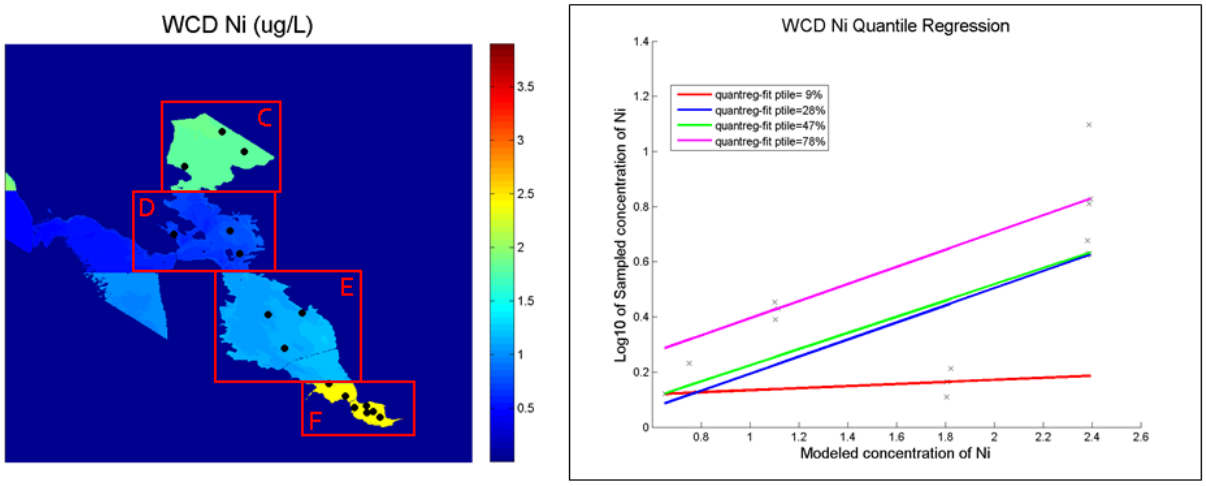


Figure 9: Four modeled geographic regions (left) and corresponding quantile regression lines (right) for WCD of Ni. The top quantile (in red) introduces non-uniform slope, indicating non-uniform variance in the model.

To determine which region was the source of the outlier group, quantile regression was run to estimate the line of best fit at each subset boundary, or quantile, in the cumulative distribution function, displaying the independent contribution of each geographic region to the full regression; an example is shown in Figure 9. Models whose quantile regression lines had uniform slopes were confirmed to have uniform variance for all regions. For models that exhibited a wide spread or large variation in the slopes of quantile regression lines, quantile regression was re-run removing each of the four regions in turn to determine which was introducing non-uniform variance. For each model, plots of every combination of three out of the four geographic regions were examined and the combination that exhibited the most uniform quantile regression line slopes was selected, as in Figure 10. Co, Fe, Hg, Mn, and Ni required the removal of region C.

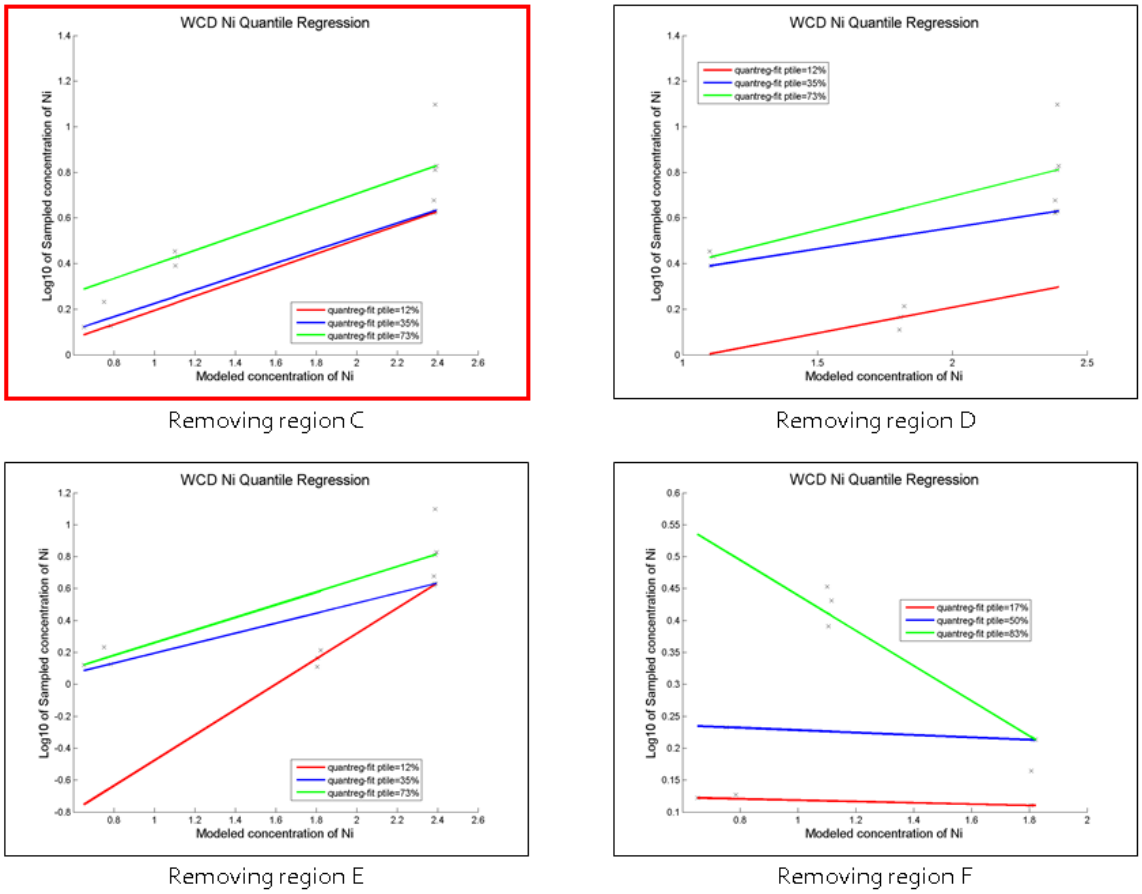


Figure 10: Quantile regression for WCD of Ni removing each of the four geographic regions in turn. Removing region C (upper left) leads to the most consistent slopes across quantile regression lines.

In every case, it was found that the quantile regression excluding region C exhibited the most uniform slopes and the narrowest spread; it was therefore determined that this region, North Bay (west), was introducing non-uniform variance and influential points into the model. Comparing the calculated SSC band with a USGS cross section of measured SSC in the water column, high values in the North Bay were identified as false-positives from sunglint. Sunglint is visible in this area in the original image, justifying the removal of points in region C from the model. With populations adjusted by transformation and regional selection as necessary, robust regression was run a final time to determine the predictive power of the models for *in situ* element concentrations. Appendix 1.

Table 2 shows significantly improved predictive power for most models. Slope (*m*) and intercept (*b*) values for the final models can be found in Appendix 1.

Table 2: Predictive power and significance of models for WCT and WCD after log transformations and sun-glint adjustments applied as necessary.

Element	WCT			WCD		
	R^2	*p*	R^2 from Benoit et al.	R^2	*p*	R^2 from Benoit et al.
Ag	-	-	0.592	-	-	0.01
As	0.70510	0.00005	0.444	0.56354	0.00081	0.47
Cd	0.49784	0.00226	0.125	0.55339	0.00095	0.08
Co	0.73653	0.00017	0.785	0.44311	0.01301	0.20
Cu	0.51844	0.00166	0.552	0.62717	0.00240	0.55
Fe	0.73009	0.00020	0.946	-	-	0.03
Hg	0.65356	0.00082	0.547	-	-	0.05
Mn	0.61231	0.00157	0.532	-	-	0.38
Ni	0.77249	0.00008	0.690	0.77498	0.00008	0.45
Pb	0.68190	0.00094	0.846	0.37348	0.01550	0.03
Se	-	-	0.351	-	-	0.59
Zn	0.58573	0.00372	0.858	0.40972	0.00995	0.06

Discussion

Statistically significant ($p \leq 0.01$) results for WCT were found for 10 elements—As, Cd, Co, Cu, Fe, Hg, Mn, Ni, Pb, and Zn—with R^2 values between 0.49 and 0.77. Predictive power was greater that of the original models in Benoit, Kudela, and Flegal (2010a) for five of these elements. WCD was successfully modeled for six elements—As, Cd, Co, Cu, Ni, and Zn—with R^2 values between 0.37 and 0.77, and predictive power for each element exceeding that of the original models. These results suggest that this method can significantly predict total and dissolved trace element concentrations in the San Francisco Bay, and demonstrate the potential of monitoring programs using remote observation of trace element concentrations. Table 3 compares modeled WCT values with EPA water quality criteria. Concentrations of three elements—As, Fe, and Pb—were found to exceed EPA water quality criteria, and four others—Cd, Cu, Mn, and Ni—were present at more than 50% of the recommended concentration (U.S. Environmental Protection Agency 2014b). Under the National Recommended Water Quality Criteria, Arsenic was present at 28425% of the concentration recommended for human health and Iron was present at 114% of the concentration recommended for aquatic life; Lead was present at 72750% of the concentration listed in National Primary Drinking Water Regulations (U.S. Environmental Protection Agency 2014a). All elements were present at high

concentrations in South Bay (south), or region F in Figure 6. Despite a high density of sampling sites in this area of the Bay, the highest concentrations of trace elements are found to the west of all *in situ* sampling sites, and therefore were not detected by RMP monitoring. This is one significant benefit of the synoptic spatial coverage that can be achieved with this monitoring method.

Table 3: Comparison of modeled WCT values with limits set by EPA water quality standards.

Element	Max. WTC (µg/L)	EPA Limit[3] (µg/L)	WCT/Limit
As	5.117	0.018^	28425%
Cd	0.138	0.250ǂ	55.0%
Cu	2.005	3.10ǂ	64.66%
Fe	1141.8	1000.0ǂ	114.18%
Hg	0.0051	0.94ǂǂ	0.54%
Mn	30.454	50.0^	60.91%
Ni	6.509	8.20ǂ	79.38%
Pb	0.728	<0.001ǂǂ	72750%
Zn	5.232	81.0ǂ	6.46%

Conclusion

This section validates a method for remote trace element monitoring using hyperspectral imagery. Successful trace element models using HICO data expand the spatial range of *in situ* measurement with synoptic coverage of the entire San Francisco Bay estuary, but are limited in in spatial resolution by the sensor's 90m pixel size and temporal resolution by its to-order imaging. More frequent images are necessary to establish typical SSC and trace element conditions, observe change over time, understand their function in biogeochemical cycling, and analyze conditions following episodic events such as storms, floods, and algal blooms (A Russell Flegal et al. 2005). For these reasons, the extension of the method developed in this section to other satellite sensors is an essential step to realizing a remote water quality monitoring program.

[3] ^ National Recommended Water Quality, Human Health Criteria
ǂ National Recommended Water Quality, Aquatic Life Criteria
ǂǂ National Primary Drinking Water Regulations

Section II: Assessing the effects of environmental conditions in Landsat 8 measurement of suspended solids concentration

Introduction

The most significant barrier to extending remote water quality monitoring to different sensors is the determination of an algorithm using remote sensing reflectance to calculate SSC, which is both the key determinant of the accuracy of remote trace element monitoring and a significant water quality parameter in itself. To detect suspended solids from satellite imagery, a relation must be established between patterns in reflectance across bands—a pixel's spectral shape—and SSC. The reflectance of water is highly influenced by its constituent components; in particular, chlorophyll from phytoplankton, color dissolved organic material (CDOM), and suspended solids control the reflectance spectrum of optically complex case II waters (Morel and Prieur 1977). Reflectance can also be influenced by environmental factors; for example: high wind can produce white-capping waves, shallow water can exhibit bottom effects from the seafloor, or thin cirrus clouds can alter reflectance across the spectrum (Frouin, Schwindling, and Deschamps 1996; Gao et al. 2002; Lyzenga 1981). Understanding the behavior of remote sensing reflectance is thus complicated by a multiplicity of factors that can influence spectral shape. However, algorithms have been developed to isolate the signature of suspended solids under specific conditions. These semi-empirical SSC algorithms are usually unique to a single location with characteristic water constituents and environmental conditions.

Whereas HICO acquired images to-order, Landsat 8 presents the opportunity to acquire regular, synoptic data on the San Francisco Bay. Launched in April 2013, Landsat 8 captures an image of the full San Francisco Bay estuary every 16 days. The new USGS sensor collects images covering over 31,000 square kilometers at a spatial resolution of 30 meters. Ultraviolet, visible, and infrared light reflecting off the Earth is captured in 11 bands (U.S. Geololgical Survey n.d.). These data are calibrated with sun radiance at the time of imaging to acquire remote sensing reflectance or brightness temperature (BT).

Landsat 8 is at present one of the most powerful satellite sensors in existence, but because it is relatively new, few algorithms have been developed to take advantage of its extended spectral range. Using Landsat 8 for water quality monitoring is advantageous because of this sensor's higher spatial and temporal resolution, but requires the development of new SSC

algorithms specific to the sensor's unique bands and calibrated to the San Francisco Bay. This stage of the project aims to develop a statistical algorithm to calculate SSC in the San Francisco Bay Estuary from Landsat 8 data accounting for environmental factors including water depth, wind speed, tide phase, and cloud conditions.

Methods

Images for this study were obtained from the USGS Earth Explorer database. Table 4 lists the five Landsat 8 images that were selected out of the data available as of October 2014 for Path 44 Row 34, containing the entire San Francisco Bay estuary. The selected images are largely cloud-free and taken within four days of USGS *in situ* sampling. The images were radiometrically corrected in ENVI 5. Bathymetry data were acquired as a regional DEM from the NOAA National Geophysical Data Center. Nine Landsat 8 bands[4] and the DEM were stacked, sampled to 30 meter resolution, and spatially subset to the area of interest. Pixels with elevation greater than -0.000001 meters—representing land area—were masked.

Table 4: Images used in this study

Image ID	Date
LC80440342013266LGN00	09/23/2013
LC80440342013298LGN00	10/25/2013
LC80440342014013LGN00	01/13/2014
LC80440342014109LGN00	04/19/2014
LC80440342014157LGN00	06/06/2014

[4] Panchromatic Band 8 was excluded because its wavelengths are already sampled by the visible bands. TIRS Band 11 was excluded due to corrupt data from a sensor malfunction.

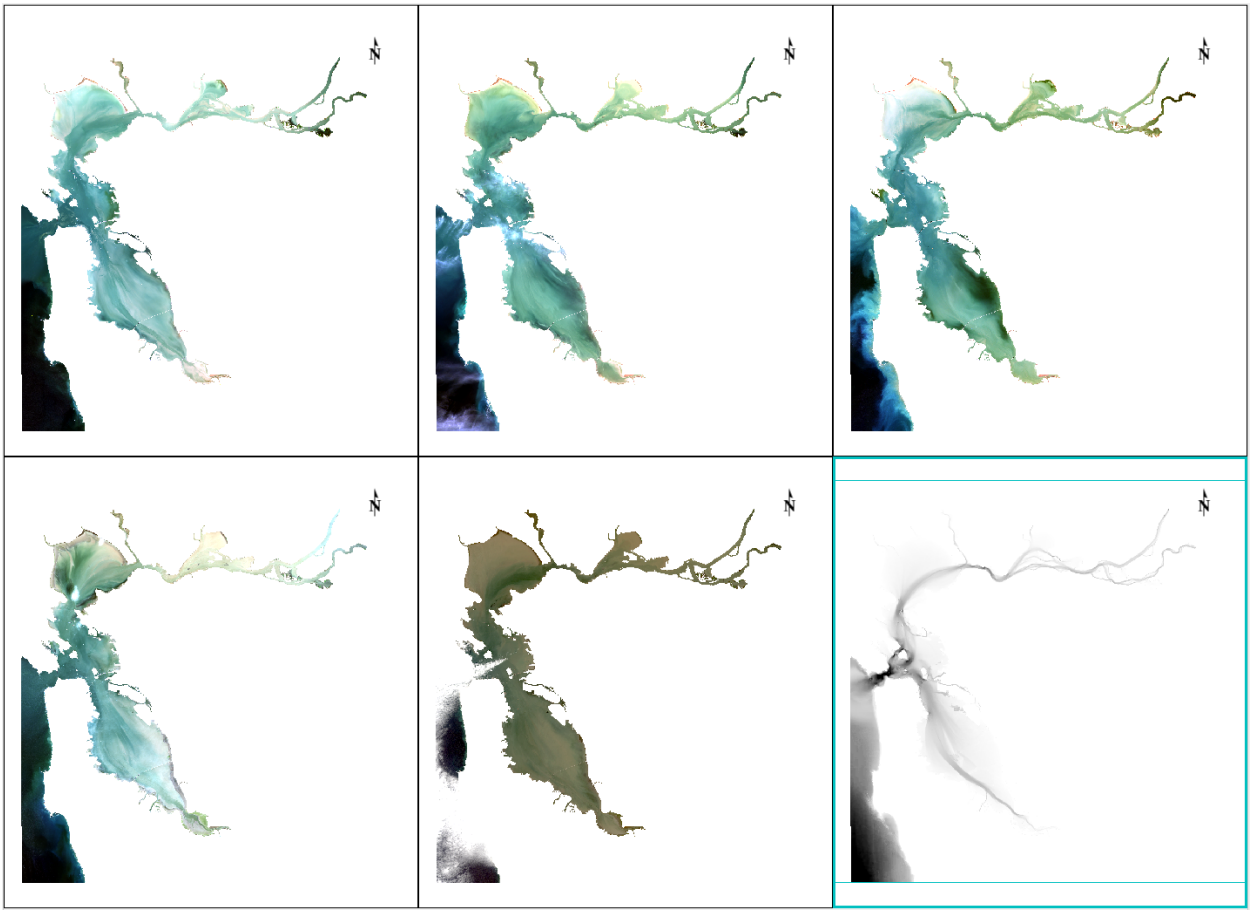


Figure 11: Landsat 8 432-RGB images, atmospherically corrected and with land pixels masked. Clockwise from top left: 09/23/2013, 10/25/2013, 01/13/2014, bathymetry, 06/06/2014, 04/19/2014.

In situ SSC data at 1 meter depth from the five corresponding dates were acquired from the USGS Water Quality of San Francisco Bay database. The dates examined span from September 2013 to June 2014, and 162 samples were collected in total. ArcGIS was used to create a vector file of the coordinates of *in situ* sampling which was overlain on the five images. Numerical data values from all bands at the pixels at those coordinates were extracted and reported as remote sensing reflectance or brightness temperature. Sampling sites were classified into seven geographic regions based on the scheme in Figure 6.

Water depths at each sampling site were extracted from the DEM acquired from NOAA. Site depths were binned into six depth categories: very shallow, less than 5 meters; shallow, 5 to 10 meters; moderate, 10 to 15 meters; deep, 15 to 25 meters; and very deep, greater than 25 meters. Because of the Bay's bathymetry, site depth is often correlated with geographic region.

Weather and tide data at each sampling point were acquired from the NOAA Center for Operational Oceanographic Products and Services from the closest of five Bay area weather stations, calculated geometrically from site coordinates. Code was written to select weather and tide data from the times nearest to imaging and sampling time. These data are available at 6-minute intervals, and times of selected data usually fall within three minutes of the corresponding image or sample data point. Variables that adjust for differences in environmental conditions between the time of imaging and sampling were calculated in Excel. Water level difference was calculated from tidal data as meters difference between water level at time of image and time of sample. Difference in time between imaging and sampling was divided into days and hours, and the magnitude of each was determined using absolute value.

Table 5: Data parameters used for model development

Bay Region	Geographic region according to Benoit, Kudela, and Flegal (2010a)
Depth Category	Very shallow, shallow, moderate, deep, or very deep
Sample Date	Sample date, MM/DD/YYYY
Sample SSC	In-situ SSC at 1m depth from USGS measurements
Day difference	Absolute value of time difference rounded to the nearest day
Hour difference	Absolute value of hours between image time of day and sample time of day
Image wind speed	Wind speed at closest weather station at time of image (m/s)
B1	Landsat 8 Band 1, Coastal Aerosol, 0.43-0.45 microns (Rrs)
B2	Landsat 8 Band 2, Blue, 0.45-0.51 microns (Rrs)
B3	Landsat 8 Band 3, Green, 0.53-0.59 microns (Rrs)
B4	Landsat 8 Band 4, Red, 0.64-0.67 microns (Rrs)
B5	Landsat 8 Band 5, NIR, 0.85-0.88 microns (Rrs)
B6	Landsat 8 Band 6, SWIR I, 1.57-1.65 microns (Rrs)
B7	Landsat 8 Band 7, SWIR II, 2.11-2.29 microns (Rrs)
B9	Landsat 8 Band 9, Cirrus, 1.36-1.38 microns (Rrs)
B10	Landsat 8 Band 10, TIRS I, 10.60-11.19 microns (BT)
Water level difference	Meters difference between water level at time of image and time of sample

Analysis of HICO data suggests that out of the visible and near infrared wavelengths, red bands have the highest predictive power for measured SSC (Figure 4). Therefore, Landsat 8 Band 4 is expected to be a significant predictor of SSC. Because the algorithm's purpose is to utilize remote sensing data, at least one band must be used as a predictor.

Because of varying environmental conditions in different areas of the bay—from shallow estuarine environments in regions A, B, F, and G to deeper water with considerable through flow in regions C, D, and E—geographic region is expected to be a significant predictor. Because reflectance of the seafloor can dominate remote sensing reflectance at shallow water depths, it is expected that water depth will affect accuracy of the SSC prediction (Lyzenga 1981).

Results

After initial scatter plots of SSC versus band values displayed heteroskedasticity, log(SSC) was used as the response variable. Figure 12 is a matrix plot with regression lines that confirms previous observations that B4 is highly correlated with SSC. Also notable is the high level of correlation between many of the bands, especially those adjacent to each other.

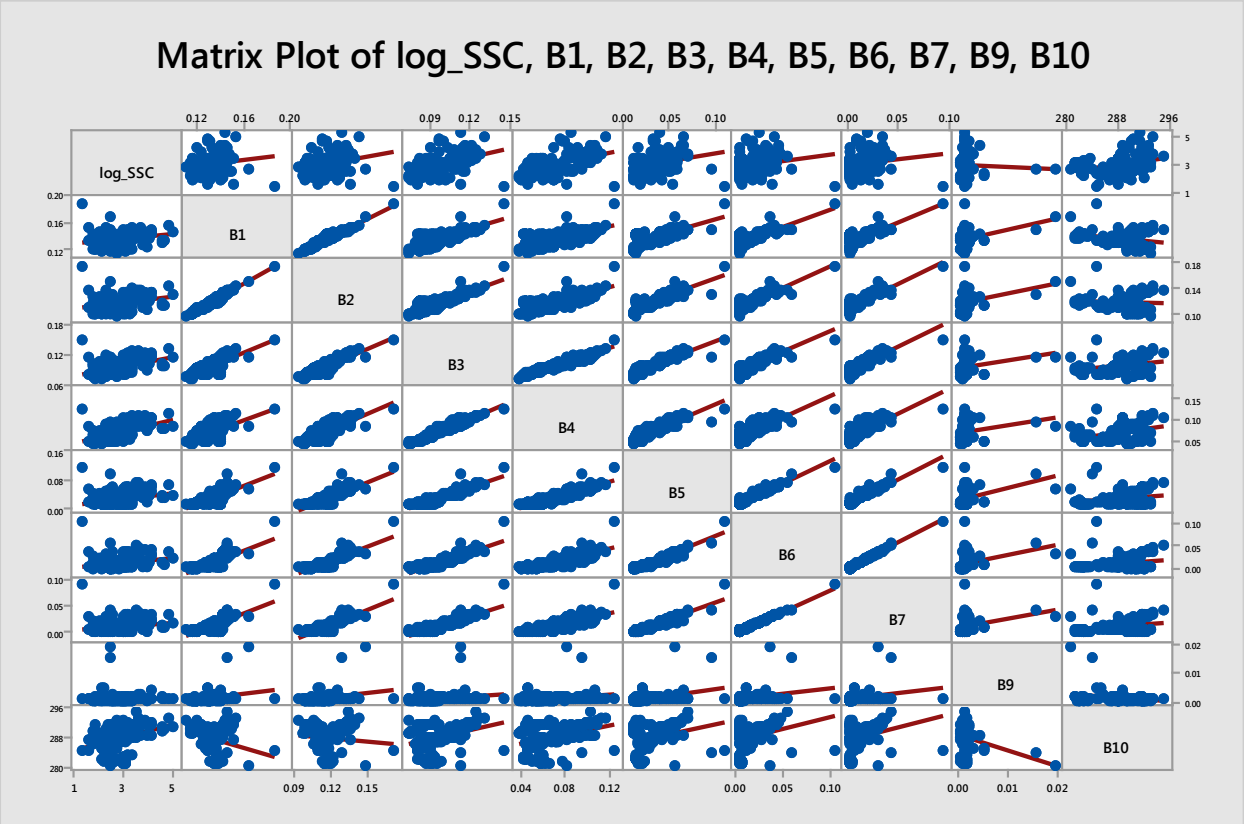


Figure 12: Correlation between log(SSC) and Landsat 8 bands shown in matrix of scatter plots with regression lines. Each cell plots the variable listed in the same column on the x-axis the variable listed in the same row on the y-axis.

Figure 13 suggests an interaction between water depth and bay region. To test if there is a significant difference in log(SSC) between sampling sites in different regions and depth

categories, one-way ANOVA was applied to a concatenated variable combining bay region and water depth category (Appendix 2).

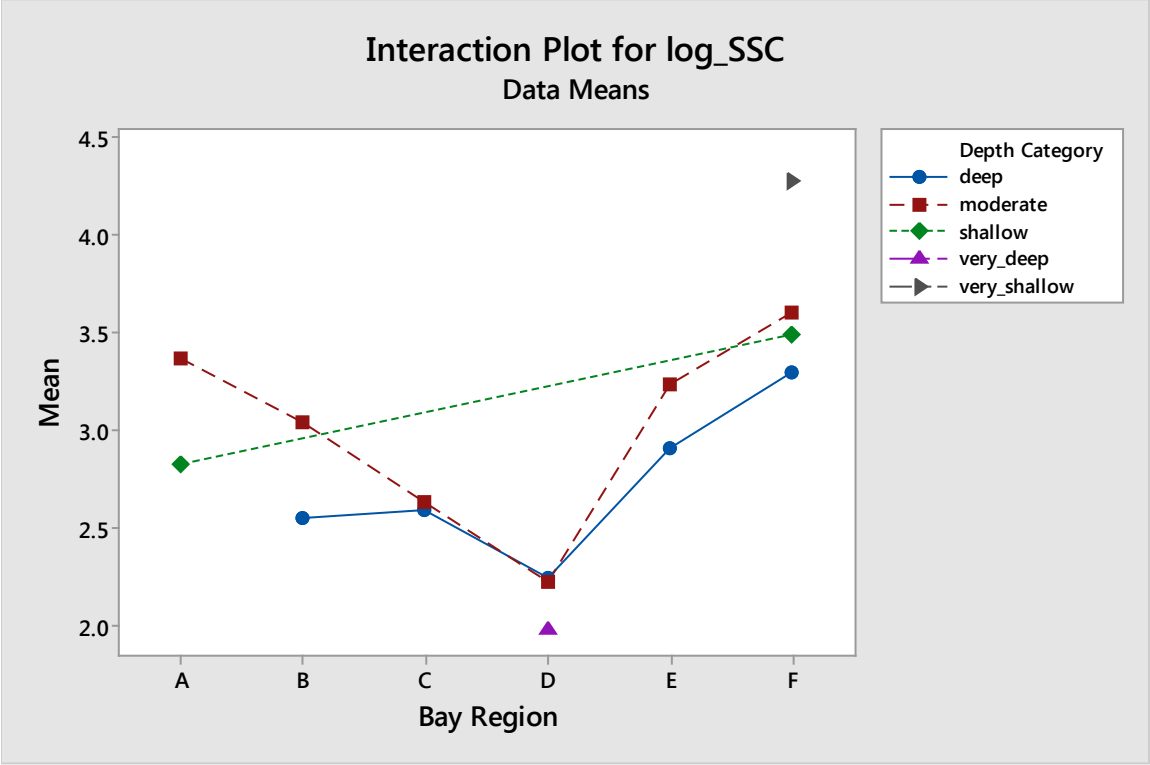


Figure 13: Interaction plot for Bay region and depth category variables. The different slope of each of the depth category lines illustrates that the effect of depth depends on the level of the Bay region variable.

The ratio of variances between categories did not exceed 3.4, which was accepted as meeting the ANOVA assumption of equal variance due to the low number of observations (as few as 2) in some categories. It was found that the combined bay region and depth categories do exhibit significant differences. The residual plots in Figure 14 also meet the assumptions of ANOVA, displaying an approximately normal distribution of residuals (bottom left); however, an order effect is visible with residuals of larger magnitude for later observations, suggesting that differences between sampling dates may also be influencing the model (bottom right). To account for this effect, a categorical variable was created for sampling date.

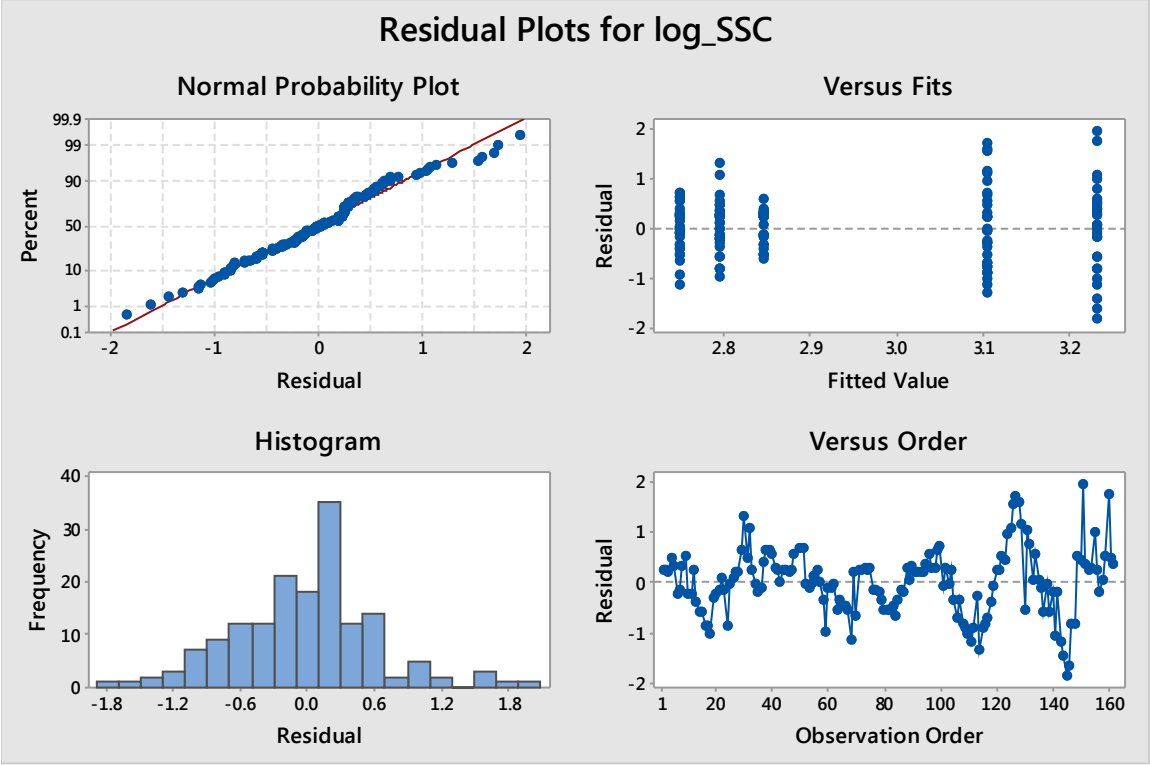


Figure 14: Residual plots for ANOVA between Bay region and depth catgory.

General linear model analysis was used to assess which variables are significant predictors of suspended solids concentration in the San Francisco Bay. A stepwise subtractive procedure was performed, beginning with bay region, depth category, and date as categorical factors and B1 through B10, day difference, hour difference, image wind speed, and water level difference as continuous covariates. A cross-effect between bay region and depth category was initially included, but could not be calculated by Minitab. At each step, the variable with the highest p-value was removed until all variables were significant to a level of α=0.05. The sequence of variables removed was as follows: image wind speed, B4, B1, B2, B3, water level difference, bay region, B9, B7. The model arrived upon includes seven significant predictors—depth category, date, B5, B6, B10, magnitude of day difference and magnitude of hour difference—and has an adjusted R^2 value of 64.87%. The standardized residual plots in Figure 15 demonstrate normally distributed residuals (bottom left), meeting general linear model assumptions. Observation 151 is an outlier with a residual greater than 4 standard deviations from the mean; it is accounted for with an indicator variable.

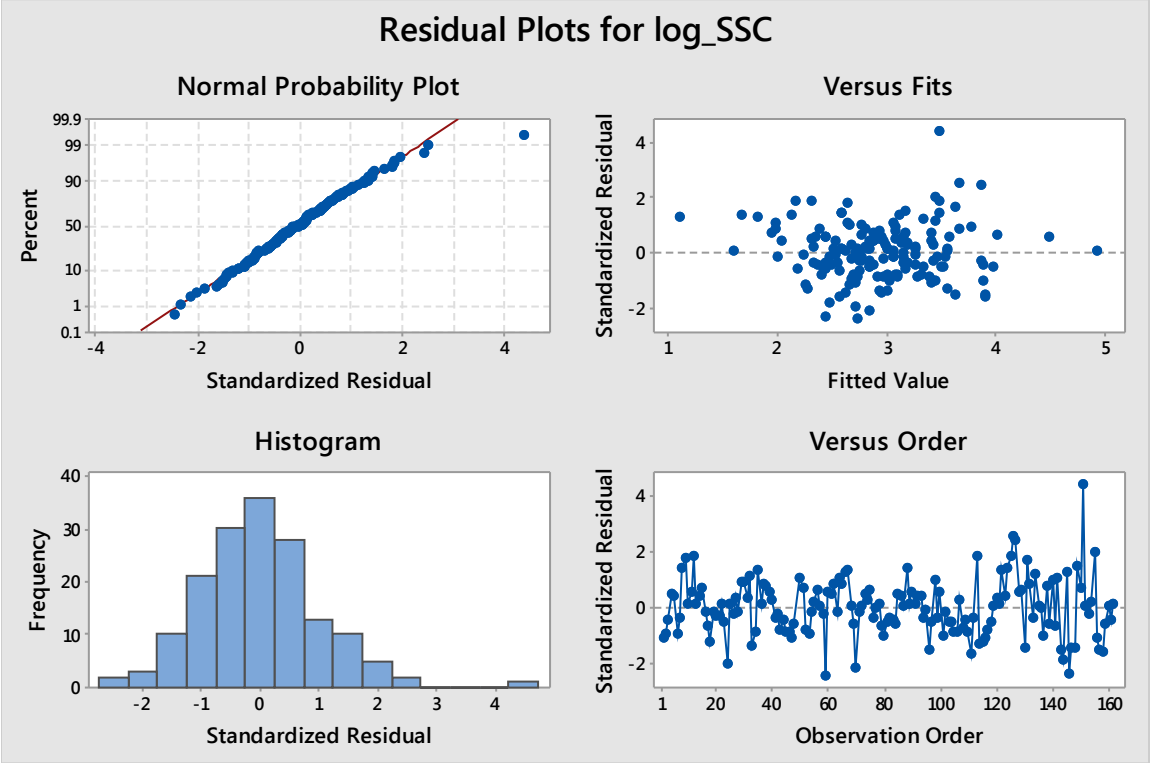


Figure 15: Standardized residual plots for general linear model.

The final model predicts log(SSC) from eight significant variables—depth category, sampling date, B5, B6, B10, magnitude of day difference, magnitude of hour difference, and an indicator variable for observation 151—and has an adjusted R^2 of 0.6933. The model coefficients are listed in Table 6 and the model is detailed in full in Appendix 3; the predicted log(SSC) is plotted against measured values in Figure 16.

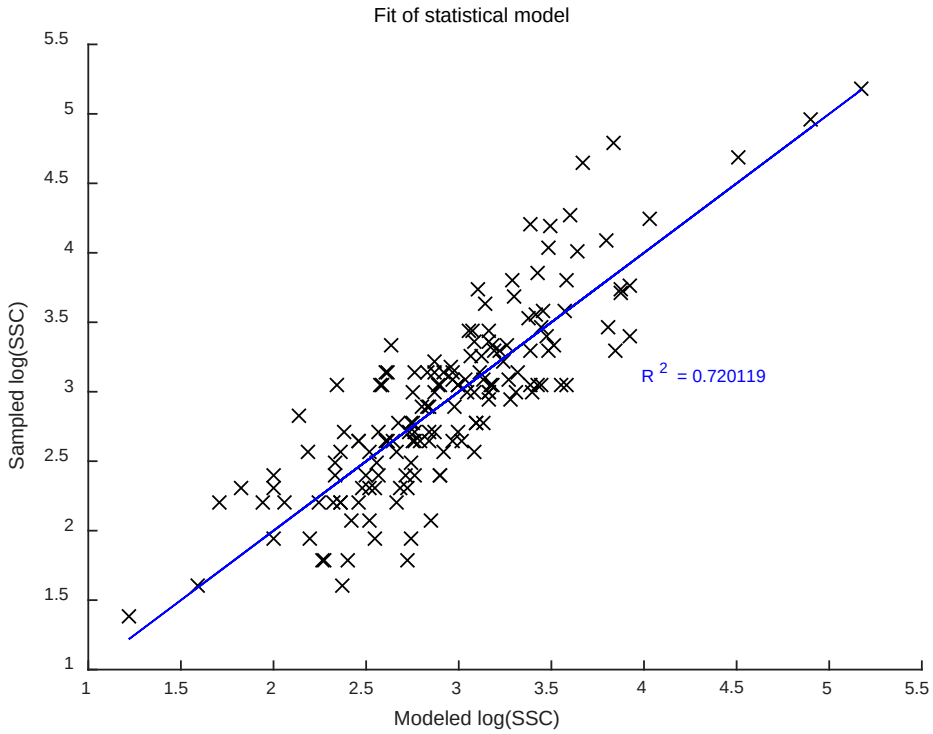
S	R-squared	R-squared (adjusted)
0.369476	72.01%	69.33%

Coefficients

Term	Coefficient	Coefficient Standard Error	T-Value	P-Value	VIF
Constant	-37.92	8.35	-4.54	0	
B5	50.3	9.16	5.49	0	33.32
B6	-54.2	11.4	-4.76	0	34.59
B10	0.133	0.0292	4.56	0	11.97
abs(day_dif)	0.624	0.127	4.92	0	43.52
abs(hour_dif)	0.132	0.0256	5.15	0	1.91
151	1.802	0.381	4.73	0	1.06
Depth Category					
deep	-0.0862	0.0721	-1.2	0.234	1.66
moderate	0.0434	0.0638	0.68	0.498	1.44
shallow	-0.413	0.123	-3.35	0.001	1.42
very deep	-0.181	0.139	-1.3	0.196	1.69
Date					
1	-0.729	0.129	-5.63	0	8.09
2	1.241	0.265	4.68	0	34.46
3	1.713	0.262	6.55	0	30.01
4	-1.091	0.206	-5.29	0	19.96

Table 6: Coefficients of general linear model terms

Figure 16: Fit of statistical model predicting SSC, detailed in Table 6.



Discussion

This model uses three Landsat 8 bands: near infrared (NIR) in B5, short wave infrared (SWIR) in B6, and thermal infrared (TIRS) in B10. Log(SSC) increases with a stronger signal from B5 and B10, and decreases with a stronger signal from B6. The exclusion of B4, expected to be highly predictive of SSC, can be explained by the high level of correlation between B4 and B5. It is also well-established that the near infrared is a good predictor of SSC (Novo, Hansom, and Curran 1989). Figure 17 shows spectra which represent characteristic spectral shapes for various ranges of SSC. The B5 NIR range is highlighted, showing R_{rs} increasing strongly with increasing SSC. Tolk, Han, & Rundquist (2000) suggest that NIR wavelengths are better than red wavelengths for measuring SSC if there is significant reflectance from the seafloor. This study found that reflectance in the 740 to 900 nanometer range is dominated by suspended solids rather than seafloor brightness, explaining why B5 has high predictive power for SSC in the predominantly shallow Bay. The physical significance of the B6 and B10 variables cannot be explained optically and are likely due to environmental conditions.

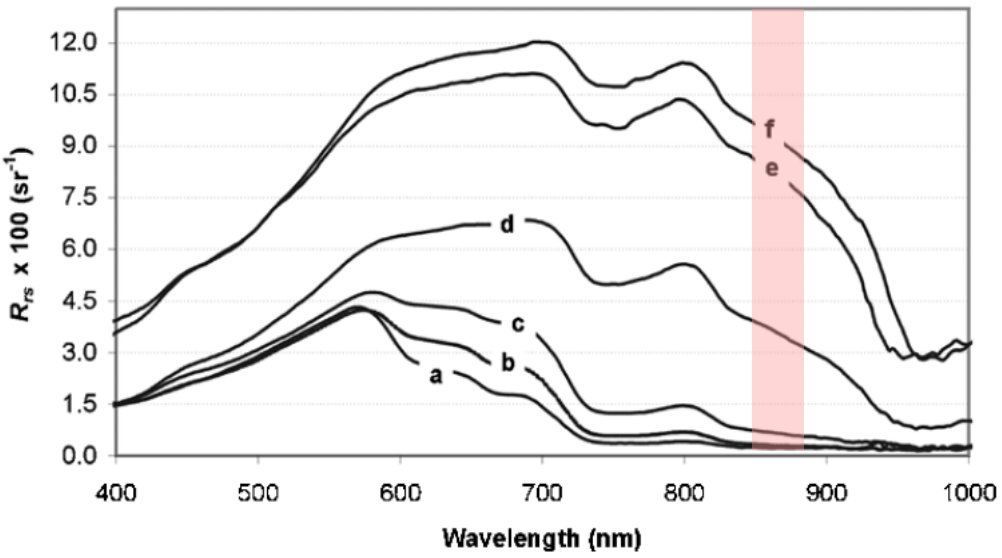


Figure 17: Examples of measured spectra at various SSC: a) 0.013 g/L; b) 0.023 g/L; c) 0.062 g/L; d) 0.355 g/L; e) 0.651 g/L; and f) 0.985 g/L. The B5 range is highlighted. Adapted from Doxaran, Froidefond, and Castaing (2002).

The effects of SSC on reflectance and brightness temperatures at wavelengths above 1100 nanometers—including the SWIR and TIRS ranges of B6 and B10—have not been significantly studied, because these wavelengths are known to be more sensitive to water absorbance and thermal emissivity, respectively (U.S. Geololgical Survey n.d.). The significance of these bands as predictors in this model cannot be explained physically, and may be due to factors correlated with the distribution of SSC in the Bay rather than SSC itself. However, since this is an empirical model designed only for local use, this is not problematic as long as environmental conditions remain similar.

Although depth is a significant variable, the shallow water category is the only significant predictor. It requires a negative correction[5], indicating that the strongest depth effects occur in shallow waters. At depths less than 10 meters, a negative correction to SSC is required to account for bottom reflectance. In water deeper than 10 meters, depth category is not significant and no correction is needed. This comports with the physical understanding of bottom effects in remote sensing (Doxaran, Cherukuru, and Lavender 2005).

[5] Although this correction is relative to very shallow water, the very shallow category has the fewest observations and therefore likely has high error. Regardless, the coefficient for shallow water is greater in magnitude than that of any other category.

Date is a significant predictor of SSC, requiring a positive correction for samples taken on some dates and a negative correction for others. This is likely due to widespread differences in SSC across the Bay between the time of *in situ* sampling and the time of remote measurement that reflect numerous environmental factors. For each date the image was captured within four days before or after *in situ* sampling, but in that time winds, tides, and currents can change significantly. While the individual effect of changes in each environmental condition cannot be quantified, the date variable encompasses their net effect. This variable may also account for systematic error in the water sampling process within each cruise date.

Both time difference variables contribute a positive correction to SSC. The direction is not meaningful; rather, the fact that as the time between imaging and sampling increases, environmental conditions change more and a correction of larger magnitude is require for SSC. Because each imaging/sampling pair had a constant day difference for all observations on that date, an attempt was made to include in the model interactions between magnitude of day difference and sampling date. Minitab was unable to estimate this variable and it was therefore not included.

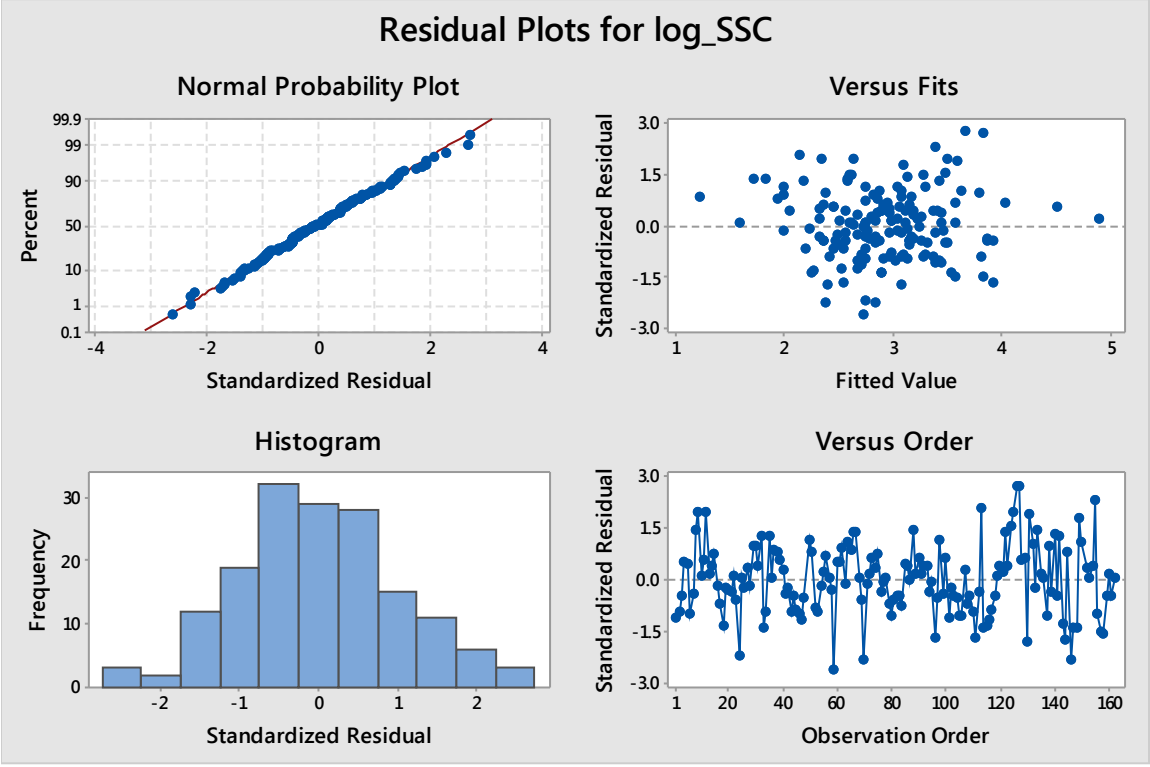


Figure 18: Standardized residual plots for final model.

Residual plots for the final model are shown in Figure 18. All residuals in the final model fall within three standard deviations of the mean (top right) and display an approximately normal distribution (bottom left), meeting assumptions for general linear models. As evident in the residual plots below, there are still trends in observation order corresponding with the order of water sampling. This suggests that hour difference did not account fully for sampling order and a correction for the order of sample collection on each date may be appropriate.

Conclusion

This section demonstrates the development of a linear statistical model for SSC based on remote optical measurement, as well as the limitations of such a model. The general linear model designed here is useful for predicting SSC throughout San Francisco Bay on the five dates sampled, but because date effects are significant in the model, SSC can be predicted for an image from a different date only if corresponding *in situ* data are available to calibrate the model. Therefore this model is more useful for spatial extension of measured SSC—where environmental conditions remain similar—than for temporal extension, where wind, tidal, and cloud conditions change. Although scientists in the field of ocean remote sensing are at present developing more advanced statistical techniques to address such dependence on training conditions, there is an alternative approach based on optics. By investigating the optical properties of water and various materials in the environment with experimental evidence and optical principles, an attempt can be made to identify within the measured remote sensing reflectance signal the contributions of water, water constituents, and environmental factors.

Section III: Empirical assessment of background effects in hyperspectral measurement of suspended solids

Introduction

In the statistical model of Section II, bottom reflectance proved the most influential environmental factor. Because it is also a largely static feature between images of the San Francisco Bay taken on different dates, understanding the optical behavior of the seafloor may enable a significant correction to measured remote sensing reflectance to better isolate the signal of water constituents. The optical effects of bottom reflectance depend on composition of the seafloor and depth of the water, as well as interaction with water constituents (Curtis D. Mobley 1994). The goal of this portion of the project is to empirically investigate the optical interaction of suspended solids and various seafloor conditions.

Visible light attenuates in water according to the Beer-Lambert law. In open water longer wavelengths are absorbed most quickly, with 90% of red light absorbed in the first five meters of water (Gordon and McCluney 1975). Blue wavelengths penetrate deeper into the water column, causing open water to appear dark blue, but coastal waters such as those in the San Francisco Bay generally display a more complex absorption pattern due to the presence of CDOM, chlorophyll, and suspended solids as well as the influence of a shallow seafloor (Tolk, Han, and Rundquist 2000). It is therefore expected that bottom reflectance may have the greatest contribution to total reflectance in the shorter visible wavelengths.

If seafloor composition is approximately uniform in the Bay, brightness contributed by seafloor reflectance should follow the trends of water depth. In shallow water the seafloor causes the most pervasive distortion of water's spectral shape, consistently adding to the reflectance from chlorophyll, CDOM, and suspended solids (Lyzenga 1981). By discerning the depth at which bottom effects become the dominant contributor to total reflectance, data in those areas can be discarded or appropriately adjusted before assessing SSC. The contribution of bottom effects to remote sensing reflectance is visible in comparing the five Landsat 8 images to the bathymetry map (lower right) in Figure 11, with deeper areas—including the ocean and the central north-south channel in the Bay—appearing darker, and shallower coastal areas appearing brighter. However, it is difficult to disentangle the influence of bottom reflectance from changes in water composition that may be related to depth through its influence on local circulation, as

evidenced by the significant cross-effects between geographic region and depth category in the general linear model developed in section II (Appendix 2). For example, shallow bathymetry in the south Bay (region F) both introduces bottom reflectance and causes the retention of high levels of suspended solids from low water circulation (Walters, Cheng, and Conomos 1985).

Reflectance patterns that do not follow trends in bathymetry, on the other hand, are likely caused by water constituents. Figure 19 shows an example of bright pixels in eddy patterns that appear to be plumes of suspended solids following local fluid motion. This pattern can be seen in different locations in multiple images and confirms that reflectance from suspended solids is not controlled exclusively by depth.

Figure 19: 432-RGB Landsat 8 image of the western inlet to the Central Bay on 9/23/2013.

Methods

Scatterplots were made in ENVI to display depth in meters versus R_{rs} in each of the visible bands for all pixels in each image, and the dancing pixel feature was used to identify areas in the images corresponding to regions in the scatterplots. Figure 20 shows scatterplots for the September 2013 image, with Band 4 (red) on the left and Band 2 (blue) on the right. Regions (a) and (d) display low reflectance values that are not correlated with depth; these regions

contain ocean pixels. Regions (b) and (e) display somewhat higher reflectance values that are also uncorrelated with depth; these regions contain pixels from the Bay's deeper central channel. Regions (c) and (f) display reflectance values positively correlated with depth; these regions contain pixels from most of the Bay excluding the central channel. Pixels outside of this region include areas near-shore ocean pixels containing sandy sediment, areas where sediment has been carried from the Bay into open water, and other anomalous areas.

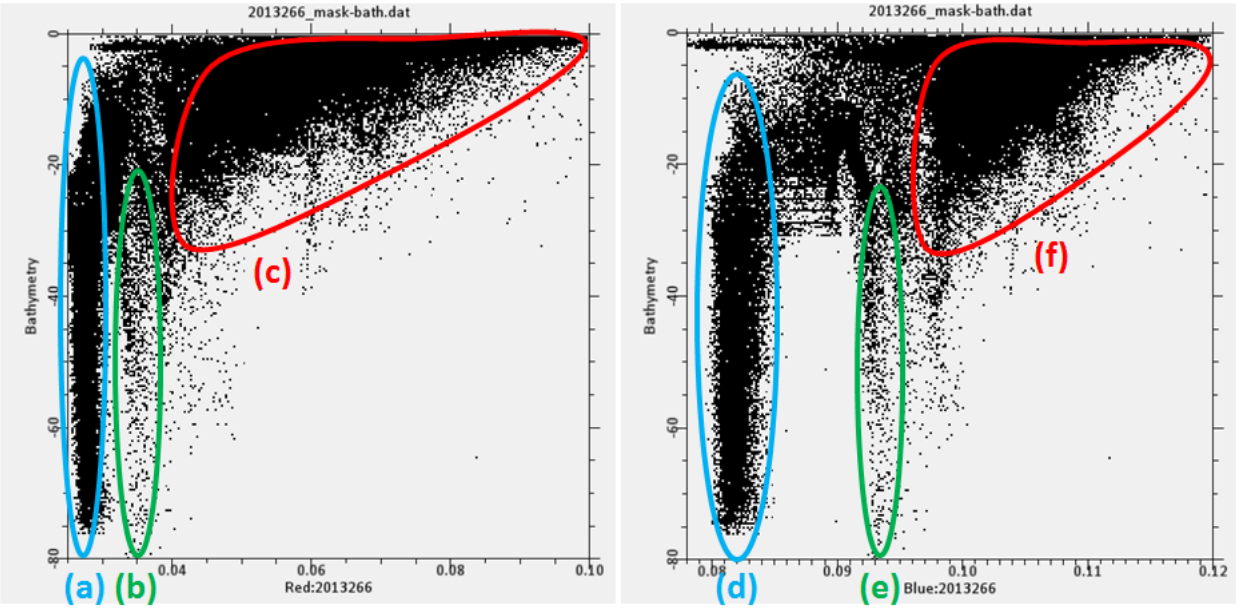


Figure 20: Scatter plots of all water pixels in the 9/23/2013 image. Left, depth versus R_{rs} in Band 4 (red); Right, depth versus R_{rs} in Band 2 (blue).

As discussed above, the open water of the ocean is darkest because it has low concentrations of CDOM, chlorophyll, and suspended solids; likewise, the central channel is darker than the rest of the Bay because it tends to contain low levels of suspended solids in accordance with bathymetric patterns. The range of reflectance values in region (c) is at least 0.06, whereas the range of reflectance values in region (f) is less than 0.03 over approximately the same depth. Although reflectance values are lower in the red band due to rapid attenuation through water, the range of reflectance values between dark and bright pixels in the Bay is greater. In the red band, pixels from the shallowest areas of the Bay have relatively higher reflectance values compared to deeper pixels than in the blue band, indicating the presence of bottom effects. A bright seafloor composed of sand and sediment is more reflective at red

wavelengths than at blue wavelengths and the effect is more pronounced at shallower depths, giving the red-band scatterplot a longer tail on the right.

These scatterplots confirm that bottom effects are significant in much of the Bay. The magnitude of the effect, however, is controlled not only by depth but also by water composition and geographic region, introducing error into the positive correlation between depth and R_{rs} and causing regions (c) and (f) to appear as triangles rather than straight lines. Therefore, it is not possible to determine the necessary correction for bottom effects from scatterplots alone.

To further investigate the effect of bottom reflectance across varying wavelengths and concentrations of suspended solids, an idealized lab-scale experiment was designed to model the interaction of reflectance from the seafloor and suspended solids. Varying levels of SSC were simulated with solutions of powdered milk in tap water, and reflectance was measured over low- and high-reflectivity backgrounds. In the low-reflectivity sequence, plywood covered in Behr Ultra Flat Black paint served as the control dark surface. In the high-reflectivity sequence, a layer of sand and gravel was added to completely cover a section of the painted plywood, forming the bright surface. A Pyrex baking dish above each surface served to hold water. Together, the plywood and dish formed the dark low-reflectivity background; the plywood, sand, and dish formed the bright high-reflectivity background. The spectra of both backgrounds are shown in Figure 21.

Figure 21: Spectra of the low- and high-reflectivity backgrounds.

The Pyrex dish was filled to a depth of 1.8 centimeters with 1 liter of tap water. Powdered milk was added incrementally and reflectance was measured with an ASD FieldSpec Pro spectroradiometer. In each sequence, 30 measurements were taken at increments of approximately 100 mg/L between concentrations of 0-3000 mg/L to model the full range SSC seen in the environment. The spectrometer measured reflectance at 1nm intervals from 350-2500nm, calibrated with a Spectralon target as the white reference. Figure 22 shows the experimental setup for both low- and high-reflectivity background sequences without suspended solids, and Figure 23 shows the experiment in progress where suspended solids obscure the high-reflectivity background.

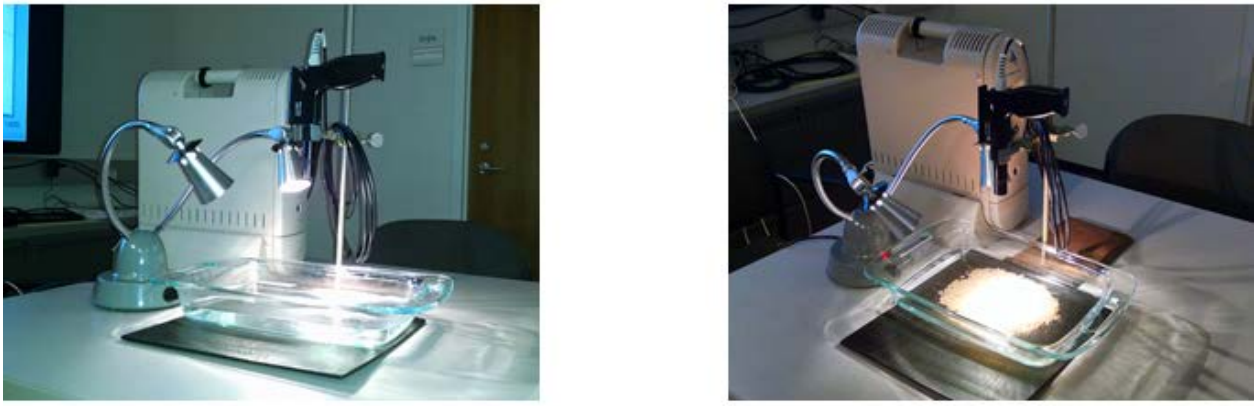


Figure 22: Experimental setup of painted plywood, Pyrex dish, and spectrometer with focused light source and pure water (0 mg/L SSC) over the low-reflectivity background (left) and high-reflectivity background (right).

Figure 23: Experiment in progress with 3000 mg/L SSC over the high-reflectivity background.

The experimental results were described with a three-layer model with total reflectance (R_t), background reflectance ($R_{b)}$, transmission of the water layer (τ), and backscattered fraction from suspended solids (s) as the wavelength-dependent parameters. As shown in Figure 24, the experimental setup is modeled as an infinitesimally thin layer of suspended solids between two water layers of equal depth ½ D. This suspension is above the background which included the painted plywood surface, Pyrex dish, and—in the bright-background case—the sand layer. Specular reflectance of the surface is ignored.

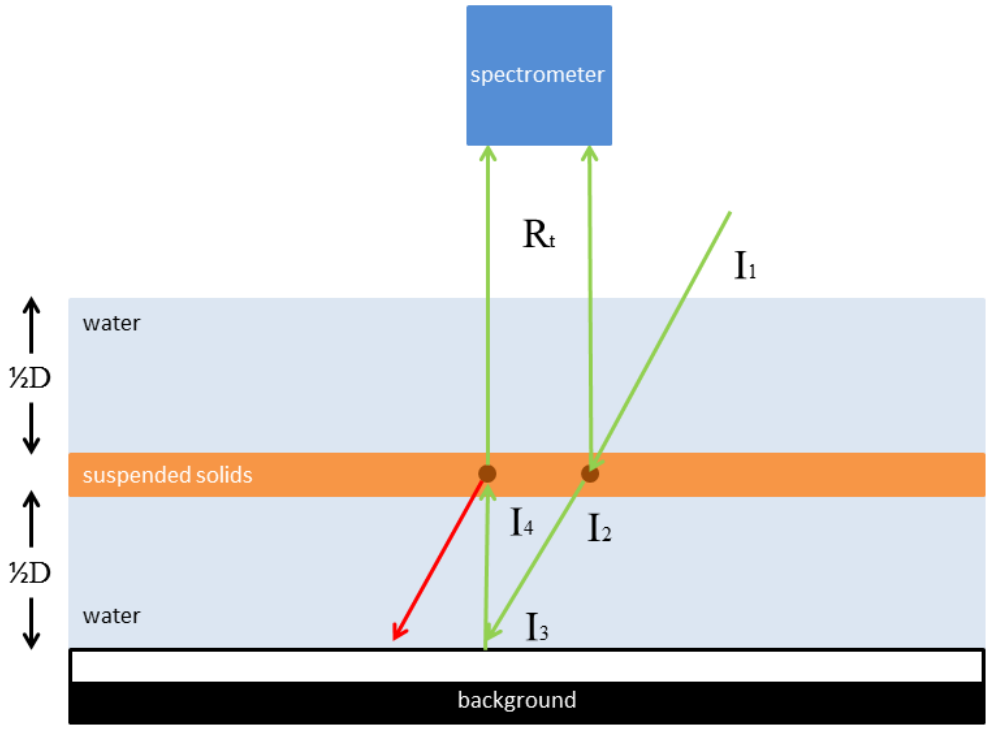


Figure 24: Schematic of the experimental setup used to develop the theoretical model.

A fraction $1-\tau$ of incoming irradiance I_1 is absorbed by the upper water layer and the remaining fraction τ is transmitted to the suspended solids layer (Equation 8). A fraction $1-s$ of the irradiance I_2 that reaches the layer of suspended solids passes through, and the second water layer again transmits a fraction τ (Equation 9). The remaining fraction s of I_2 is backscattered, passing through the upper water layer to reach the surface (Equation 11). A fraction R_b of the irradiance I_3 that reaches the background is reflected (Equation 10), and the remainder absorbed. As the light again travels the reverse path through two water layers and the layer of solids to return to the surface, the irradiance is similarly modified. Equations 8 through 11 are used to derive Equation 12 for total reflectance as measured by the spectrometer.

$$I_2 = I_1\tau \quad (8)$$

$$I_3 = I_2\tau(1-s) \quad (9)$$

$$I_4 = I_3 R_b \tau \quad (10)$$

$$R_t = I_4\tau(1-s) + I_2\tau s \quad (11)$$

$$R_t = R_b\tau^4(1-s)^2 + s\tau^2 \quad (12)$$

Once total reflectance has been measured at each wavelength, parameters R_b and τ can be determined from the boundary cases of Equation 12. Using the background alone, R_b is measured with no fluid or suspended solids where $s = 0$ and $\tau = 1$ (Equation 13).

$$R_t = R_b \quad (13)$$

Using water with no suspended solids, τ is calculated with $s = 0$ according to Equation 14.

$$R_t = R_b \tau^4 \quad (14)$$

Solving for these parameters enables the calculation of the backscattered fraction from suspended solids, the quantity most pertinent to understanding the effects of suspended solids *in situ*. The value of *s* is both wavelength and concentration dependent. Solving Equation 12 at each wavelength with the calculated R_b and τ yields two roots, and the fractional real root is designated as *s*.

Results

Background and Total Reflectance

Spectra were assessed in the visible and near infrared (NIR) range, from 400-1400nm; this range captured all significant trends. Above 1400nm, absorption by water overwhelms the reflectance of suspended solids, and below 400nm the signal is dominated by noise.

Figure 25 shows selected measured spectra of the low-reflectivity background. The background itself has a reflectance of less than 0.1 throughout the visible and NIR range. The addition of 1 L water to the dish lowers reflectance in the visible wavelengths by 0.05 and even more significantly affects the NIR range, damping reflectance to almost zero at 975nm and above 1150nm. The addition of even a small amount of solids yields a spectral shape characteristic of the milk powder solution: bright in the visible wavelengths, with NIR showing the signature of water absorption. The solution's highest reflectance is at 400nm, with an even descent to a trough at 975nm, a smaller peak at 1075nm, and consistent minimum reflectance above 1150nm. The smaller peak at 1075nm ranges from 23% to 36% of the maximum reflectance at 400nm. Figure 26 shows that as SSC increases reflectance increases across the spectrum. The slight concave shape of the curves for the visible wavelengths (400-700nm) indicate that the suspension is reaching optical saturation, where increasing concentration has a smaller marginal effect on total reflectance.

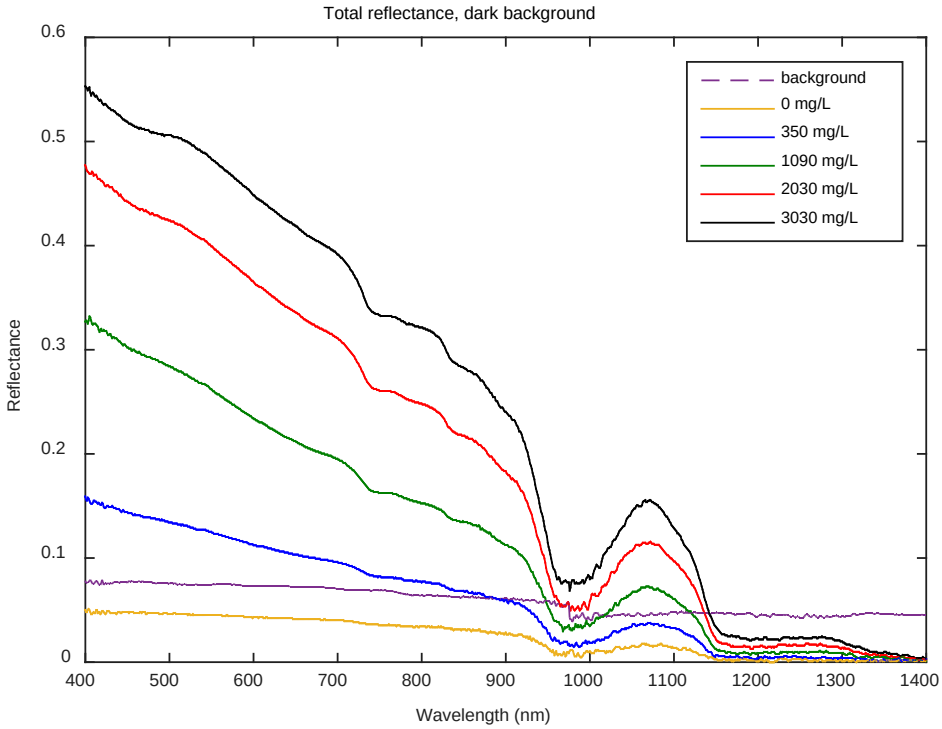


Figure 25: R_t vs. wavelength for a range of SSC values over the low-reflectivity background.

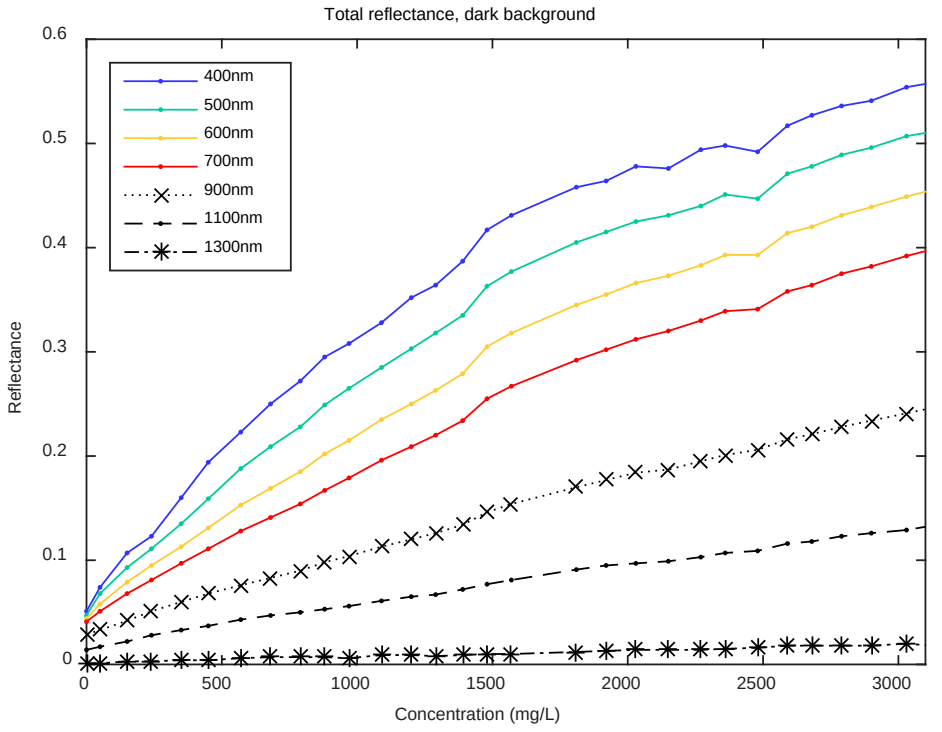


Figure 26: R_t vs. SSC for a range of wavelengths over the low-reflectivity background.

Figure 27 and Figure 28 shows measured spectra of the high-reflectivity background, made up of the black-painted shingle, a layer of sand, and a Pyrex dish. The spectrum of the background is dominated by the optical properties of sand: a minimum reflectance of 0.12 at 400nm, steadily increasing reflectance in the visible wavelengths, and then generally flat around 0.4 above 600nm. This explains the sand's orange appearance. There is a shallow trough and rise in the NIR, but reflectance remains in a range of 0.4-0.5. The addition of 1.8 cm water to the dish does not significantly affect reflectance in the visible wavelengths, but damps reflection considerably above 800nm so that the spectrum begins to exhibit a similar shape in the NIR range to the spectral shape of the low-reflectivity background. For 0 mg/L SSC, the characteristic trough at 975nm is 20% of the background reflectance, the peak at 1075nm is 68% of the background value, and the reflectance above 1150nm evens out at less than 9% of the background value. The addition of a small amount of solids raises reflectance compared to the background in the blue and green wavelengths, but lowers reflectance in the red and NIR wavelengths.

Each solution has similar reflectance at 400nm as the solution of corresponding concentration over the low-reflectivity background. However, for low-concentration solutions on the high-reflectivity background, this is not the maximum reflectance; solutions with SSC up to 1160 mg/L exhibit an increase in reflectance from 400nm to a maximum in the red wavelengths (600-700nml) and maintain elevated reflectance within 0.1 of the maximum until the trough at 975nm. This elevated reflectance at 500-900nm evidences the persistent effect of the high-reflectivity background overpowering the signal of the suspended solids. Higher concentration solutions have their maximum reflectance at 400nm, like the low-reflectivity sequence, but decrease more shallowly through the visible range. Above 700nm, the high-concentration spectra exhibit decreasing reflectance in a similar shape to spectra from the control sequence. In the peak from 975-1150nm, additional suspended solids actually diminish reflectance from the pure water curve.

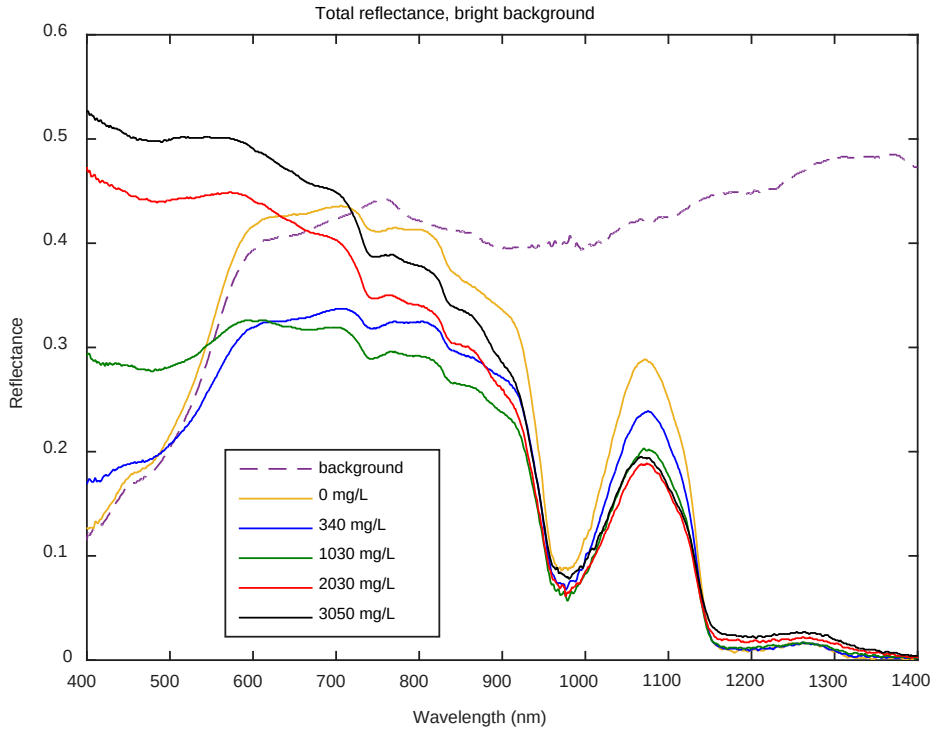


Figure 27: R_t vs. wavelength for a range of SSC values over the high-reflectivity background.

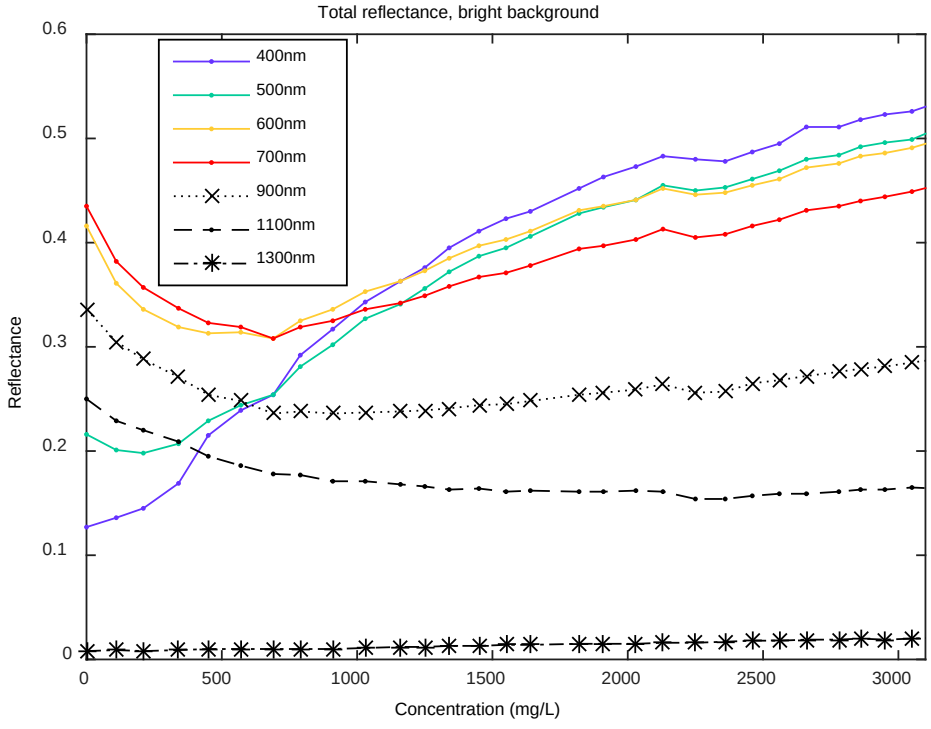


Figure 28: R_t vs. SSC for a range of wavelengths over the high-reflectivity background.

Water Absorptivity

Figure 29 shows the absorption spectrum of water—calculated as $1 - \tau$ where τ is derived from Equation 14—normalized per centimeter and compared to literature values (Palmer and Williams 1974; Pope and Fry 1997). Both measured spectra show similar spectral shapes, although the dark background curve is offset by 0.15 and contains more noise, likely as a result of precision limits of the spectrometer for measurement of low-magnitude signals. The absorption curve for the bright background closely follows established values from 400-950nm. From 950-1150nm there is a difference of less than 0.1, but above 1150nm the gap rapidly increases, indicating that measured τ is higher than established values. Therefore calculations of s in the NIR range are expected to be higher than actual values.

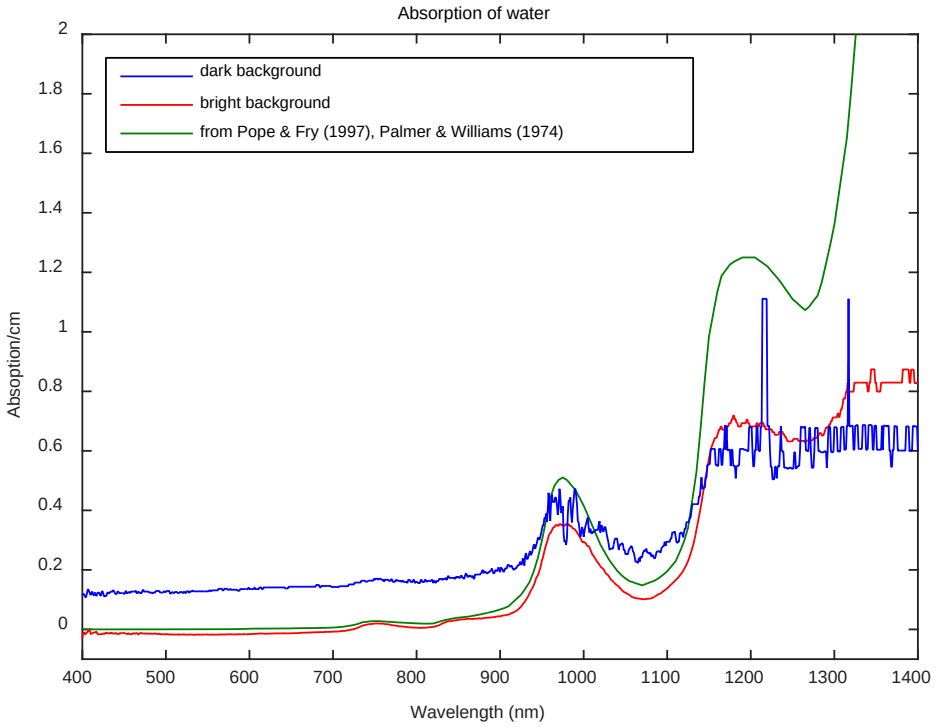


Figure 29: Calculated percent absorption per centimeter by water over the low- and high-reflectivity backgrounds compared to literature values.

Scattering by Solids

The backscattered fraction s is a property of a suspension of a particular concentration, calculated according to Equation 12 with the measured values for R_b and τ. In this model, backscatter is attributed exclusively to the thin suspended solids layer, and light not backscattered by the solids layer is transmitted. With the low-reflectivity background, this

appears a reasonable assumption. Figure 30 shows calculated *s* values from the low-reflectivity sequence compared to wavelength. Here *s* has larger values at higher concentrations and lower wavelengths, with a maximum value of 0.7 for the 3030 mg/L solution at 400nm. An approximately linear decrease in *s* is observed in all solutions—steeper for those of higher concentrations—from 400-900nm. Above 900nm, *s* exhibits the same characteristic trough and peak pattern as the total reflectance curves, with local maxima ranging from 25% to 36% of the maximum values at 400nm. Figure 31 displays the roughly linear correlation of *s* with SSC throughout the visible and NIR wavelengths, becoming shallower as wavelength increases.

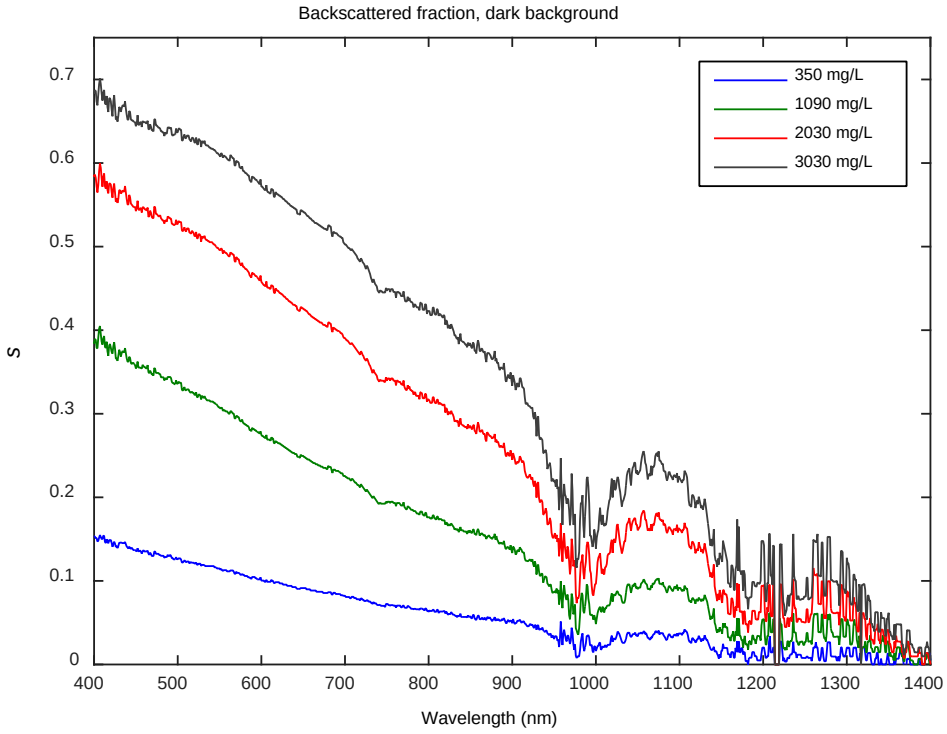


Figure 30: *s* vs. wavelength for a range of SSC values over the low-reflectivity background.

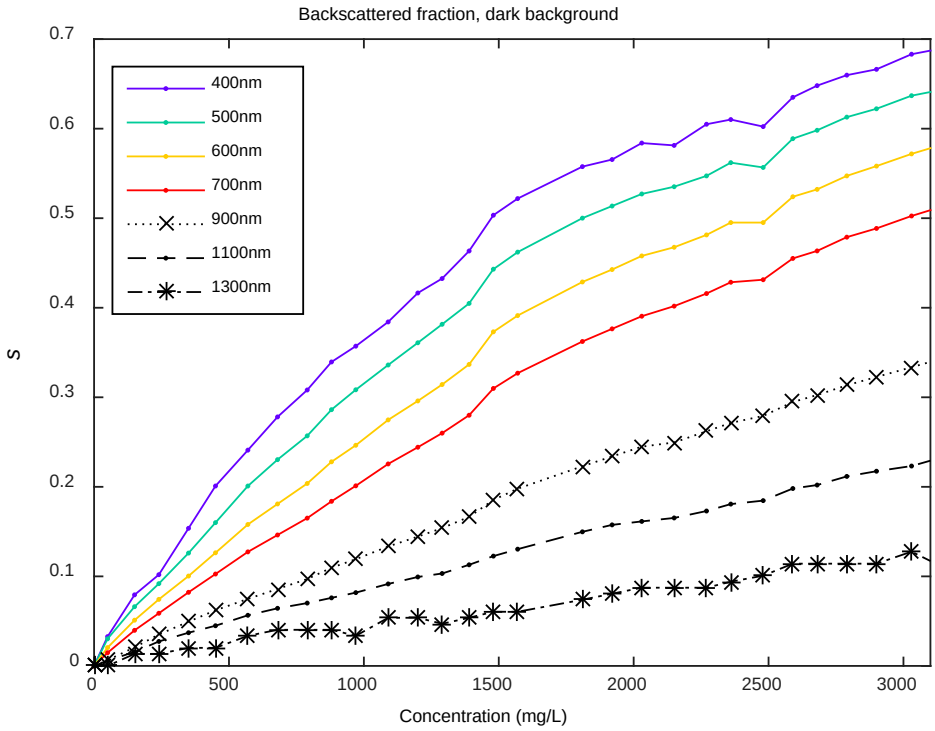


Figure 31: *s* vs. SSC for a range of wavelengths over the low-reflectivity background.

Figure 32 shows calculated *s* values compared to wavelength from the high-reflectivity sequence. Unlike in the low-reflectivity background sequence, these *s* curves do not exhibit the same trends as the total reflectance spectra: values are not exclusively positive; curves for various concentrations converge on asymptotic limits between 550-950nm and 1025-1100nm. However, some vestiges of the total reflectance curves remain. In the visible range, greater *s* values are found at higher concentrations and lower wavelengths. Curves level off above 1150nm at low positive *s* values. Figure 33 shows that only the blue, green, and high NIR wavelengths—400nm, 500nm, and 1300nm respectively—exhibit the expected linear correlation between *s* and SSC. Other wavelengths maintain a negative minimum s value over a range of concentrations.

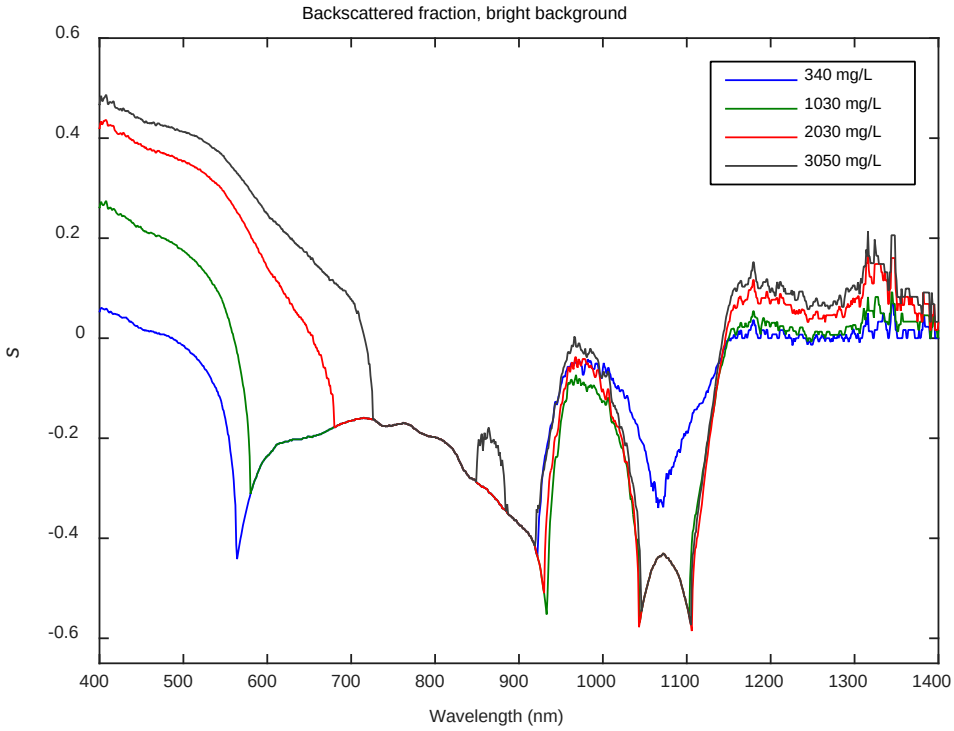


Figure 32: *s* vs. wavelength for a range of SSC values over the high-reflectivity background.

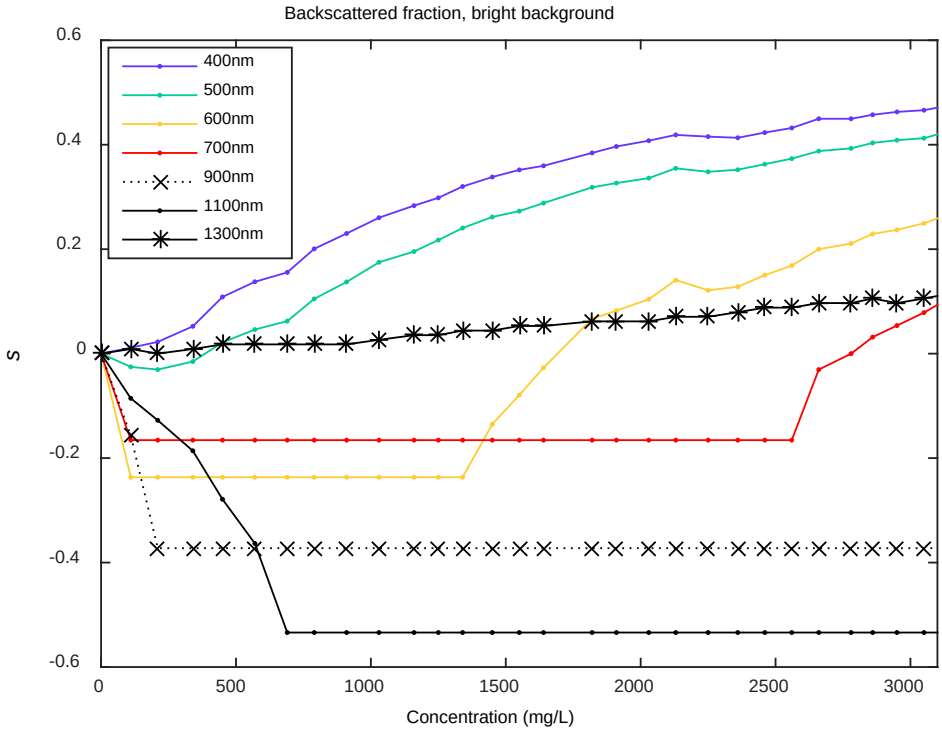


Figure 33: *s* vs. SSC for a range of wavelengths over the high-reflectivity background.

Discussion

Above 1150nm, absorption of water is so strong that reflectance of every solution is less than that of the dry background (Figure 25 and Figure 27). It is expected that increasing suspended solids concentration introduces more particles near the surface of the water, thereby increasing the backscattered fraction of light and total reflectance, but at these wavelengths suspended solids do not succeed in scattering light before it is absorbed by water. Figure 26 and Figure 28 show that reflectance at 1100nm increases slightly with increasing concentration over the low-reflectivity background and decreases over the high-reflectivity background, and reflectance at 1300nm is approximately constant. This is indicative of multiple scattering, where light is not scattered back on the incoming path but rather scattered at numerous angles to encounter other scattering particles. This increases path length through water, enabling additional absorption. Here the model's treatment of the scattering medium as an infinitely thin layer between two water layers may compensate somewhat for multiple scattering effects by imposing a path through the absorption medium regardless of concentration. The dominance of water absorption at longer wavelengths may explain the success of the model to derive *s* at 1300nm over the high-reflectivity background (Figure 33); regardless of suspension concentration, most light is absorbed before reaching the bottom of the dish.

Figure 30 displays curves for the backscattered fraction *s* that mirror trends in total reflectance: *s* increasing with SSC and generally decreasing with wavelength. Whereas the low-reflectivity sequence yields expected results for the backscattered fraction, results from the high-reflectivity background do not correspond with the sequence's total reflectance in a straightforward way. In the theoretical model *s* should represent a positive fraction, but the high-reflectivity background unexpectedly yields negative *s* values that do not have physical significance. The curves in Figure 32 exhibit asymptotic behavior increasingly as SSC decreases, suggesting that the calculation for *s* in the model does not isolate the optical signal of the suspended solids but rather captures and misinterprets the interaction of the background and the suspended solids. The model is not robust when the reflectance of the background is of a similar magnitude to that of the suspended solids, as seen with low SSC over the high-reflectivity background. However when the background reflectance is significantly lower than that of the suspended solids—such as in the low-reflectivity background sequence (Figure 31), or at higher

SSC in the high-reflectivity background sequence (Figure 33)—the problem of background-solids interaction is not as extensive, and *s* more frequently exhibits meaningful positive values.

The findings of this experiment have several implications for environmental remote sensing. The success of the model over the low-reflectivity background gives insight into the kind of environmental conditions that can be best assessed by simple single-constituent models. The requirement that reflectance from suspended solids exceed that of the background may be met where high SSC overpowers the optical signal of the seafloor: where there is a low-reflectivity seafloor material such as mud or where bottom depth exceeds the penetration depth of light (Spitzer and Dirks 1987). However, it is difficult to affirm these conditions over a broad spatial and spectral extent given the complexities of the coastal and estuarine environment, which is characterized by the presence of multiple water constituents, non-uniform seafloor composition, and variation in water depth.

With these limitations in mind, it is informative to examine a sample case of the model applied to remote sensing data. Pixels were selected from HICO image 8054 at several RMP sampling points with known *in situ* SSC values. *In situ* SSC ranged from 5 – 23 mg/L, with values of 7mg/L, 16mg/L, and 23mg/L at stations 21, 31, and 36, respectively. Water depth at these stations is 17.4m, 13.7m, and 7.9m. Figure 34 shows remote sensing reflectance spectra compared to measured reflectance of the lowest SSC suspensions over the low- and high-reflectivity backgrounds. The spectra cover the range of the 87 HICO bands, spanning the visible and low NIR wavelengths from 400 – 900nm. Because SSC in the image is an order of magnitude smaller than the lowest concentrations in the lab experiment, reflectance is displayed on a logarithmic scale.

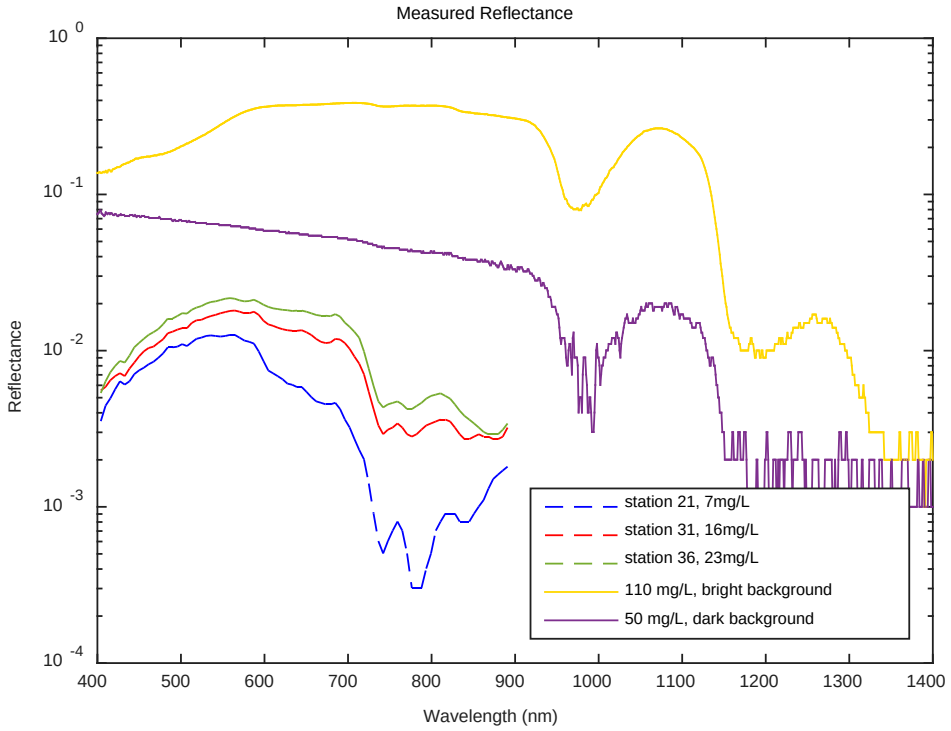


Figure 34: R_{rs} spectra for selected SSC sampling stations in HICO image 8054, compared to measured reflectance of the lowest SSC suspensions over the low- and high-reflectivity backgrounds.

Remote sensing reflectance in the image does not align completely with the spectral shape of either of the lowest concentration suspensions. In the blue and green wavelengths R_{rs} displays a similar increase to the spectrum from the high-reflectivity sequence, both displaying a minimum at 400nm that is 26% to 37% of the maximum reflectance around 600nm. Whereas the image spectra then decline shallowly from 600-700nm and reach a steep trough at 750nm, the spectrum from the high-reflectivity sequence remains level and the spectrum from the low-reflectivity sequence shallowly decreases until a trough at 950nm. The unique spectral shape of the image pixels may reflect some combination of the two background types, but also surely results from the optical influence of other water constituents and environmental factors.

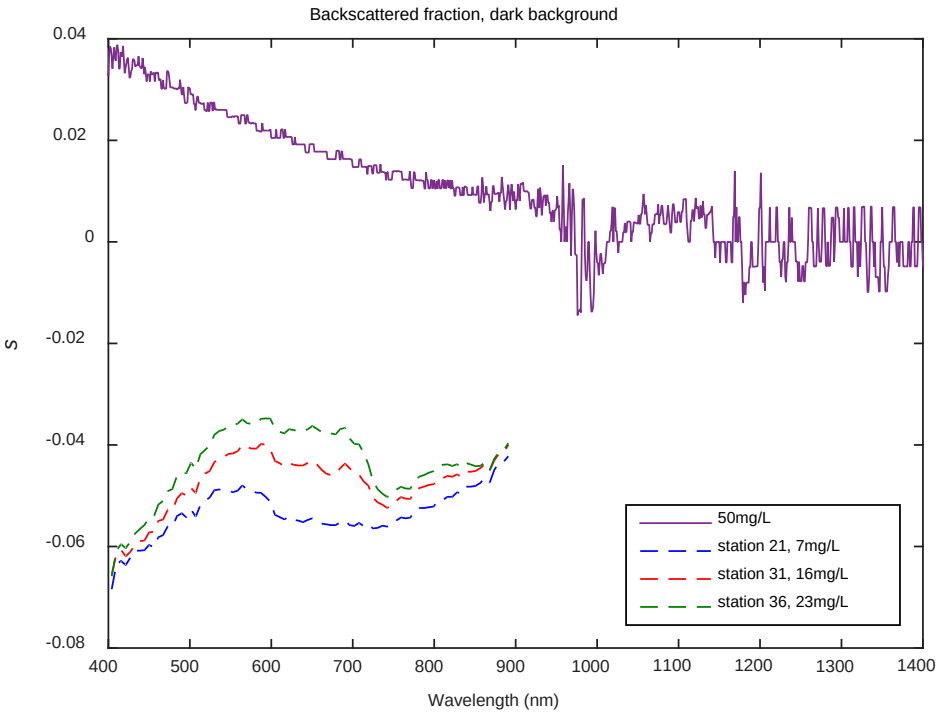


Figure 35: *s* curves for selected SSC sampling stations in HICO image 8054 calculated with parameters of the low-reflectivity background, compared to *s* for the lowest SSC suspension over the low-reflectivity background.

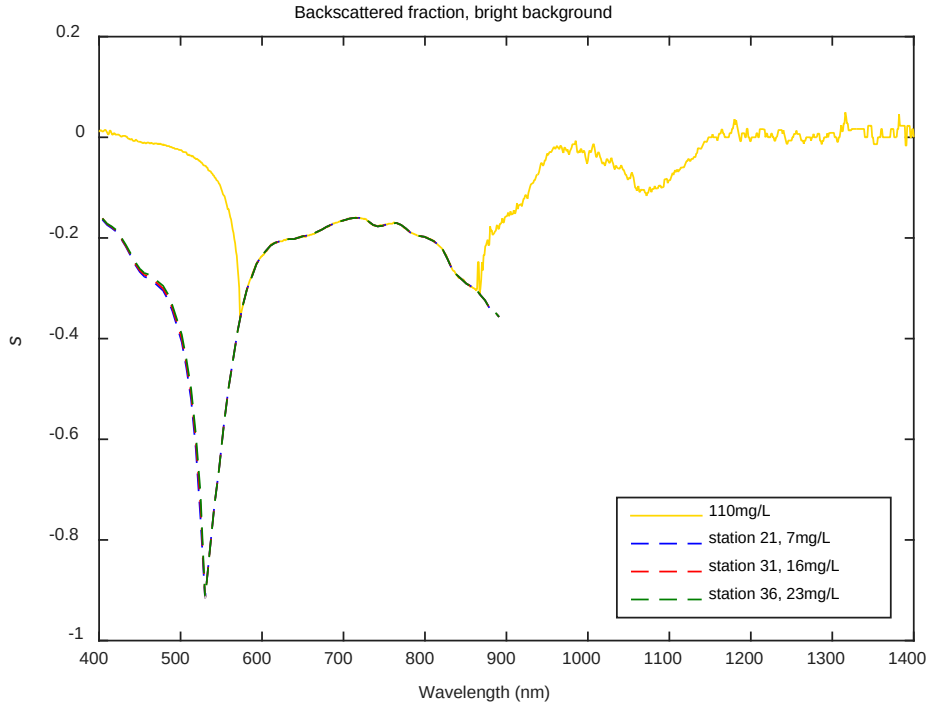


Figure 36: *s* curves for selected SSC sampling stations in HICO image 8054 calculated with parameters of the high-reflectivity background, compared to *s* for the lowest SSC suspension over the high-reflectivity background.

Figure 35 and Figure 36 display the backscattered fraction *s* for the HICO pixel spectra calculated with parameters (R_b and τ) from the low- and high-reflectivity sequences, respectively. All calculated *s* values are negative, likely because R_b measured in the lab experiment exceeded the remotely measured total reflectance of low-SSC conditions. Although the low-reflectivity *s* curves from the HICO pixels do not at all align with the shape of the *s* curves calculated from the experimental results, like the experimental low-reflectivity sequence, they mirror the shape of their total reflectance spectra (Figure 35). As expected, pixels at locations of higher SSC exhibit greater backscattering throughout the visible wavelengths. Despite the fact that negative *s* values do not have physical significance, the consistent shape between total reflectance and *s* curves validates that this model could be appropriate for assessing environments with a low-reflectivity background given the proper tuning. The high-reflectivity *s* curves extend the asymptotic shape of the low SSC experimental curves (Figure 36), confirming that the model is not robust under conditions where background reflectance greatly exceeds total reflectance.

Conclusion

These results highlight the importance of accounting for bottom reflectance, as the interaction between reflectance of suspended solids and background is dependent upon the optical properties of each. The San Francisco Bay seafloor has varied composition including sand, bedrock, and mud and the sand used in this experiment is not representative (Barnard and Kvitek 2010); likewise, suspended solids in the Bay are unlikely to resemble milk powder. Rather than represent the environment, this experiment models the type of optical behavior expected with presence of suspended solids over two generalized seafloor conditions. Although the presence of additional water constituents and environmental factors as well as a smaller magnitude of SSC prohibit direct comparison of results from the remote sensing data with those from the experimental sequences, the fact that the HICO pixels' *s* curves calculated with low-reflectivity parameters display a similar shape to their remote sensing reflectance spectra and increase in value with increasing SSC indicates that the model does capture the signal of suspended solids, even if it is unable to isolate their optical effect from that of a bright background.

Regarding the work in Section I on various algorithms for remote measurement of suspended solids in the San Francisco Bay, the optical model in this section illuminates two possible explanations for the linear relationship found between the 634nm HICO band and SSC. It is clear that bottom reflectance is not relevant for the SSC *in situ* sampling points used to develop the model. The shallowest sampling point has a depth of 7.9m, and the penetration depth—defined as the layer in which 90% of the radiance is absorbed—of red light in case II water is less than 5m (Gordon and McCluney 1975). Thus the contribution of seafloor reflectance at this wavelength is minimal in these locations and water constituents are the dominant contributors to backscatter. At 634nm the minerals composing suspended solids have approximately twice the scattering capacity as chlorophyll, suggesting that the model is indeed capturing suspended solids in deep water (Curtis D. Mobley 1994). However, the model is in this case unlikely to be accurate where seafloor depth is less than the penetration depth of water, complicating the assessment of SSC in shallower waters.

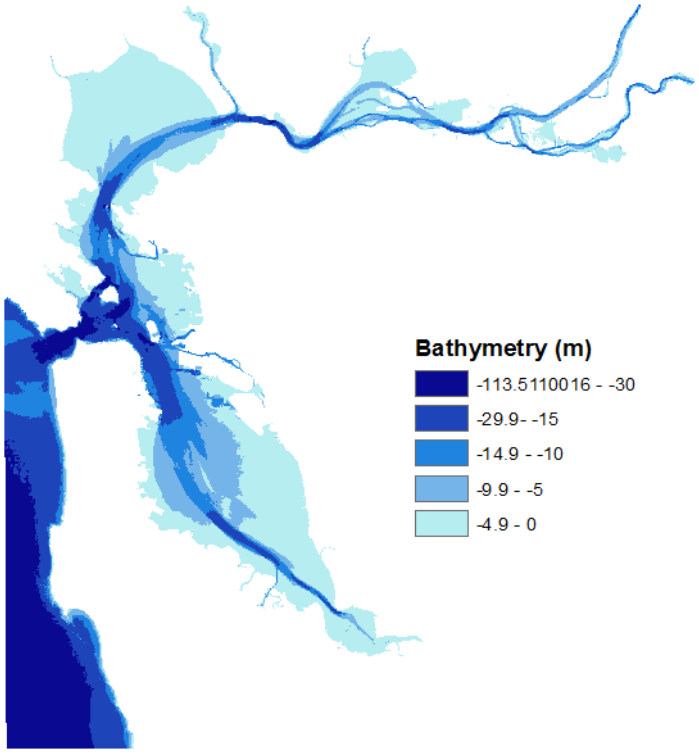


Figure 37: Bathymetry of the San Francisco Bay extracted from a regional DEM from the NOAA National Geophysical Data Center

The bathymetry in Figure 37 shows the full extent of the shallow area with the potential for bottom effects at 634nm. If the seafloor of here is composed of low-reflectivity materials such as mud, suspended solids will persist as the dominant contributor to scattering, and remote sensing reflectance or the backscattered fraction may be used as linear predictors of SSC (C. D. Mobley 2001; Spitzer and Dirks 1987). This behavior is demonstrated where R_t and *s* are plotted against SSC over a low-reflectivity background in Figure 26 and Figure 31. On the other hand, if the seafloor is composed of high-reflectivity materials such as sand or vegetation, bottom reflectance may be a significant contributor to measured R_{rs}, interacting with reflectance from suspended solids in a complicated, wavelength-dependent manner. This non-linear behavior is demonstrated where R_t and *s* are plotted against SSC over a high-reflectivity background in Figure 28and Figure 33.

Given the potential for vastly varying environmental conditions across the Bay, spatial extension of a linear model for SSC must therefore be taken with a grain of salt. This proves a significant limitation for remote monitoring as shallow bathymetry often fosters low circulation that supports retention of suspended solids, making shallow areas the most critical to monitor (Walters, Cheng, and Conomos 1985).

Future Work

This project is among the first to use data from contemporary satellite sensors to assess SSC in coastal waters. In the five years that HICO was operational, several projects were undertaken to retrieve SSC but few papers been published. For the even younger Landsat 8 there is perhaps one publication at this time. While this project succeeds in identifying potential locations of high SSC in the San Francisco Bay estuary and highlighting environmental conditions that could produce these results, it encounters the limitations of both statistical and optical methods of modeling optically complex waters.

Beyond the general linear modeling techniques applied here, principal component analysis (PCA) is often used to design models for remote sensing applications. This technique can better capture essential information from remote sensing spectra when spectral shape depends on numerous correlated factors. PCA can be used to design SSC algorithms that characterize spectral shapes of certain combinations of water constituents and environmental conditions, and may be an appropriate method of analysis for Landsat 8 and HICO data (Sathyendranath, Prieur, and Morel 1989). Partial least squares regression (PLSR) is a newer statistical method being tested for constituent retrieval in optically complex waters by Fichot et al. (2014) and Ryan & Ali (2014). PLSR combines features from PCA and multiple linear regression as well as optical information, and can perhaps be used to incorporate data from multiple sensors and across time series in algorithm development (Abdi 2010). By integrating hyperspectral data with Landsat 8 imagery and additional *in situ* measurement of optical properties, spectral profiles of the estuary waters under certain conditions may be established. The product would be a spectral library that can be searched using data from any sensor to extract turbidity conditions. Research in these methods in ongoing and much of the groundbreaking work is still unpublished.

Further development of this project and other novel work in coastal remote sensing promise many exciting applications. A more robust algorithm for SSC retrieval in the San Francisco Bay not only supports thorough and cost-effective water quality monitoring, but also enables analysis of SSC following episodic events such as storms, floods, and algal blooms whose impact on biogeochemical cycling is not well understood. Consistent monitoring is essential to assess temporal trends in SSC and increase understanding of the effects of anthropogenic activity on distribution and transport of suspended solids in the estuary. As work

in this field continues within the next few years, remote SSC monitoring may become a valuable tool not only for scientific investigation, but also for water resource management in the Bay area.

Acknowledgements

A number of people made it possible for me to complete this project, and I sincerely thank them for their support over the past year. I would like to thank the whole program staff of the NASA Student Airborne Research Program for providing me with the resources to undertake this project over the summer and present my progress at the American Geophysical Union Fall Meeting. Thank you to Emily Schaller from the National Suborbital Education and Research Center for abundant guidance about scientific research and entering the field of earth science. My advisor at the start of this research was Raphe Kudela from UC Santa Cruz. Thank you for introducing me to ocean optics, for your expert advice throughout all phases of this project, and for your support in all of my scientific endeavors. I cannot overstate the significance of Jennifer Broughton's constant presence as a mentor and a friend. Thank you for all of the work you put in on my behalf learning new methods and for your unparalleled commitment to providing feedback on my work.

I am grateful to my advisor in the Geology & Geophysics Department, Ron Smith, for enabling me to continue this project at Yale and to broaden the scope of the work over a full academic year. Thank you for overseeing the direction of my research and for providing assistance in every detail—from designing experimental methods to interpreting data. To Larry Bonneau at the Yale Center for Earth Observation, thank you for teaching me everything I know about ENVI, for providing me with amazing preparation for future work in remote sensing, and for welcoming me into your workspace with a smile. Your deep knowledge of remote sensing and positive attitude solved every problem I encountered in the lab and even convinced me that it would be worthwhile to try this research thing again sometime. I would also like to acknowledge Jonathan Reuning-Scherer for his advice on statistical modeling.

Thanks to Kush Patel for helping with code design and execution, for serving as a scientific sounding-board, and for repeating words of assurance daily. Many thanks to my friends at Yale—especially Mira Korber, Allie Schechter, Isabel Singer, Micah Sukol—for listening to me talk about the trials and triumphs of this work and for providing encouragement day in and day out. Last but not least, thank you to my parents Eric and April and my sister Elena for your editing prowess and even more for your unending love and support. Your confidence in what I can achieve and the difference I can make in the world inspires me to reach ever higher.

Works Cited

Abdi, H. 2010. “Partial Least Squares Regression and Projection on Latent Structure Regression (PLS Regression).” *Wiley Interdisciplinary Reviews: Computational Statistics*.

Barnard, Patrick, and Rikk Kvitek. 2010. “Anthropogenic Influence on Recent Bathymetric Change in West-Central San Francisco Bay.” *San Francisco Estuary and Watershed Science* 8(3).

Benoit, Michelle D, Raphael M Kudela, and A Russell Flegal. 2010a. “Modeled Trace Element Concentrations and Partitioning in the San Francisco Estuary, Based on Suspended Solids Concentration.” *Environmental science & technology* 44(15): 5956–63.

Benoit, Michelle D., Raphael M. Kudela, and A. Russell Flegal. 2010b. “Supporting Information.” *Environmental Science & Technology* 44(15).

Binding, C. E., J. H. Jerome, R. P. Bukata, and W. G. Booty. 2010. “Suspended Particulate Matter in Lake Erie Derived from MODIS Aquatic Colour Imagery.” *International Journal of Remote Sensing* 31(19): 5239–55.

Buchanan, P.A. et al. 2014. *Continuous Water-Quality and Suspended-Sediment Transport Monitoring in the San Francisco Bay, California: U.S. Geological Survey Fact Sheet*.

Domagalski, Joseph L, and Kathryn M Kuivila. 1993. “Distributions of Pesticides and Organic Contaminants between Water and Suspended Sediment, San Francisco Bay, California.” *Estuaries* 16(3): 416–26.

Doxaran, D., R. C. N. Cherukuru, and S. J. Lavender. 2005. “Use of Reflectance Band Ratios to Estimate Suspended and Dissolved Matter Concentrations in Estuarine Waters.” *International Journal of Remote Sensing* 26(8): 1763–69.

Doxaran, D., J.-M. Froidefond, and P. Castaing. 2002. “A Reflectance Band Ratio Used to Estimate Suspended Matter Concentrations in Sediment-Dominated Coastal Waters.” *International Journal of Remote Sensing* 23(23): 5079–85.

Fichot, Cedric et al. 2014. “Remote Sensing of Water Quality and Contaminants in the California Bay-Delta.” In *Proceedings of the American Geophysical Union 2014 Fall Meeting*,.

Flegal, A Russell et al. 2005. “A Review of Factors Influencing Measurements of Decadal Variations in Metal Contamination in San Francisco Bay, California.” *Ecotoxicology* 14(6): 645–60.

Flegal, A.R. et al. 1991. “Dissolved Trace Element Cycles in the San Francisco Bay Estuary.” *Marine Chemistry* 36(1-4): 329–63.

Frouin, Robert, Myriam Schwindling, and Pierre-Yves Deschamps. 1996. “Spectral Reflectance of Sea Foam in the Visible and near-Infrared: In Situ Measurements and Remote Sensing Implications.” *Journal of Geophysical Research* 101(C6): 14361.

Gao, Bo-Cai et al. 2002. “An Algorithm Using Visible and 1.38-m Channels to Retrieve Cirrus Cloud Reflectances From Aircraft and Satellite Data.” *IEEE Transactions on Geoscience and Remote Sensing* 40(8): 1659–68.

Gordon, Howard R., and W. R. McCluney. 1975. “Estimation of the Depth of Sunlight Penetration in the Sea for Remote Sensing.” *Applied optics* 14(2): 413–16.

Hammond, Douglas E et al. 1985. “Benthic Fluxes in San Francisco Bay.” In *Temporal Dynamics of an Estuary: San Francisco Bay*, Springer, 69–90.

“HICO - Sensor and Data Characteristics.” http://hico.coas.oregonstate.edu/datasets/datacharacteristics.shtml (April 27, 2015).

Lucke, Robert L et al. 2011. “Hyperspectral Imager for the Coastal Ocean: Instrument Description and First Images.” *Applied Optics* 50(11): 1501–16.

Luoma, Samuel N, and D.J.H Phillips. 1988. “Distribution, Variability, and Impacts of Trace Elements in San Francisco Bay.” *Marine Pollution Bulletin* 19(9): 413–25.

Lyzenga, David R. 1981. “Remote Sensing of Bottom Reflectance and Water Attenuation Parameters in Shallow Water Using Aircraft and Landsat Data.” *International Journal of Remote Sensing* 2(1): 71–82.

McKee, Lester J, Neil K Ganju, and David H Schoellhamer. 2006. “Estimates of Suspended Sediment Entering San Francisco Bay from the Sacramento and San Joaquin Delta, San Francisco Bay, California.” *Journal of Hydrology* 323(1): 335–52.

Miller, Richard L., and Brent a. McKee. 2004. “Using MODIS Terra 250 M Imagery to Map Concentrations of Total Suspended Matter in Coastal Waters.” *Remote Sensing of Environment* 93(1-2): 259–66.

Mobley, C. D. 2001. “Radiative Transfer in the Ocean.” *Encyclopedia of ocean sciences*: 2321–30.

Mobley, Curtis D. 1994. *Light and Water: Radiative Transfer in Natural Waters*. Academic Press.

Morel, Andre, and Louis Prieur. 1977. “Analysis of Variations in Ocean Color.” *Limnology and Oceanography* 22(4): 709–22.

National Oceanic and Atmospheric Administration. “Center for Operational Oceanographic Products and Services.” http://tidesandcurrents.noaa.gov.

Nechad, B., K.G. Ruddick, and Y. Park. 2010. “Calibration and Validation of a Generic Multisensor Algorithm for Mapping of Total Suspended Matter in Turbid Waters.” *Remote Sensing of Environment* 114(4): 854–66.

Nichols, F H, J E Cloern, S N Luoma, and D H Peterson. 1986. “The Modification of an Estuary.” *Science* 231(4738): 567–73.

Novo, E.M.M., J.D. Hansom, and P.J. Curran. 1989. “The Effect of Viewing Geometry and Wavelength on the Relationship between Reflectance and Suspended Sediment Concentration.” *International Journal of Remote Sensing* 10(8): 1357–72.

Odermatt, Daniel, Anatoly Gitelson, Vittorio Ernesto Brando, and Michael Schaepman. 2012. “Review of Constituent Retrieval in Optically Deep and Complex Waters from Satellite Imagery.” *Remote Sensing of Environment* 118: 116–26.

Palmer, Kent F., and Dudley Williams. 1974. “Optical Properties of Water in the near Infrared.” *Journal of the Optical Society of America* 64(8): 1107–10.

Pope, Robin M., and Edward S. Fry. 1997. “Absorption Spectrum (380–700 Nm) of Pure Water. II. Integrating Cavity Measurements.” *Applied Optics* 36(33): 8710–23.

Ruhl, C A, D H Schoellhamer, R P Stumpf, and C L Lindsay. 2001. “Combined Use of Remote Sensing and Continuous Monitoring to Analyse the Variability of Suspended-Sediment Concentrations in San Francisco Bay, California.” *Estuarine, Coastal and Shelf Science* 53(6): 801–12.

Ryan, K, and KA Ali. 2014. “Evaluating Existing Ocean Color Algorithms in Retrieving Chlorophyll-a Concentrations in the Turbid Waters of Long Bay South Carolina.” In *Proceedings of the American Geophysical Union 2014 Fall Meeting,*.

San Francisco Estuary Institute. “SF Bay Regional Monitoring Program.” http://www.sfei.org/rmp.

Sathyendranath, S., L. Prieur, and A. Morel. 1989. “A Three-Component Model of Ocean Colour and Its Application to Remote Sensing of Phytoplankton Pigments in Coastal Waters.” *International Journal of Remote Sensing* 10(8): 1373–94.

Schoellhamer, David H. 1996. “Factors Affecting Suspended-solids Concentrations in South San Francisco Bay, California.” *Journal of Geophysical Research: Oceans (1978–2012)* 101(C5): 12087–95.

Spitzer, D., and R. W. J. Dirks. 1987. “Bottom Influence on the Reflectance of the Sea.” *International Journal of Remote Sensing* 8(3): 279–308.

Tolk, B. L., L. Han, and D. C. Rundquist. 2000. “The Impact of Bottom Brightness on Spectral Reflectance of Suspended Sediments.” *International Journal of Remote Sensing* 21(11): 2259–68.

U.S. Environmental Protection Agency. 2014a. “Drinking Water Contaminants.”

———. 2014b. “National Recommended Water Quality Criteria.”

U.S. Geological Survey. 2014a. “Continuous Monitoring of Water Quality and Suspended-Sediment Transport in the San Francisco Bay and Delta.” http://ca.water.usgs.gov/projects/baydelta/methods.html.

———. 2014b. “San Francisco Bay Suspended Sediment Monitoring.”

U.S. Geololgical Survey. “Landsat 8 Instruments.” http://landsat.usgs.gov/landsat8.php.

———. “Water Quality of the San Francisco Bay.” http://sfbay.wr.usgs.gov/access/wqdata.

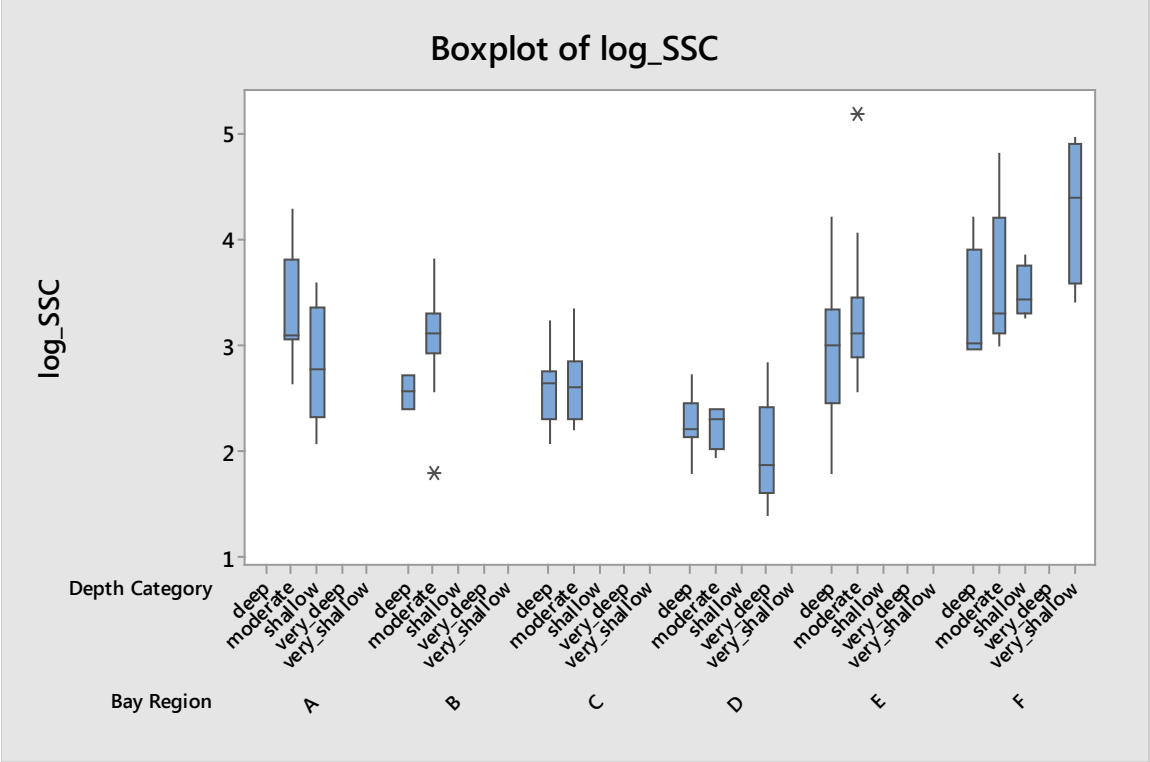
Walters, R A, R T Cheng, and T J Conomos. 1985. “Time Scales of Circulation and Mixing Processes of San Francisco Bay Waters.” *Hydrobiologia* 129: 13–36.

Appendices

Appendix 1: Calculated coefficients for trace element models

Element	WCT		WCD	
	m	*b*	*m*	*b*
Ag	-	-	-	-
As	0.40562	0.48647	0.45098	0.56984
Cd	0.51661	0.03440	0.34158	0.05617
Co	6.41650	-3.33550	9.26500	-2.30970
Cr	-	-	-	-
Cu	0.42029	0.02746	1.85720	1.13560
Fe	0.00194	1.62660	-	-
Hg	784.7100	-4.68050	-	-
Mn	0.18828	-2.85880	-	-
Ni	0.22095	0.11029	0.31271	-0.00334
Pb	3.42600	-1.73200	45.74800	-2.12330
Se	-	-	-	-
Zn	6.78320	-15.31500	0.90826	-0.87517

Appendix 2: ANOVA test of geographic region and depth category variables



Descriptive Statistics: log_SSC

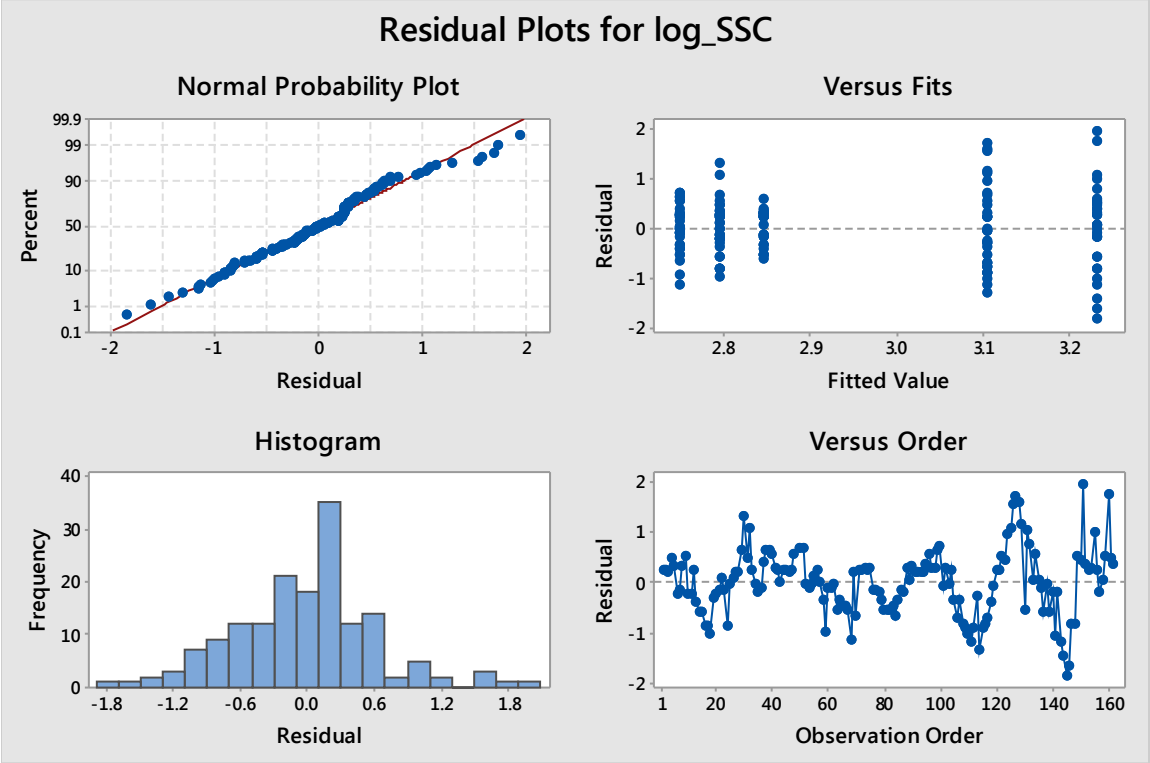
Variable	Combined	N	N*	Mean	SE Mean	StDev	Minimum	Q1	Median	Q3	Maximum
log_SSC	Amoderate	15	0	3.369	0.140	0.543	2.639	3.045	3.091	3.807	4.277
	Ashallow	5	0	2.827	0.255	0.570	2.079	2.322	2.773	3.360	3.584
	Bdeep	2	0	2.553	0.155	0.219	2.398	-	2.553	-	2.708
	Bmoderate	20	0	3.0439	0.0891	0.3984	1.7918	2.9167	3.1133	3.2864	3.8067
	Cdeep	20	0	2.5931	0.0736	0.3290	2.0794	2.3026	2.6391	2.7565	3.2189
	Cmoderate	10	0	2.635	0.111	0.352	2.197	2.303	2.602	2.841	3.332
	Ddeep	10	0	2.2395	0.0847	0.2679	1.7918	2.1344	2.1972	2.4397	2.7081
	Dmoderate	5	0	2.2247	0.0908	0.2030	1.9459	2.0127	2.3026	2.3979	2.3979
	Dvery_deep	8	0	1.982	0.174	0.492	1.386	1.609	1.869	2.413	2.833
	Edeep	21	0	2.912	0.132	0.603	1.792	2.441	2.996	3.331	4.205
	Emoderate	24	0	3.237	0.112	0.548	2.565	2.890	3.111	3.441	5.176
	Fdeep	4	0	3.294	0.299	0.599	2.944	2.957	3.020	3.903	4.190
	Fmoderate	9	0	3.604	0.225	0.674	2.996	3.107	3.296	4.203	4.796
	Fshallow	4	0	3.494	0.126	0.251	3.258	3.302	3.434	3.746	3.850
	Fvery_shallow	4	1	4.283	0.345	0.689	3.401	3.574	4.388	4.887	4.956

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Combined	N	Mean	Grouping
Fvery_shallow	4	4.283	A
Fmoderate	9	3.604	A B
Fshallow	4	3.494	A B C D E F
Amoderate	15	3.369	A B C
Fdeep	4	3.294	A B C D E F G H
Emoderate	24	3.237	B C E
Bmoderate	20	3.0439	B C D E F G H
Edeep	21	2.912	C D E F G H
Ashallow	5	2.827	B C D E F G H I
Cmoderate	10	2.635	E F G H I
Cdeep	20	2.5931	D F H I
Bdeep	2	2.553	B C D E F G H I
Ddeep	10	2.2395	I
Dmoderate	5	2.2247	G H I
Dvery_deep	8	1.982	I

Means that do not share a letter are significantly different.



Appendix 3: General Linear Model for SSC

General Linear Model: log_SSC versus B5, B6, B10, abs(day_dif), abs(hour_dif, 151,

Method

Factor coding (-1, 0, +1)
Rows unused 1

Factor Information

Factor	Type	Levels	Values
Depth Category	Fixed	5	deep, moderate, shallow, very_deep, very_shallow
Date	Fixed	5	1, 2, 3, 4, 5

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
B5	1	4.120	4.1195	30.18	0.000
B6	1	3.099	3.0995	22.70	0.000
B10	1	2.841	2.8411	20.81	0.000
abs(day_dif)	1	3.309	3.3089	24.24	0.000
abs(hour_dif)	1	3.619	3.6186	26.51	0.000
151	1	3.050	3.0503	22.34	0.000
Depth Category	4	3.579	0.8946	6.55	0.000
Date	4	7.741	1.9352	14.18	0.000
Error	146	19.931	0.1365		
Total	160	71.212			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0.369476	72.01%	69.33%	-

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	-37.92	8.35	-4.54	0.000	
B5	50.30	9.16	5.49	0.000	33.32
B6	-54.2	11.4	-4.76	0.000	34.59
B10	0.1330	0.0292	4.56	0.000	11.97
abs(day_dif)	0.624	0.127	4.92	0.000	43.52
abs(hour_dif)	0.1320	0.0256	5.15	0.000	1.91
151	1.802	0.381	4.73	0.000	1.06
Depth Category					
deep	-0.0862	0.0721	-1.20	0.234	1.66
moderate	0.0434	0.0638	0.68	0.498	1.44
shallow	-0.413	0.123	-3.35	0.001	1.42
very_deep	-0.181	0.139	-1.30	0.196	1.69
Date					
1	-0.729	0.129	-5.63	0.000	8.09
2	1.241	0.265	4.68	0.000	34.46
3	1.713	0.262	6.55	0.000	30.01
4	-1.091	0.206	-5.29	0.000	19.96

Regression Equation

Depth Category	Date	
deep	1	log_SSC = -38.74 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif) + 0.1320 abs(hour_dif) + 1.802 151
deep	2	log_SSC = -36.77 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif) + 0.1320 abs(hour_dif) + 1.802 151
deep	3	log_SSC = -36.30 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif) + 0.1320 abs(hour_dif) + 1.802 151
deep	4	log_SSC = -39.10 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif) + 0.1320 abs(hour_dif) + 1.802 151
deep	5	log_SSC = -39.15 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif) + 0.1320 abs(hour_dif) + 1.802 151
moderate	1	log_SSC = -38.61 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif) + 0.1320 abs(hour_dif) + 1.802 151
moderate	2	log_SSC = -36.64 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif) + 0.1320 abs(hour_dif) + 1.802 151
moderate	3	log_SSC = -36.17 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif) + 0.1320 abs(hour_dif) + 1.802 151
moderate	4	log_SSC = -38.97 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif) + 0.1320 abs(hour_dif) + 1.802 151
moderate	5	log_SSC = -39.02 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif) + 0.1320 abs(hour_dif) + 1.802 151
shallow	1	log_SSC = -39.07 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif) + 0.1320 abs(hour_dif) + 1.802 151

```
shallow         2     log_SSC = -37.10 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif)
                              + 0.1320 abs(hour_dif) + 1.802 151

shallow         3     log_SSC = -36.62 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif)
                              + 0.1320 abs(hour_dif) + 1.802 151

shallow         4     log_SSC = -39.43 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif)
                              + 0.1320 abs(hour_dif) + 1.802 151

shallow         5     log_SSC = -39.47 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif)
                              + 0.1320 abs(hour_dif) + 1.802 151

very_deep       1     log_SSC = -38.83 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif)
                              + 0.1320 abs(hour_dif) + 1.802 151

very_deep       2     log_SSC = -36.86 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif)
                              + 0.1320 abs(hour_dif) + 1.802 151

very_deep       3     log_SSC = -36.39 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif)
                              + 0.1320 abs(hour_dif) + 1.802 151

very_deep       4     log_SSC = -39.20 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif)
                              + 0.1320 abs(hour_dif) + 1.802 151

very_deep       5     log_SSC = -39.24 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif)
                              + 0.1320 abs(hour_dif) + 1.802 151

very_shallow    1     log_SSC = -38.02 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif)
                              + 0.1320 abs(hour_dif) + 1.802 151

very_shallow    2     log_SSC = -36.05 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif)
                              + 0.1320 abs(hour_dif) + 1.802 151

very_shallow    3     log_SSC = -35.58 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif)
                              + 0.1320 abs(hour_dif) + 1.802 151

very_shallow    4     log_SSC = -38.38 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif)
                              + 0.1320 abs(hour_dif) + 1.802 151

very_shallow    5     log_SSC = -38.42 + 50.30 B5 - 54.2 B6 + 0.1330 B10 + 0.624 abs(day_dif)
                              + 0.1320 abs(hour_dif) + 1.802 151


Fits and Diagnostics for Unusual Observations
Obs  log_SSC    Fit   Resid  Std Resid
 24    1.946  2.744  -0.798      -2.21  R
 30    4.094  3.792   0.302       0.96     X
 35    2.565  2.182   0.383       1.28     X
 59    1.792  2.727  -0.935      -2.61  R
 70    2.079  2.846  -0.766      -2.30  R
 96    3.401  3.926  -0.524      -1.67     X
113    2.833  2.137   0.696       2.07  R
126    4.644  3.668   0.977       2.73  R
127    4.796  3.835   0.961       2.69  R
128    4.682  4.511   0.171       0.55     X
130    2.565  3.082  -0.517      -1.76     X
145    1.386  1.219   0.167       0.80     X
146    1.609  2.376  -0.766      -2.28  R
151    5.176  5.176  -0.000          -     X
155    4.205  3.389   0.816       2.28  R
160    4.956  4.904   0.052       0.17     X
162    3.584  3.573   0.011       0.03     X

R  Large residual
X  Unusual X
```