Is Ocean Reflectance Acquired by Ferry Passengers Robust for Science Applications?

by

Yuyan Yang

B.Sc., South China Agricultural University, 2015

A Thesis Submitted in Partial Fulfillment of the

Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Mathematics and Statistics

Is Ocean Reflectance Acquired by Ferry Passengers Robust for Science Applications?

by

Yuyan Yang

B.Sc., South China Agricultural University, 2015

Supervisory Committee

Dr. Laura Cowen, Supervisor

(Department of Mathematics and Statistics)

Dr. Maycira Costa, Outside Member

(Department of Geography)

Supervisory Committee

Dr. Laura Cowen, Supervisor

(Department of Mathematics and Statistics)

Dr. Maycira Costa, Outside Member

(Department of Geography)

ABSTRACT

Monitoring the dynamics of the productivity of ocean water and how it affects fisheries is essential for management. It requires data on proper spatial/temporal scales, which can be provided by operational ocean colour satellites. However, accurate productivity data from ocean colour imagery is only possible with proper validation of, for instance, the atmospheric correction applied to the images. In situ water reflectance data is of great value due to the requirements for validation and it is traditionally measured with the Surface Acquisition System (SAS) solar tracker system. Recently, an application, 'HydroColor', was developed for mobile devices to acquire water reflectance data. We examine the accuracy of the water reflectance acquired by HydroColor with the help of trained and untrained citizens under different environmental conditions. We used water reflectance data acquired by SAS solar tracker and HydroColor onboard the BC ferry Queen of Oak Bay from July to September 2016. Monte Carlo permutation F-tests were used to assess whether the differences between measurements collected by SAS solar tracker and HydroColor with citizens

were significant. Results showed that the HydroColor measurements collected by 447 citizens were accurate in red, green, and blue bands, as well as red/green and red/blue ratios under different environmental conditions. Piecewise models were developed for correcting HydroColor blue/green water reflectance ratios based on the SAS solar tracker measurements. In addition, we found that training and environmental conditioins had impacts on the data quality. A trained citizen obtained higher quality HydroColor data especially under clear skies at noon run (12:50-2:30 pm).

Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

I would first like to express my sincere gratitude to my supervisors, Dr. Laura Cowen and Dr. Maycira Costa for the opportunity to do this project and their technical guidance, patience, and support during my master study.

I would like to acknowledge: BC Ferries for helping citizen data collection; Ocean Network Canada for providing data downloading system of the ocean science data; MEOPAR for providing funding for this research.

I would like to thank all members in the Spectral remote sensing lab. They are so nice and provided lots of useful technical advice. I would also like to thank the fellow graduate students in statistics for their encouragement.

Thank you to my friends for listening to me when I was upset. Thank you to my boyfriend, Weishuo Lin, for giving me strong emotional support. Thank you to my parents for their endless love throughout my life.

DEDICATION

To my grandmother,

For teaching me to be curious and optimistic all the time.

Chapter 1

Introduction

Improved understanding of the long-term spatio-temporal productivity of coastal oceans is of fundamental importance for the management of natural resources, especially fisheries (Perry and Masson, 2013). The ocean's productivity, or its proxy, concentration of chlorophyll, can be derived from water reflectance measured by operational ocean colour satellites such as MODIS-AQUA, VIIRS, and Sentinel-3 (Blondeau-Patissier et al., 2014). However, limitations on the use of satellite-derived reflectance measurements to retrieve accurate chlorophyll concentrations, specifically on the west coast of Canada, arise from the lack of sufficient in situ water reflectance data for validation of atmospherically corrected imagery (Komick et al., 2009; Carswell et al., 2017). Hyperspectral Surface Acquisition System (HyperSAS) is one of the common instruments used to measure water reflectance, and has been used extensively in the Salish Sea (Komick et al., 2009; O'Neill and Costa, 2013; Phillips and Costa, 2017). Recently, the Surface Acquisition System (SAS) solar tracker system, which tracks the position of the sun for an optimal geometry of data acquisition, has also been used on the local water (Del Bel Belluz et al., 2016). A disadvantage of these instruments is that they are both expensive and technologically complex, which

means it is difficult to acquire ocean reflectance on a large spatial-temporal scale. An alternative is to engage the public, or "citizen scientists", in ocean reflectance data acquisition to improve the spatial and temporal domain of the data.

Citizen science generally refers to the participation of non-scientists in a scientific research project (Buytaert et al., 2014). Citizen science has been widely used in different fields, predominantly terrestrial projects, which study the plants or animals on land. Volunteers on most projects are asked to help assess biodiversity and species distribution on a large scale, for example sampling terrestrial savannah invertebrates in South Africa (Lovell et al., 2009), or reporting the presence/absence of koalas in South Australia (Sequeira et al., 2014). In addition, citizen scientists have been used to acquire quantitative measurements, for example tree census measurements in forests (Butt et al., 2013), or total phosphorus, total nitrogen, chlorophyll concentrations and Secchi depth, which is related to water turbidity, in the Florida Lakewatch Program (Canfield Jr et al., 2002). Although the engagement of citizens allows us to gain data on a larger scale with less cost, the quality of citizen scientist data is a concern. Therefore, many researchers who use volunteer-collected data have first assessed the data effectiveness for the scientific framework (Fuccillo et al., 2015; Butt et al., 2013; Lovell et al., 2009), with consideration of accuracy based on providing specialized training (Lovell et al., 2009), and statistical solutions for potential biases (Bird et al., 2014).

Although citizen science is more often used in the terrestrial field, thousands of citizen scientists are involved in many marine research projects (Thiel et al., 2014). The majority of published marine citizen science studies focus on fauna data collection followed by flora and contamination data collection (Thiel et al., 2014). Due to the fact that most of the instruments for water reflectance measurement are cumbersome, there have been few citizen science projects involving data collection and

evaluation of water reflectance measurements. With increasing popularity of mobile devices, many mobile sensing applications have been developed. Todays mobile devices are sensor-equipped (e.g., inclinometer, digital compass, GPS, camera) and programmable (Lane et al., 2010). In addition, the public can easily access these applications, providing opportunities for using mobile devices as sensors to acquire water data for scientific research. Presently, many useful mobile sensing applications are available to citizens, such as (1) 'EyeonWater' (formerly 'Citclops') developed based on the Forel-Ule colour index system for distinguishing the colours of water bodies with a user-friendly interface and guidance (http://www.eyeonwater.org) (Novoa et al., 2014); (2) 'Greek watch' for the public to report the water flow rate and the amount of water trash with a photo of the waterway (Kim et al., 2011); (3)'Marine Litter Watch' for monitoring different kinds of trash in the ocean by crowdsourcing (http://www.eea.europa.eu/themes/coast sea/marine-litterwatch); (4) 'Secchi3000' water quality measurement system for measuring the Secchi depth and turbidity by Toivanen et al. (2013); (5) 'HydroColor' for monitoring water quality and providing data on turbidity (0-80 NTU), the concentration of SPM (g/m^3) and the backscattering coefficient in the red spectra (m^{-1}) calculated according to Gordon et al. (1988) and Neukermans et al. (2012) along with the water reflectance of red, green, and blue bands (Leeuw, 2014); and (6) 'Algal Watch' for evaluating occurrences of algae with both spatial and temporal scales (Kotovirta et al., 2014). To the best of our knowledge, EyeonWater, Algal Watch, and Marine Litter Watch have been used for citizen marine data acquisition in published citizen science studies (Novoa et al., 2015; Kotovirta et al., 2014; Ave, 2017). The comparison between the Forel-Ule colour values reported by citizens and derived from the images in EyeonWater shows a high correlation ($r^2 = 0.89$; Novoa et al., 2015). Algal Watch citizen data shows more dispersion than the expert data with means of correlation equal to 0.72, 0.65,

and 0.56 using bootstrapping methodology for years 2011-2013 in Finland (Kotovirta et al., 2014). It implies that the citizen data is not accurate enough, but can be used as additional information on algal monitoring. Beach marine litter data, collected by secondary school children using Marine Litter Watch after training, has been directly used for analyzing the percentage of different kinds of litter on Flora Mamaia beach, Romania (Ave, 2017). In addition, the accuracy of EyeonWater measurements has been demonstrated in controlled laboratory conditions (Mahama, 2016) and ocean satellite reflectance (Busch et al., 2016). Other than EyeonWater and Algal Watch, the accuracy of the derived information with these applications by the public is still unknown.

We assessed the feasibility of using the HydroColor application to acquire water reflectance data under different environmental conditions with the help of trained and untrained citizens onboard the BC ferry Queen of Oak Bay across the Salish Sea. This assessment was performed on the citizen HydroColor data along with the SAS solar tracker data acquired on the ferry under different time and cloud cover conditions using Monte Carlo permutation F-tests, a nonparametric statistical test to compare the means among groups. This project can potentially help with validating atmospheric correction of satellite imagery by increasing the spatial-temporal match-ups between in situ water reflectance and satellite imagery. If citizens' HydroColor data proves to be accurate, we will expand the data acquisition to fishing boats. The fishers can help with the data acquisition as "fisher scientists". Therefore a large spatio-temporal scale of accurate information can be derived for addressing fisheries management in the Salish Sea.

Chapter 2

Methods and Sampling

2.1 Study area

Sampling was conducted along the ferry route from Departure Bay, Nanaimo to Horseshoe Bay, West Vancouver in the area of the Salish Sea, west coast of Canada (Figure 2.1). The regions of the Salish Sea are the estuarine waters of the Strait of Juan de Fuca, the Strait of Georgia, and Puget Sound. The ferry route of this study was within the Strait of Georgia (SoG), which is a wide semi-enclosed sea approximately 220 km in length between the mainland of British Columbia and Vancouver Island (Masson and Peña, 2009). The water productivity of the SoG is profoundly affected by the outflow of the Fraser River, which generally peaks in June (Masson, 2006). Phytoplankton abundance in the SoG, which is associated with the Fraser River plume and tidal mixing, varies seasonally and is generally highest in the spring (Masson and Peña, 2009). Turbidity in the SoG has been found to be high at the water surface, particularly in the spring and summer (Johannessen et al., 2006). In this region, measurements of above-water reflectance at visible and near infra-red wavelengths show a large magnitude range from 0.001 to 0.027 sr^{-1} due to the variable concentrations of

water optical constituents (Phillips and Costa, 2017). Spectrally, for the SoG waters, Phillips and Costa (2017) show that the above-water reflectance (R_{rs}) is typically low in the blue wavelengths, high in the green and low again in the red wavelengths. This trend can be compared with the trend of the citizen R_{rs} to see whether the citizen R_{rs} can show the feature of the water of the SoG.

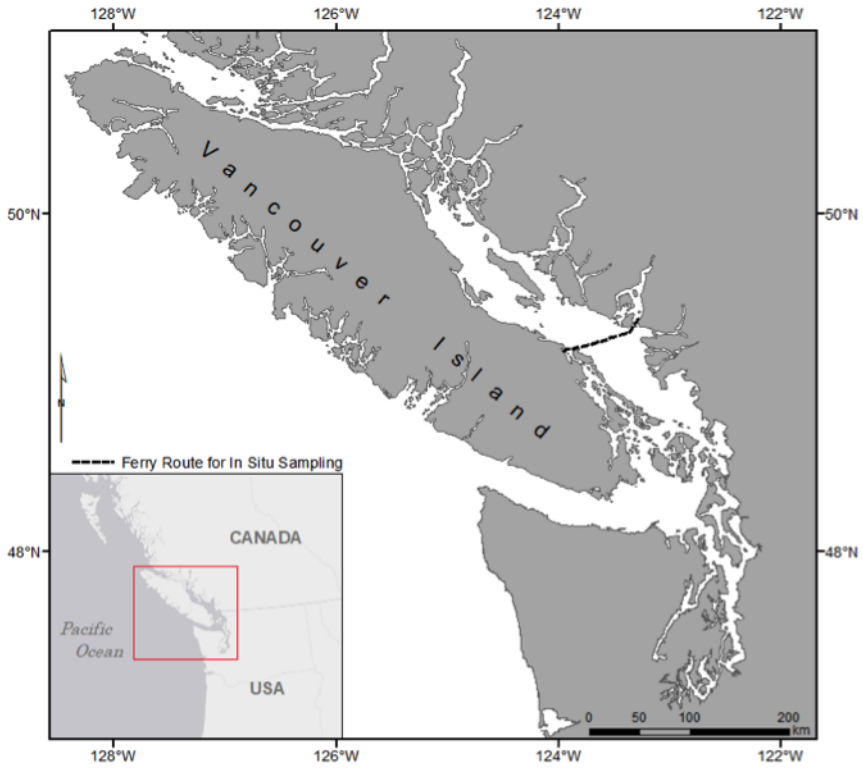

Figure 2.1: Map showing the study ferry route from Departure Bay, Nanaimo to Horseshoe Bay, West Vancouver, in the Salish Sea.

2.2 Data set

Above-water reflectance, turbidity, and chlorophyll-α measurements were acquired along the Queen of Oak Bay ferry on the SoG between July 1 and September 5, 2016.

Above-water reflectance data was collected with both high precision measurements and ferry passengers to assess whether the citizen data was accurate. We also collected turbidity and chlorophyll-α to evaluate the impact of these measurements on the error of the citizen data.

2.2.1 Above-water reflectance measurements

Above-water reflectance measurements were acquired with the HydroColor application and SAS solar tracker, an autonomous high precision data acquisition system.

HydroColor - citizens

Citizen water reflectance data was collected by the water quality application, HydroColor, which is available on both IOS and Android mobile devices (Leeuw, 2014). This application uses the camera to detect the light intensity coming from natural water bodies and skies, and then calculates the red, $R_{rs,H}(R)$, green, $R_{rs,H}(G)$, and blue reflectance, $R_{rs,H}(B)$, respectively. The reflectance of measured water is determined by the number of optical water constituents (OWCs) and how they absorb and scatter light in the water. The main components of OWCs are coloured dissolved organic matter (CDOM), inorganic suspended particulate material (SPM), and phytoplankton chlorophyll-α (Mobley, 1999).

In addition to measuring light, HydroColor leverages the GPS, compass, and inclinometer in mobile devices to ensure users obtain precise measurements with a required geometry. The geometric setting is a solar azimuth angle of 135°, which is a horizontal angle measured clockwise from the sun, and a solar zenith angle of 40°, which is the angle between the zenith and the centre of the sun's direction (Leeuw, 2014). For measuring water reflectance, the user needs to take a photo of an 18%

exposure photography grey card, a photo of the sky, and a photo of the water in the correct direction denoted by green arrows (illustrated in Figure 2.2). The photos of the grey card, sky, and water measure downward plane irradiance, sky radiance, and water leaving radiance respectively (Leeuw, 2014). After capturing photos, the above-water reflectance, $R_{rs,H}$, and other water properties are instantly computed. All photos and measurements are saved in the application and can be downloaded to computers for later analysis.

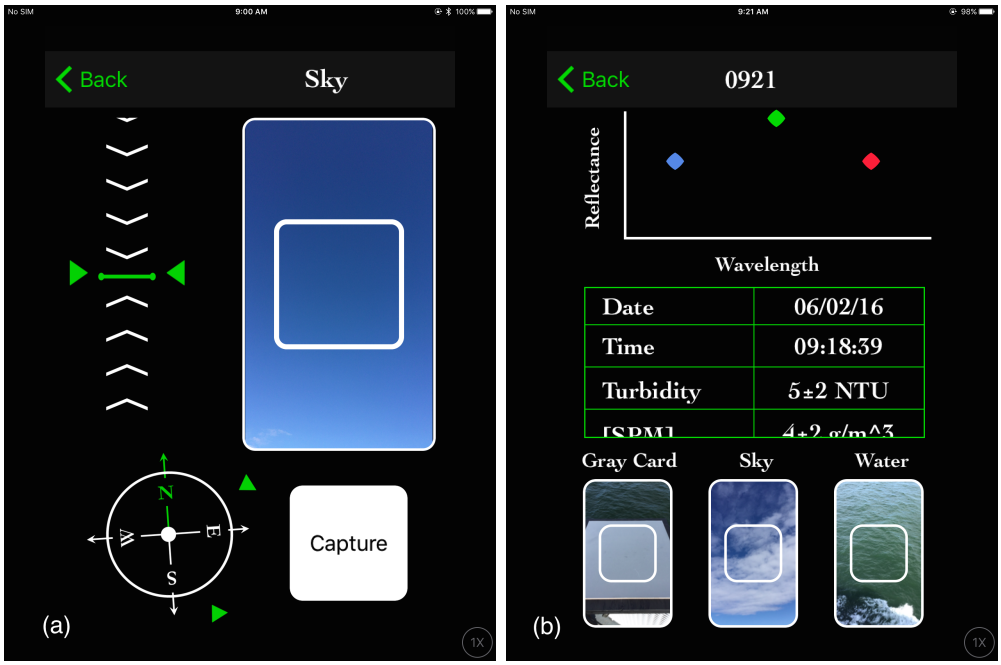

Figure 2.2: HydroColor interfaces of (a) data collection, (b) analysis. The correct direction is set when the lower green arrow rotates to align with the north arrow on the compass and the upper green arrows are lined up with the short green line on the inclinometer in (a).

HydroColor measurements were acquired every Friday to Monday during two ferry routes morning (8:30-10:10 am) with a solar elevation range from 22.4° to 49.5° and noon (12:50-2:30 pm) with a solar elevation range from 50.7° to 72.6°. Three iPad mini 4 tablets with the HydroColor application were used to acquire data, and two 18% photography grey cards were installed on the railing of the ferry deck for photo

taking (Figure 2.3).

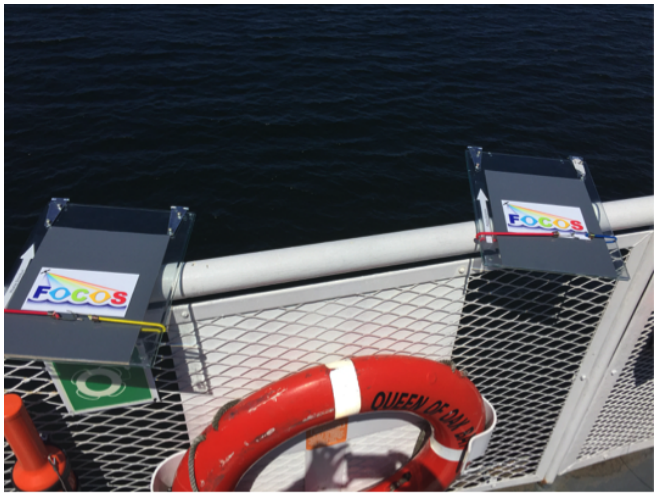

Figure 2.3: The setting of 18% photography grey cards attached to the ferry railing.

As part of the BC Ferries and Parks Canada Coastal Naturalist Program, a biologist, called a Coastal Naturalist (CN) (the same person for the entire sampling period), was trained to use the HydroColor application (trained citizen) before working with ferry passengers (FP, untrained citizens). Thiel et al. (2014) found that over 40% of reviewed marine citizen science projects provide necessary training and around 28% of studies have experts accompany citizen scientists for the research activities. Also, HydroColor samples must follow specific angles as mentioned above, and to do so requires proper training. Therefore, both trained and untrained citizen data were acquired for analyzing the training effect on the HydroColor data quality. Moreover, only adults (age 18 or older) were invited to collect data for quality control. The established protocol was that the CN gave a 25-minute presentation prior to inviting FP to collect data, so a minimum level of information was provided to FP (Figure 2.4). A total of 446 FP were involved in data collection resulting in 1270 HydroColor samples acquired by CN and FP. The University of Victoria's Human Research Ethics Board reviewed and approved the study (Ethics Approval Number: 16-221).

Figure 2.4: Costal Naturalist doing presentation and helping passengers collect data.

SAS Solar Tracker - autonomous

The SAS solar tracker was installed aboard the ferry to obtain precise water reflectance measurements for comparing with HydroColor measurements. The system was mounted in the front of the upper deck of the vessel to avoid capturing the vessel or its shadow in the measurements (Costa et al., 2016). This instrument has three sensors to simultaneously acquire high precision hyperspectral measurements of water-surface radiance, $L_t(\lambda)$, sky radiance, $L_s(\lambda)$, and sky irradiance, $E_s(\lambda)$, where λ is wavelength from 350- 800nm. The SAS solar tracker is able to adjust its sensors' pointing angles with respect to the vessel's heading angle and the local sun azimuth angle, which is calculated by the GPS location and time, and therefore automatically adjusts to the correct geometry setting according to Hooker and Morel (2003). Measurements were acquired continuously when the ferry was travelling and the solar elevation angle was greater than 30° (Gordon, 2005). From these measurements, the corresponding above-water remote sensing reflectance, $R_{rs,SAS}$ was calculated (see section 2.3.2).

2.2.2 Turbidity and chlorophyll-α

Surface water turbidity and concentration of chlorophyll-α data were downloaded from the Ocean Network Canada website (http://dmas.uvic.ca/DataSearch). This data was acquired with a WETLabs ECO Triplet installed on the BC ferry Queen of Oak Bay. The downloaded turbidity and chlorophyll- data were first averaged in one-minute intervals and then matched with the observed water reflectance dataset by time tags for analyses.

2.3 Data processing

After acquiring the raw data acquired by SAS solar tracker and HydroColor, all measurements were processed to remove unqualified measurements and calculated R_{rs} into a comparable format between two instruments prior to analyzing the error of $R_{rs,H}$.

2.3.1 $R_{rs,H}$

All HydroColor data and images acquired by citizens (both the CN and FP) were first sorted by (1) cloud cover, and (2) route time. The level of cloud cover was defined as clear (no cloud in the sky), overcast (clouds covering the sky), and cloudy (patchy clouds in the sky). Cloud cover conditions affect the quality of the R_{rs} data because the reflectance of clouds are brighter than the background sky and therefore produce cloud glitter effects (Mobley, 1999). In addition, there were slight time differences between taking three photos with HydroColor, while the SAS solar tracker acquired $L_t(\lambda)$, $L_s(\lambda)$, and $E_s(\lambda)$ simultaneously. Therefore the $L_t(\lambda)$ and $L_s(\lambda)$ derived from the water and sky photos in HydroColor might be different from those measurements by SAS solar tracker, especially when having passing clouds in the sky. The level of

route time was defined as morning (8:30-10:10 am) and noon (12:50-2:30 pm). The time effect was considered because as the solar zenith angle becomes larger, there is less radiance that is backscattered from the ocean water (Wang, 2006); so fewer signals can be detected in the morning. Therefore there were six levels of time and cloud cover as environmental conditions: clear sky morning, cloudy morning, overcast morning, clear sky noon, cloudy noon, and overcast noon.

Subsequently, HydroColor data were classified into three categories according to photo quality in the white square (Figure 2.5): perfect, good, and bad. A sample was classified as perfect if all three photos (water, sky, and grey card) had no shadows or other contaminants (Figure 2.5a); as good if only one of them was less than 1/3 contaminated (Figure 2.5b); and as bad if more than one photos was contaminated, or any of the water reflectance values were zero (Figure 2.5c). Typical reasons for bad photos were (1) shaded grey card or water, (2) white foam or sun glint in the water photo, and/or (3) the photo is entirely different from the object; for example, a sky image was captured for the photo titled 'water'. Conditions (1) and (2) are generally considered as main sources of contamination for R_{rs} measurements (Garaba et al., 2012; Leeuw, 2014; Zibordi et al., 2015). Only perfect and good samples were considered in the statistical analyses.

2.3.2 $R_{rs,SAS}$

The raw SAS solar tracker data, $L_t(\lambda)$, $L_s(\lambda)$, and $E_s(\lambda)$, was processed by an automated raw data processing software PySciDon, which included data calibration, dark data correction, time interpolation, and wavelength interpolation (Vandenberg et al., 2017). All raw SAS solar tracker data was processed in a 90-second interval and the above-water remote sensing reflectance, $R_{rs,SAS}(\lambda)$, was calculated according to Ruddick et al. (2006) using:

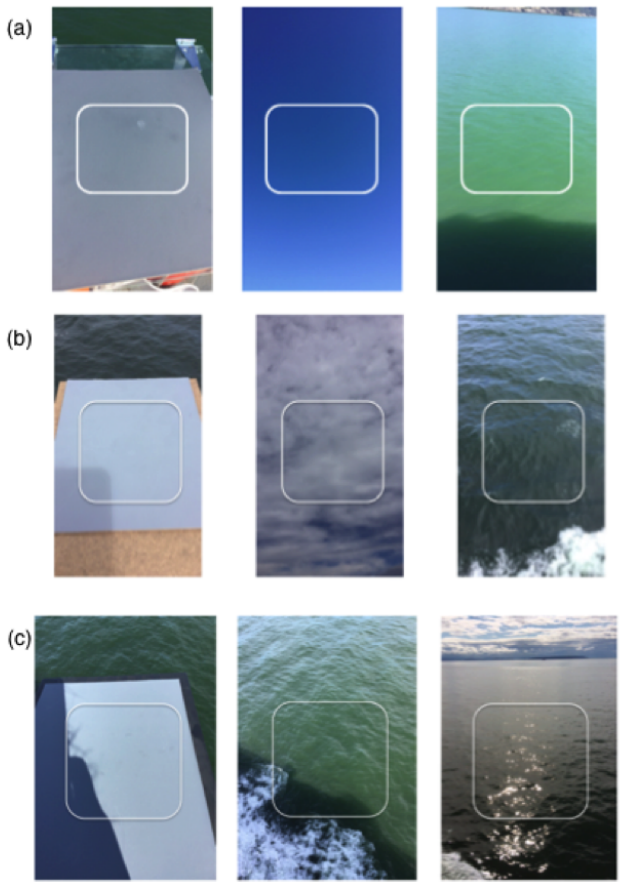

Figure 2.5: Example of HydroColor measurements by quality (a) perfect, (b) good, and (c) bad.

$$R_{rs,SAS}(\lambda) = \frac{L_t(\lambda) - \rho_{sky} L_s(\lambda)}{E_s(\lambda)} \tag{2.3.1}$$

$$\rho_{sky} = \begin{cases} 0.0256 + 0.00039W + 0.000034W^2, & \text{when } \frac{L_s(750)}{E_s(750)} < 0.05 \\ 0.0256, & \text{when} \frac{L_s(750)}{E_s(750)} \geq 0.05 \end{cases} \tag{2.3.2}$$

where ρ_{sky} is the proportion of sky radiance that is reflected off the surface of the water and is dependent on wind speed, W, and the proportion of cloud cover in the sky radiance measurements. A similar approach has been effectively used by other researchers in this region (Komick et al., 2009; Phillips and Costa, 2017). The wind

speed obtained from the Government of Canada website (http://climate.weather.gc.ca/) was the average of the hourly wind speeds at buoy 46146 (49.340N 123.730W), which is near the mainland, and at Entrance Island (49.209N 123.811W), near Nanaimo.

Since $R_{rs,SAS}(\lambda)$ is a function of wavelength with 451 narrow spectral bands, the signal was convolved into three broader spectral bands (red, green, and blue) to match with the HydroColor data. Therefore, the spectral sensitivity function of the iPad was used in this reduction:

$$R_{rs,SAS}(i) = \frac{\sum_{\lambda=400}^{800} R_{rs,SAS}(\lambda) SSF_i(\lambda)}{\sum_{\lambda=400}^{800} SSF_i(\lambda)}, i = R, G, B; \tag{2.3.3}$$

where $R_{rs,SAS}(i)$ is the water reflectance of the i^{th} band, $i = R, G, B$ represents red, green, and blue channels respectively; $SSF_i(\lambda)$ is the spectral sensitivity function of the i^{th} band; $R_{rs,SAS}(\lambda)$ is the above-water remote sensing reflectance acquired by SAS solar tracker and computed by Equation (2.3.1) and (2.3.2).

The sensitivity function of the iPad mini 4 was not available from the manufacturer, and therefore the sensitivity function of the iPhone 5 (Leeuw, 2014) was used instead (Figure 2.6). The feasibility of using this alternative sensitivity function was examined by comparing the HydroColor data acquired with the iPad mini 4 and the iPhone 5 under a clear sky and green water conditions during 11:30 am- 12:30 pm on October 10th, 2016. The mean difference and standard deviation of the water reflectance between the iPad mini 4 and the iPhone 5 were small (0.0002 and 0.0008 for red; -0.0014 and 0.0025 for green; -0.0013 and 0.0024 for blue), particularly for the red band (Table 2.1). The mean absolute percentage differences of bands were relatively small compared with a similar study (69%, 67%, and 77% for red, green, and blue respectively; Mahama, 2016), which compared the differences of three water reflectance bands generated by the iPhone 5 sensitivity function and the CIE1931 2-degree colour matching function. Therefore, we assumed that using the iPhone 5

cameras sensitivity function as an alternative was acceptable when the actual function was unavailable.

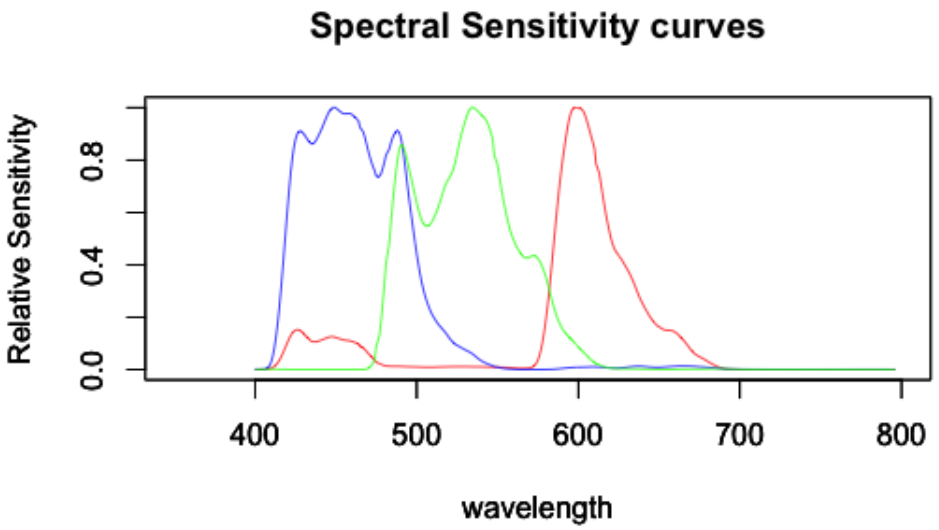

Figure 2.6: The spectral sensitivity function of the iPhone 5 from Leeuw (2014).

Table 2.1: Descriptive statistic of the average difference in reflectance between HydroColor on the iPhone 5 and iPad mini 4. PD is the mean absolute percentage difference.

Band	n	Mean	S.D	PD(%)
Red	48	0.0002	0.0008	40
Green	48	-0.0014	0.0025	28
Blue	48	-0.0013	0.0024	30

2.4 Statistical analyses

Data analysis was conducted in four parts. First, the effects (time, cloud cover, training) on the quality of HydroColor photos and measurements were evaluated. Second, the significance level of the error of $R_{rs,H}$ compared to $R_{rs,SAS}$ was assessed. Third, the error of $R_{rs,H}$ compared to $R_{rs,SAS}$ was evaluated under different environmental

conditions. Finally, we evaluated the influence of other possible effects on the difference between $R_{rs,H}$ and $R_{rs,SAS}$, such as the magnitude of R_{rs}, concentration of chlorophyll-α, and turbidity. We used RStudio version 1.0.136 (RStudio Team, 2016) for all analyses.

2.4.1 Effects on HydroColor measurements by citizens

Contingency tables with the quality of HydroColor citizen data and other factors (sky condition, time, data acquired by the CN or FP) were first analyzed to obtain a preliminary understanding of the effects of those factors on the data quality of $R_{rs,H}$ itself. Chi-squared tests were performed to assess whether these factors were significant.

2.4.2 Assessment of $R_{rs,H}$ compared with $R_{rs,SAS}$

To assess the feasibility of using HydroColor as a substitute to SAS solar tracker with the help of citizens, we analyzed both $R_{rs,H}$ and $R_{rs,SAS}$. The ratios of any two bands reflectance (red/green, blue/green, red/blue) were also computed for both SAS solar tracker (denoted as $R_{rs,SAS}(R/G)$, $R_{rs,SAS}(B/G)$, $R_{rs,SAS}(R/B)$ respectively) and HydroColor (denoted as $R_{rs,H}(R/G)$, $R_{rs,H}(B/G)$, $R_{rs,H}(R/B)$ respectively) for analysis. Band ratios are generally used in ocean colour chlorophyll algorithms, especially the blue/green ratio (O'Reilly et al., 1998; Brewin et al., 2015).

The significance of mean differences between $R_{rs,H}$ and $R_{rs,SAS}$ was evaluated after computing the descriptive statistics for $R_{rs,H}$ and $R_{rs,SAS}$ respectively. All bands and band ratios were considered to be independent and identically distributed random variables; therefore, we analyzed each band or ratio separately. For this analysis, the water reflectance differences between SAS solar tracker and HydroColor was calculated as:

Table 2.2: Data structure of water reflectance for each band or band ratio obtained by SAS solar tracker and HydroColor as well as their difference in reflectance.

Environmental Condition	SAS Solar Tracker $R_{rs,SAS,ijk}$	Hydrocolor $R_{rs,H,ijk}$	Difference d_{ijk}
Clear Sky Morning	$R_{rs,SAS,\cdot 11}, ..., R_{rs,SAS,\cdot 1n_1}$	$R_{rs,H,\cdot 11}, ..., R_{rs,H,\cdot 1n_1}$	$d_{\cdot 11}, ..., d_{\cdot 1n_1}$
Cloudy Morning	$R_{rs,SAS,\cdot 21}, ..., R_{rs,SAS,\cdot 2n_2}$	$R_{rs,H,\cdot 21}, ..., R_{rs,H,\cdot 2n_2}$	$d_{\cdot 21}, ..., d_{\cdot 2n_2}$
Overcast Morning	$R_{rs,SAS,\cdot 31}, ..., R_{rs,SAS,\cdot 3n_3}$	$R_{rs,H,\cdot 31}, ..., R_{rs,H,\cdot 3n_3}$	$d_{\cdot 31}, ..., d_{\cdot 3n_3}$
Clear Sky Noon	$R_{rs,SAS,\cdot 41}, ..., R_{rs,SAS,\cdot 4n_4}$	$R_{rs,H,\cdot 41}, ..., R_{rs,H,\cdot 4n_4}$	$d_{\cdot 41}, ..., d_{\cdot 4n_4}$
Cloudy Noon	$R_{rs,SAS,\cdot 51}, ..., R_{rs,SAS,\cdot 5n_5}$	$R_{rs,H,\cdot 51}, ..., R_{rs,H,\cdot 5n_5}$	$d_{\cdot 51}, ..., d_{\cdot 5n_5}$
Overcast Noon	$R_{rs,SAS,\cdot 61}, ..., R_{rs,SAS,\cdot 6n_6}$	$R_{rs,H,\cdot 61}, ..., R_{rs,H,\cdot 6n_6}$	$d_{\cdot 61}, ..., d_{\cdot 6n_6}$

$$d_{ijk} = R_{rs,SAS,ijk} - R_{rs,H,ijk},$$

$$i = 1, 2, 3, 4, 5, 6; j = 1, 2, 3, 4; k = 1, ..., n_j$$

where $R_{rs,SAS,ijk}$ and $R_{rs,H,ijk}$ are the water reflectance acquired by SAS solar tracker and HydroColor under the i^{th} band or ratio and the j^{th} environmental condition, respectively. The index $i = 1, ..., 6$ represents the red, green, blue bands and red/green, red/blue, blue/green band ratios respectively; index $j = 1, ..., 6$ represents clear sky morning, cloudy morning, overcast morning, clear sky noon, cloudy noon, overcast noon respectively; index k represents the sample number of the j^{th} environmental condition.

The following analyses were applied to each band or band ratio (Table 2.2). First, we used Q-Q plots to test the normality of the difference of time-matched water reflectance between HydroColor and SAS solar tracker, d_{ijk}, for each i. Subsequently, we used a Monte Carlo permutation F test to examine whether the difference, d_{ijk}, was significantly different from zero under all environmental conditions for each i (Manly, 2006).

A permutation test is a nonparametric statistical test used to perform statistical inference for data that do not conform to the required distributional assumption, for instance, normal distribution of standard statistical methods (Legendre and Legendre,

2012). The distribution for inference in permutation tests is generated from the observed data themselves by rearranging the labels of the data. In exact permutation tests, the test statistics of all possible arrangement of the labels are computed to construct an experiential distribution. There were $\frac{n!}{n_1!n_2!...n_i!}$ possible arrangements of labels, where n is the overall sample size, and $n_1, n_2, ..., n_i$ are the sample sizes of each label group (Mielke Jr and Berry, 2007). However, the sample size was not small in this research (n=214, see section 3.2). Therefore, it was difficult to consider all permutations. To solve this problem, we used a Monte Carlo permutation test that randomly selects a large subset of all possible arrangements as an alternative and computes their test statistics (Mielke Jr and Berry, 2007). The Monte Carlo permutation F test was constructed as follows:

1. State the null and alternative hypothesis.

   H_0: $\mu_{D\cdot 1} = \mu_{D\cdot 2} = \mu_{D\cdot 3} = \mu_{D\cdot 4} = \mu_{D\cdot 5} = \mu_{D\cdot 6} = 0$

   H_1: $\mu_{D\cdot j} \neq 0$ for any j

   where $\mu_{D\cdot j}$ is the population mean difference between $R_{rs,SAS}$ and $R_{rs,H}$ with j^{th} environmental condition.

2. Construct a test statistics.

   A reasonable choice to construct a test statistic for the permutation test is to refer to the parametric counterpart (Basso et al., 2009). Based on the F test for comparing a full model and a reduced model, a full model,

   $$d_{\cdot jk} = \mu_{\cdot j} + \varepsilon_{\cdot jk} \quad ,$$

   and a reduced model were built,

   $$d_{\cdot jk} = \varepsilon_{\cdot jk} \quad .$$

Therefore, the error sum of squares for the full model was,

$$SSE(F) = \sum_{j=1}^{6} \sum_{k=1}^{n_j} (d_{\cdot jk} - \mu_{\cdot j})^2 \quad ,$$

and for the reduced model,

$$SSE(R) = \sum_{j=1}^{6} \sum_{k=1}^{n_j} d_{\cdot jk}^2 \quad .$$

Moreover, the F test statistics was,

$$F_{6,(n-6)} = \frac{(SSE(R) - SSE(F))/6}{SSE(F)/(n-6)} \quad , \tag{2.4.1}$$

where n=214 is the overall sample size.

3. Perform the Monte Carlo permutation test.

   The first step was to compute the observed test statistic, F_{obs}, as Equation (2.4.1) based on the observed dataset. Second, the environmental conditions' labels were resampled without replacement to have a rearranged dataset. Third, the F value was computed based on the resample dataset. Then we repeated the second and third steps N_{rep}=10 000 times to obtain an asymptotic distribution of the F test statistic generated by the observed data. The p-value of this Monte Carlo permutation test was,

$$p = \frac{number \ \ of \ \ F_i \geq F_{obs}}{N_{rep}} \quad .$$

   where F_i is the F value of the i^{th} resample dataset.

2.4.3 Evaluation of the error of $R_{rs,H}$ from $R_{rs,SAS}$

For those bands and band ratios that showed a difference in the true average water reflectance between HydroColor and SAS solar tracker, we used Wilcoxon signed-rank tests to determine under which of the six environmental conditions this occurred.

To correct for multiple testing, we used a Bonferroni correction for the significant level of each hypothesis (Dunn, 1961). The corrected significance level was α/m, where α is the overall significant level and m is the number of hypotheses. However, the Bonferroni method is conservative and lacks the power to detect true differences. The Holm-Bonferroni method, which increases the power of the true difference detection, was also used to confirm the result of the Bonferroni method (Blair et al., 1996). The Holm-Bonferroni method is a 'step-down' test that first requires the descending order of p-values such as $p_{(1)} \leqslant p_{(2)} \leqslant ... \leqslant p_{(m)}$ and the corresponding null hypotheses as $H_{(1)}, H_{(2)}, ..., H_{(m)}$. Denote the i^{th} smallest p-value as $p_{(i)}$. Then, we reject $H_{(i)}$ if and only if,

$$p_{(j)} = \frac{\alpha}{m - j + 1}$$

for all $j \leq i$ where α is the desired overall significant level, and m is the number of hypotheses.

For those $R_{rs,H}$ and $R_{rs,SAS}$ matches of bands and band ratios that were significantly different, piecewise models by environmental conditions were built (Equation (2.4.2)) to correct the $R_{rs,H}$ based on the corresponding $R_{rs,SAS}$ for future HydroColor data acquisition.

$$E(R_{rs,corrected}) = \beta_0 + \beta_1 R_{rs,H} + \beta_2 Cloud + \beta_3 Time + \beta_4 R_{rs,H} \times Cloud + \beta_5 R_{rs,H} \times Time \tag{2.4.2}$$

where $Time = \begin{cases} 0 & \text{, morning} \\ 1 & \text{, noon} \end{cases}$; $Cloud = \begin{cases} 0 & \text{, clear} \\ 1 & \text{, cloudy} \end{cases}$.

Moreover, the correction error was examined by leave-one-out cross-validation method (Han et al., 2011). The basic idea of evaluating model performance is to partition the observed dataset into a training dataset and a testing dataset. The training dataset is used to construct a model, and we applied this model to the testing dataset for estimating the prediction error. Leave-one-out cross-validation is a prediction error evaluation techniques used for a limited observed dataset (Witten et al., 2016). We ran the cross-validation N rounds, where N is the sample size of the observed dataset. For each round, one of the observations is partitioned as the testing set, and the rest of observations are the training set. The prediction leave-one-out root mean squared error ε is defined in Equation (2.4.3).

$$\varepsilon = \sqrt{\frac{\sum_{i=1}^{N}(\hat{y}_i - y_i)^2}{N}} \tag{2.4.3}$$

where i is the round number; y_i is the observation in i^{th} testing set; $\hat{y}_i$ is the prediction based on the i^{th} training set.

2.4.4 Evaluation of the effects of R_{rs} magnitude, turbidity, and chlorophyll-α on the difference of $R_{rs,H}$ and $R_{rs,SAS}$

In addition to the effects on acquiring R_{rs}, it is critical to evaluate whether the measured objects impact the measurements collected by HydroColor. For example, the turbidity provided within HydroColor was computed by a simplified programmed model through the measured red band water reflectance (Leeuw, 2014); however, Mahama (2016) indicated that this model had a major limitation in which the tur-

bidity values given by HydroColor were negative when the actual turbidity concentration exceeded 22.57 NTU. Therefore, we explored the effects of the concentration of chlorophyll-α, turbidity, and the magnitude of the R_{rs} signal on the quality of the $R_{rs,H}$. With these needs, the following scatter plots were used: the difference between $R_{rs,SAS}$ and $R_{rs,H}$ of three bands versus the (1) turbidity, (2) concentration of chlorophyll-α, and (3) the corresponding $R_{rs,SAS}$ magnitude.

2.5 Summary

We obtained samples of the R_{rs}, chlorophyll-α, and turbidity data from a BC ferry traveling between Departure Bay and Horseshoe Bay. Both SAS solar tracker and HydroColor were used to acquire R_{rs} data. After sampling, the raw $R_{rs,H}$ and $R_{rs,SAS}$ data were processed to a comparable format of three bands and three band ratios. The $R_{rs,H}$ were also sorted by image quality and environmental condition. The effects of training and environmental condition on image quality were assessed by chi-square tests. The accuracy of each band or band ratio of $R_{rs,H}$ was examined by Monte Carlo F permutation tests separately. A correction model was fitted if the $R_{rs,H}$ data was not accurate. The leave-one-out cross-validation technique was used to assess the performance of this correction model. In addition, we evaluated whether the measured objects impacted the measurements collected by HydroColor. The impacts of turbidity, chlorophyll-α, and the magnitude of R_{rs} were examined using scatter plots.

Chapter 3

Results

This chapter shows the results of (1) the significance levels of the effects on HydroColor data quality; (2) the significance levels of the HydroColor data accuracy; (3) the correction model of the $R_{rs,H}(B/G)$ and the performance of this model; (4) the effects of turbidity, chlorophyll-α, and magnitude of R_{rs} on the difference between $R_{rs,H}$ and $R_{rs,SAS}$.

3.1 HydroColor data quality

All HydroColor data collected were classified as perfect, good, or bad based on the data quality as described in section 2.3.1. There was a total of 1270 HydroColor citizen samples, in which the number of perfect, good, and bad samples were 791, 276, and 203, respectively.

The trained citizen (CN) acquired a higher percentage of perfect data than the untrained citizens (FP) (Table 3.1). Based on a chi-squared test (${X_2}^2$=236.3, p <0.001), there was a significant association between HydroColor data quality and the data collector.

We observed that data acquired at noon showed slightly higher quality than data

Table 3.1: Contingency table of the number (percentage) of HydroColor images by data quality and image collector.

	Image Quality			
Image Collector	Perfect	Good	Bad	Total
CN	636 (77%)	128 (16%)	60 (7%)	824
FP	155 (35%)	148 (33%)	143 (32%)	446
Total	791	276	203	1270

acquired in the morning (65% verse 58% of the perfect category) (Table 3.2). The association between HydroColor data quality and the time of day was also significant ($X_2{}^2$=14.2, p <0.001).

Table 3.2: Contingency table of the number (percentage) of HydroColor images by data quality and time of day.

	Image Quality			
Time	Perfect	Good	Bad	Total
Morning	313 (58%)	115 (21%)	110 (21%)	538
Noon	478 (65%)	161 (22%)	93 (13%)	732
Total	791	276	203	1270

Further, there was a significant association between HydroColor image quality and cloud cover ($X_4{}^2$=48.5, p <0.001). Table 3.3 shows the percentage of perfect data acquired under sky conditions. Clear conditions had a higher percentage of perfect data than under cloudy or overcast conditions, while the highest percentage of bad data (28%) was with overcast conditions since most of the $R_{rs,H}$ values (58 samples, 89% of the bad data) were zero with this condition.

Table 3.3: Contingency table of the number (percentage) of HydroColor images by data quality and cloud cover.

	Image Quality			
Cloud Cover	Perfect	Good	Bad	Total
Cloudy	397 (61%)	154 (23%)	105 (16%)	656
Clear	260 (67%)	93 (24%)	33 (9%)	386
Overcast	134 (59%)	29 (13%)	65 (28%)	228
Total	791	276	203	1270

Table 3.4: Descriptive statistics of $R_{rs,H}$ and $R_{rs,SAS}$ for different bands and band ratios.

Instrument	n	Band	Mean	S.D	Minimum	Maximum
SAS Solar Tracker	214	Red	0.007	0.005	0.001	0.039
		Green	0.010	0.008	0.001	0.058
		Blue	0.008	0.005	0.001	0.041
		Red/Green	0.689	0.075	0.530	0.971
		Red/Blue	0.872	0.144	0.543	1.341
		Blue/Green	0.801	0.085	0.599	1.090
HydroColor	214	Red	0.011	0.008	0.001	0.046
		Green	0.016	0.010	0.002	0.064
		Blue	0.012	0.008	0.001	0.049
		Red/Green	0.681	0.105	0.250	1.011
		Red/Blue	0.954	0.309	0.333	4.000
		Blue/Green	0.740	0.120	0.200	1.286

3.2 Comparison of $R_{rs,H}$ and $R_{rs,SAS}$

The SAS solar tracker data were first matched with the HydroColor citizens' data classified as perfect or good by time stamps. HydroColor data classified as bad were removed from analyses; a total of n=214 matches (97 matches for clear noon, 75 matches for cloudy noon, 22 matches for clear morning, 19 matches for cloudy morning, 1 match for overcast noon, 0 match for overcast morning) were considered in the analysis. The overcast conditions were not included in the following analyses given the low number of samples.

Table 3.4 shows descriptive statistics for (a) $R_{rs,SAS}$ that were matched with perfect and good categories $R_{rs,H}$, and (b) perfect and good categories $R_{rs,H}$. The range and trend of both the $R_{rs,SAS}$ and $R_{rs,H}$ was similar to another study in this region (Phillips and Costa, 2017). The mean and standard deviation of all bands of $R_{rs,H}$ were slightly higher than the corresponding statistics of $R_{rs,SAS}$.

The differences of $R_{rs,SAS}$ and perfect and good category $R_{rs,H}$ samples were first tested using Q-Q plots for normality to determine whether parametric methods could be used (Figure 3.1). The plots indicate that the differences in most bands and band

ratios were not normally distributed given many points in these plots departed from the Q-Q line, especially in the tails. Therefore, nonparametric tests were deemed appropriate.

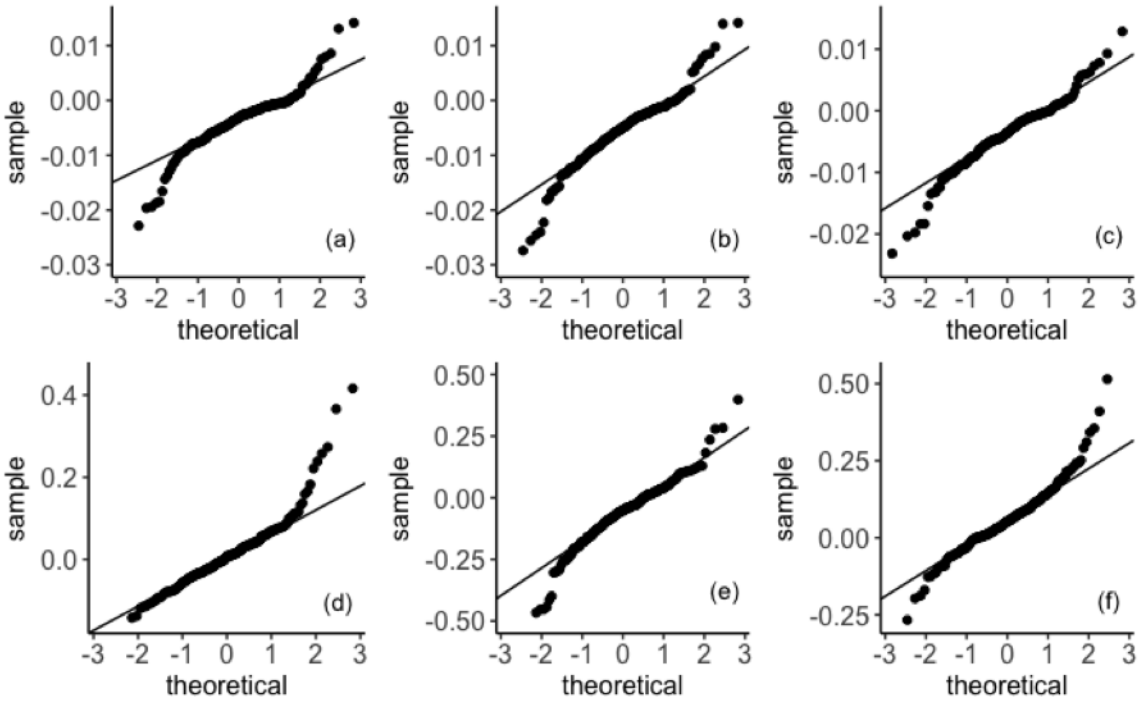

Figure 3.1: Q-Q plots checking the normality assumption of the difference between $R_{rs,SAS}$ and perfect and good quality $R_{rs,H}$ for (a) red band, (b) green band, (c) blue band, (d) red/green ratio, (e) red/blue ratio, and (f) blue/green ratio.

A nonparametric Monte Carlo permutation F test was used to determine whether the true average difference between $R_{rs,SAS}$ and $R_{rs,H}$ differed from zero. Results of the permutation F tests showed that there were no significant differences between the mean difference and zero under any environmental conditions for red, green bands, and the red/green, red/blue ratios. However, the true mean difference was significant under at least one environmental condition for the blue band and the blue/green band ratio (Table 3.5). In addition, $R_{rs,H}$ tended to overestimate the R_{rs} signals for all bands compared to $R_{rs,SAS}$, although the overestimations were tiny given the mean difference in bands (Table 3.5). When perfect quality only $R_{rs,H}$ were evaluated, the

Table 3.5: Statistical summary of the difference between SAS solar tracker and perfect and good quality HydroColor water reflectance in all bands and ratios. F_{obs} and p value are the result of the nonparametric Monte Carlo permutation F test for testing whether the true mean difference in water reflectance from SAS and HydroColor differed from zero.

Band/Band Ratio	n	Mean	S.D	Minimum	Maximum	F_{obs}	p
Red	214	-0.004	0.005	-0.04	0.01	23.00	0.44
Green	214	-0.006	0.006	-0.04	0.01	34.61	0.13
Blue	214	-0.004	0.005	-0.02	0.01	29.55	0.02
Red/Green	214	0.008	0.081	-0.31	0.42	0.70	0.83
Red/Blue	214	-0.082	0.242	-2.88	0.40	6.78	0.07
Blue/Green	214	0.060	0.109	-0.28	0.56	17.71	0.01

Table 3.6: Statistical summary of the difference between SAS solar tracker and perfect quality HydroColor water reflectance in blue bands and blue/green ratio. F_{obs} and p value are the result of the nonparametric Monte Carlo permutation F test for testing whether the true mean difference in water reflectance from SAS and HydroColor differed from zero.

Band/Band Ratio	n	Mean	S.D	Minimum	Maximum	F_{obs}	p
Blue	151	-0.004	0.005	-0.02	0.01	17.46	0.26
Blue/Green	151	0.060	0.094	-0.27	0.52	13.66	0.01

difference of blue band $R_{rs,H}$ and $R_{rs,SAS}$ were not significant; however, the blue/green ratios $R_{rs,H}$ were still significant (Table 3.6).

3.3 $R_{rs,H}(B/G)$ correction with different environmental conditions

To determine which environmental condition(s) affected the average difference in the blue/green ratio, we used Wilcoxon signed-rank tests with a Bonferroni α=0.0125 correction and confirmed using the Holm-Bonferroni method correction (Blair et al., 1996). For the blue/green ratio, the mean differences between SAS solar tracker and HydroColor water reflectance measurements were not significant in the morning

Table 3.7: Multiple comparisons for the difference in the blue/green ratio water reflectance between SAS solar tracker and perfect quality HydroColor data under different environmental conditions. W is the wilcoxon signed rank test statistic, and p is the associated p-value.

Environmental Condition	n	Mean	S.D	W	p	Holm-Bonferroni α
Clear Morning	22	-0.013	0.116	111	0.633	0.050
Cloudy Morning	19	0.042	0.125	127	0.210	0.025
Clear Noon	97	0.053	0.091	3932	<0.001	0.017
Cloudy Noon	75	0.094	0.115	2653	<0.001	0.013

and were significant at noon as shown in Table 3.7. The smaller mean difference occurred under clear morning conditions (-0.013), and the lowest standard deviation of differences was under clear noon conditions (0.091).

Given the statically significant observed differences between $R_{rs,H}$ and $R_{rs,SAS}$ at noon for blue/green ratio, linear models were built to estimate $R_{rs,corrected}(B/G)$ (Equation (3.3.1)) based on the statistical relationship between $R_{rs,H}(B/G)$ and $R_{rs,SAS}(B/G)$.

$$\begin{aligned} R_{rs,corrected}(B/G) = 0.599 + 0.241R_{rs,H}(B/G) + 0.097Cloud - 0.146Time \\ -0.084R_{rs,H}(B/G) \times Cloud - 0.213R_{rs,H}(B/G) \times Time \end{aligned} \tag{3.3.1}$$

$$\text{where } Time = \begin{cases} 0 & \text{, morning} \\ 1 & \text{, noon} \end{cases}; \qquad Cloud = \begin{cases} 0 & \text{, clear} \\ 1 & \text{, cloudy} \end{cases}.$$

Which can also be written as piecewise models respectively:

$$R_{rs,corrected}(B/G) = \begin{cases} 0.453 + 0.028R_{rs,B/G,H} & \text{, clear noon} \\ 0.550 - 0.056R_{rs,B/G,H} & \text{, cloudy noon} \\ 0.599 + 0.241R_{rs,B/G,H} & \text{, clear morning} \\ 0.696 - 0.157R_{rs,B/G,H} & \text{, cloudy morning} \end{cases} .$$

The linear relationship for blue/green ratio was significant ($F_{5,207}$=16.18, p <0.001). The adjusted R^2 for the linear model was weak (${R_{adj}}^2$ =0.26). The adjusted R^2 was used because there are multiple variables in the model (Montgomery et al., 2015). The prediction performance of this model was assessed by the leave-one-out cross-validation (Han et al., 2011). The value of leave-one-out root mean squared error seems small (0.074). However, the percentage of this leave-one-out root mean squared error compared with the mean of $R_{rs,SAS}(B/G)$ was slightly high (124%).

Most of the $R_{rs,H}(B/G)$ underestimated the blue/green R_{rs} band ratio compared to $R_{rs,SAS}(B/G)$ (Figure 3.2a). The linear model solves this underestimation problem especially for the measurements at noon (Figure 3.2b). Figure 3.2b also shows a horizontal trend and points scattering comparatively far away from the one-to-one best-fit line for the morning conditions. The discrepancy may be a result of the small sample sizes of these conditions for parameter estimation. Overall, most of the points scatter around the one-to-one line.

3.4 Effects of magnitude, turbidity and chlorophyll on $R_{rs,H}$

Besides the impact of level of training, time of data acquisition and cloud conditions on the difference between $R_{rs,SAS}$ and $R_{rs,H}$, the magnitude of $R_{rs,SAS}$, water turbidity,

and chlorophyll were explored as further impacts.

Figure 3.3 shows the relationship between the turbidity and the observed differences between $R_{rs,SAS}$ and $R_{rs,H}$ for three bands. As the values of turbidity change, the variability of the differences for three bands do not alter.

Figure 3.4 shows the relationship between the difference in $R_{rs,SAS}$ and $R_{rs,H}$ for three bands and the concentration of chlorophyll-α. Similar to the turbidity analysis, the data distribution indicates that the variability of the differences of three bands did not change with the increase in the concentration of chlorophyll-α.

Figure 3.5 shows the relationship between the magnitude of $R_{rs,SAS}$ and the observed differences between $R_{rs,SAS}$ and $R_{rs,H}$ for the three bands. The distribution of the differences between $R_{rs,SAS}$ and $R_{rs,H}$ was skewed and the majority of the $R_{rs,SAS}$ were below 0.01 sr^{-1} for the blue and red bands, and below 0.02 sr^{-1} for the green band. Also, regardless of the $R_{rs,SAS}$ magnitude, the differences were negative, which indicates that $R_{rs,H}$ magnitudes were generally higher than $R_{rs,SAS}$. The variability of the differences became slightly larger as the magnitude of $R_{rs,SAS}$ became larger for all three bands, especially the red band.

3.5 Summary

In summary, the training and environmental conditions (time of day and cloud cover) had a significant effect on the quality of HydroColor images. Higher quality images were collected by a trained citizen under clear skies at noon. Moreover, most of the bands and band ratios of perfect quality $R_{rs,H}$ were accurate except the blue/green band ratio. A correction model of blue/green ratio was fitted. However, the evaluation of this model shows the performance iw weak. In addition, the magnitude of the R_{rs} has a notable effect on the variability of the difference between $R_{rs,H}$ and $R_{rs,SAS}$.

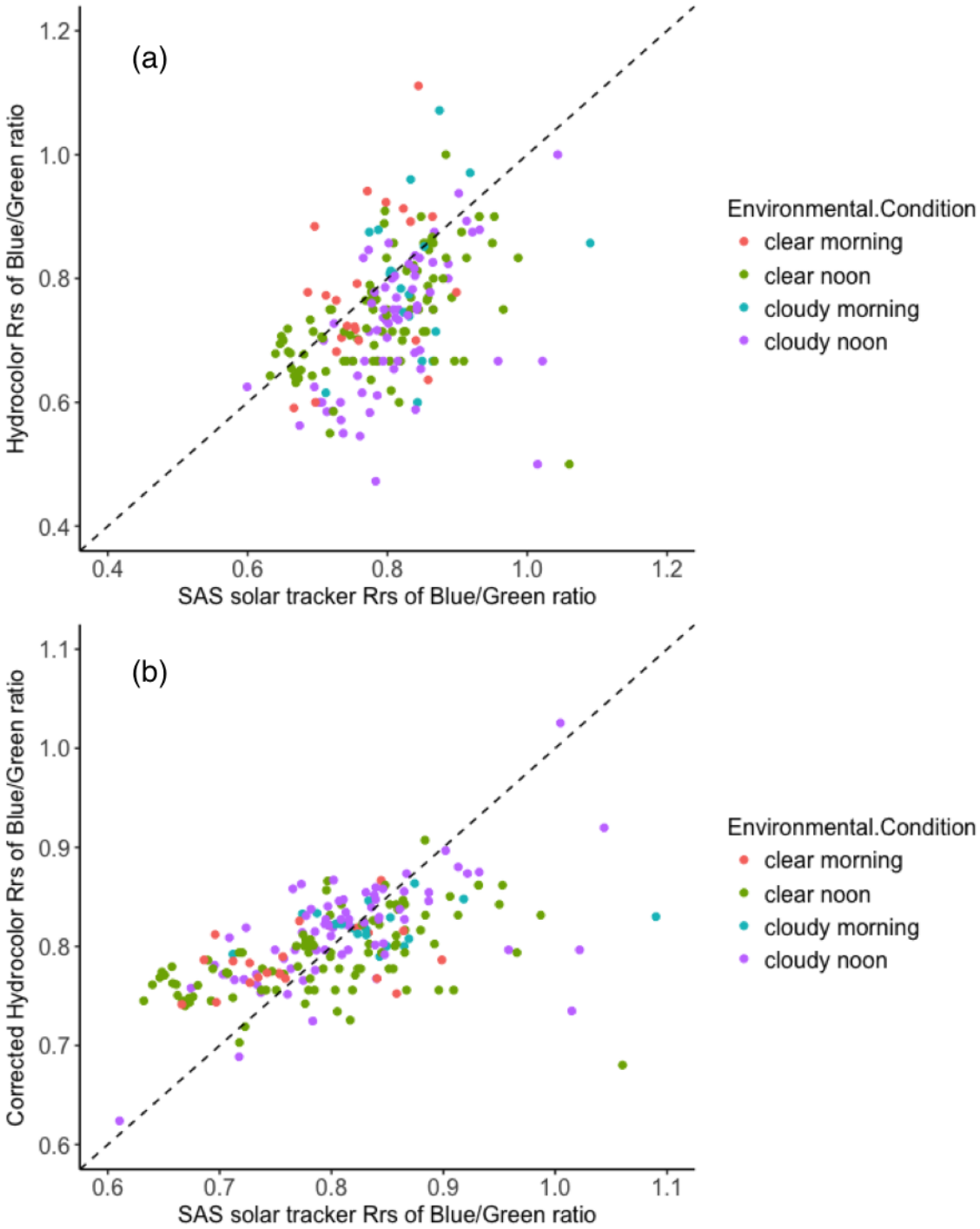

Figure 3.2: Comparison of the $R_{rs,SAS}(B/G)$ with (a) the original $R_{rs,H}(B/G)$ and (b) the corrected $R_{rs,corrected}(B/G)$. The dashed line represents a one-to-one relationship.

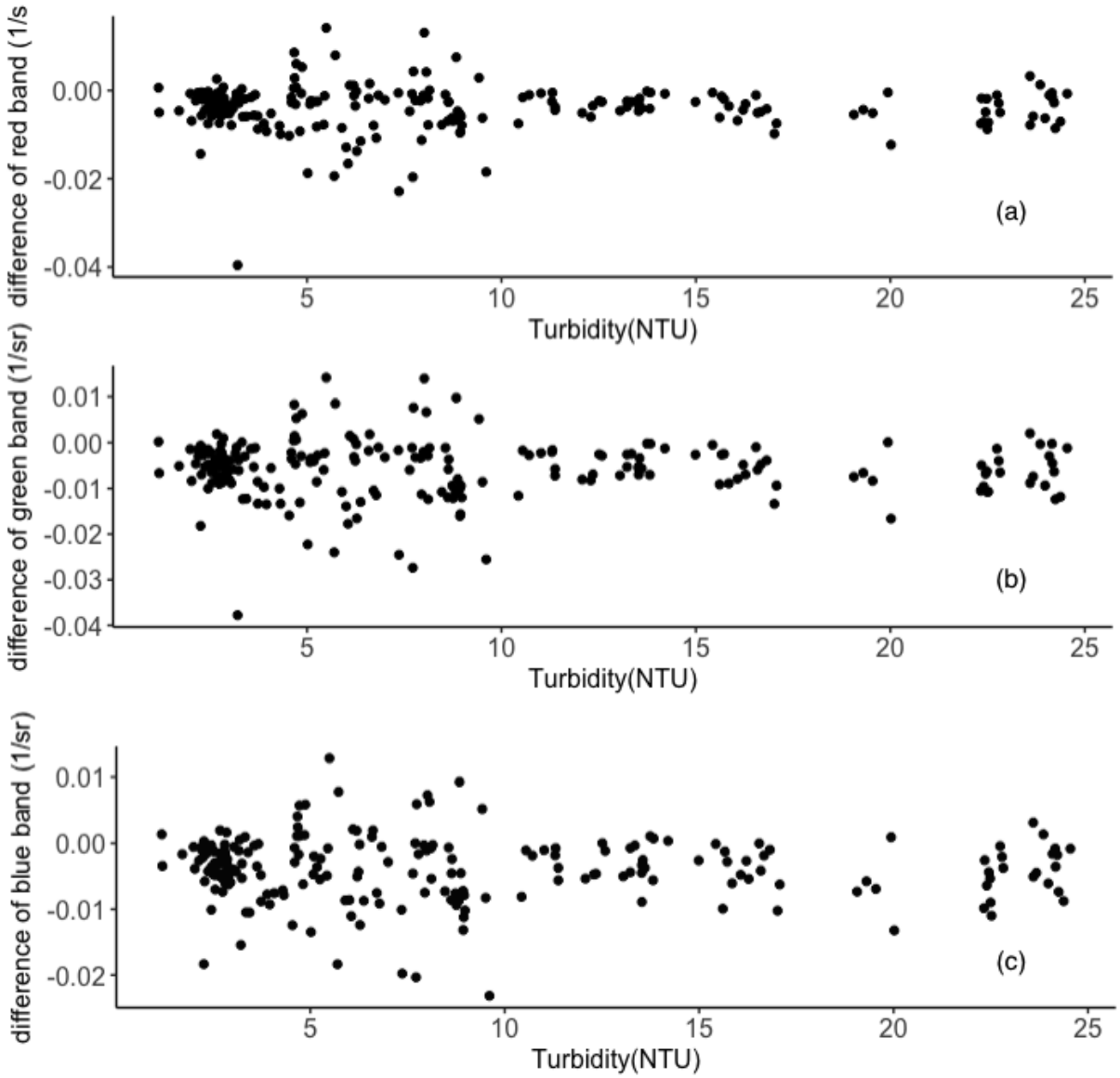

Figure 3.3: Scatter plots of the difference of $R_{rs,SAS}$ and $R_{rs,H}$ for (a) red band, (b) green band, and (c) blue band versus the turbidity.

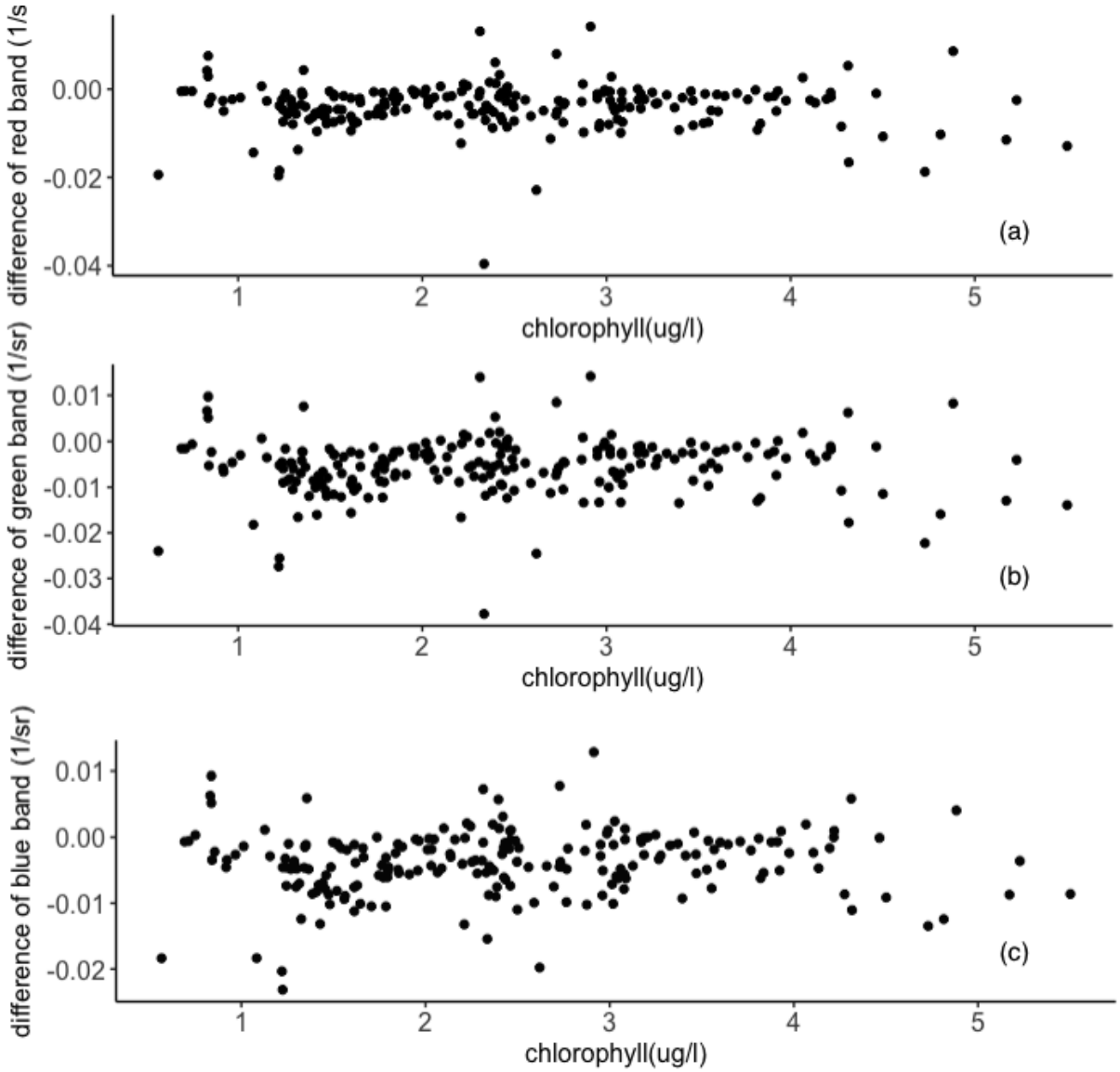

Figure 3.4: Scatter plots of the difference of $R_{rs,SAS}$ and $R_{rs,H}$ for (a) red band, (b) green band, and (c) blue band versus the concentration of chlorophyll-α.

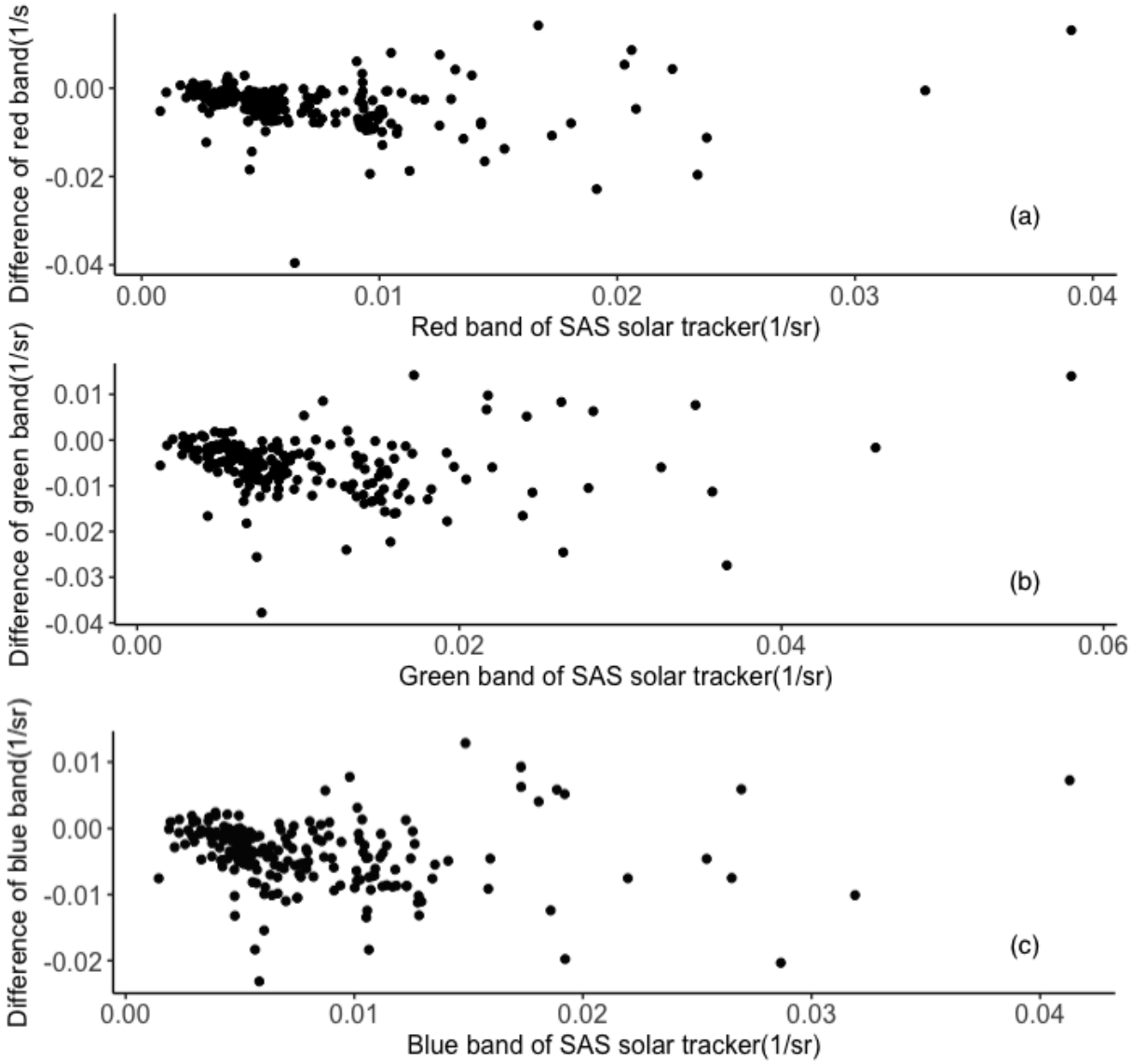

Figure 3.5: Scatter plots of the difference of $R_{rs,SAS}$ and $R_{rs,H}$ versus $R_{rs,SAS}$ for (a) red band, (b) green band, (c) blue band.

Chapter 4

Disscussion

The objective of this study was to examine the accuracy of water reflectance measurements acquired by citizens using the HydroColor application aiming to provide recommendations for extending the use of Hydrocolor to 'fisher scientists'. Water reflectance samples from HydroColor and SAS solar tracker were collected on the BC Ferry Queen of Oak Bay crossing the Salish Sea daily during July and September 2016. The main findings show that the HydroColor citizen data is accurate compared with hyperspectral instrument data for most bands and band ratios; however, citizen level of training and environmental conditions play a role in data quality. The circumstances of citizen participation and how they affect the data quality is first discussed in section 4.1, and then we discuss how the environmental conditions (time, sky condition) impact HydroColor sample quality, followed by the comparisons of SAS solar tracker data and HydroColor data with corresponding explanations in section 4.2.

4.1 Citizen participation and data quality

During the study period, which corresponds to the regional tourist season, over 200,000 passengers per month traveled along the ferry route from Departure Bay

to Horseshoe Bay (traffic data provided by BC Ferries[1]). The total number of passengers participating in the data collection was lower than the number attending the educational talk just before the measurements. Also, the total sample number acquired by regular citizens (FP) was lower than the trained citizen (CN) over the whole period of sampling (Table 3.1, 446 versus 824). This result suggests that either (1) volunteer participation for this type of data acquisition may not be as effective, similar to the findings by Kotovirta et al. (2014), or that (2) the used methodology prevents larger volunteer participation. The first can be dealt with incentive mechanism, such as micro-payments (Reddy et al., 2010). The second, most likely our case, could be prevented by having more trained citizens aboard the ferry to help passengers with data acquisition.

In our methodological framework, only one trained citizen, "the coastal naturalist", was allowed to guide passengers interested in participating in the experiment (a requirement from BC Ferries). Frequently, the CN had several people asking questions after the presentation, making it difficult to help all passengers who were willing to acquire data. It would have been useful to have more than one trained citizen (CN) to help with the data collection. Also, only invited adults (over age 18) were allowed to participate in data collection. Interestingly, children showed a lot of interest in using HydroColor for data acquisition, and it would be a valuable experiment to obtain data quality measures for various age groups in future studies.

The CN (trained citizen) collected a higher number of samples with higher quality than passengers (untrained citizen) (77% versus 35%, Table 3.1). This result was not a surprise as the CN had an educational background in environmental science and also received training on the use of HydroColor before working on this program. When classifying the untrained citizens' samples, the main errors were associated with cap-

[1]http://www.bcferries.com/about/traffic.html

turing the ships shadow and white foam in water photos (Figure 2.5c). Given the required geometry for data acquisition (azimuth angle of 135° from the sun in Leeuw (2014)) and the size of the ferry structure, the ships shadow and foam are likely to appear in some of the water photos if care is not taken. Although the CN gave short oral instructions when passengers acquired data, some passengers failed to learn the usage of HydroColor and collected poor quality samples. Focused training such as that given to the CN, can help with improving skills to avoid these errors. Training effect is highly significant ($p < 0.001$), which implies that a certain degree of training is necessary for high-quality water reflectance sample acquisition using HydroColor. This finding is similar to other studies working with plant species/phenophase (Fuccillo et al., 2015) and terrestrial invertebrate biodiversity (Lovell et al., 2009). Volunteers, having around 6 hours of previous training, could correctly identify 91% of the plant phenophases for a variety of species compared with experts (Fuccillo et al., 2015). The volunteers' sampling performance for invertebrates, with training in the field is similar to ($p > 0.05$) expert researchers (Lovell et al., 2009). Thiel et al. (2014) indicated that appropriate training varies from study to study and should be considered before instructing volunteers.

Besides training, the patience of citizen scientists is also a concern for data quality; for instance, some volunteers did not finish crabs size measurement in the study by Delaney et al. (2008), since these volunteers thought this process was tedious. In our case, passengers needed the patience to find the correct direction for taking qualified samples. Another approach to improve crowdsourcing data quality suggested by Rogstadius et al. (2011) is to emphasize intrinsic motivation such as helping other people.

4.2 Environmental variables and data quality

Similar to many studies (Mobley, 1999; Hooker and Morel, 2003; Wang, 2006; Garaba and Zielinski, 2013), our results demonstrated the importance of environmental variables (time of day and sky condition) on above-water reflectance quality. HydroColor data quality was found to be better (p <0.05) during the noon (time: 8:30-10:10 am, solar zenith: 40.5° to 67.6°) ferry run than the morning (time: 12:50-2:30 pm, solar zenith: 17.4° to 39.3°) run (Table 3.2). A higher percentage of bad samples in the morning (21% versus 13%) was acquired compared with noon most often due to contamination by the shadow of the ship on the water photos. As sun zenith angle increased, the shadow of the vessel became larger and consequently more likely to be detected as part of a sample (Garaba and Zielinski, 2013). In addition to the shadowing effect, the water-leaving radiance is highly related to the time of day or solar zenith (Salisbury, 1998; Wang, 2006). Atmospheric attenuation of solar irradiance in the visible and near-infrared spectra decreases as the solar zenith angle decreases (Salisbury, 1998). At lower zenith angles (noon run), higher irradiance reaches the ocean surface, and therefore more water-leaving radiance backscatter into the sensor (Wang, 2006). In our study, morning runs in late August and the beginning of September happened when zenith angles were between to 60.0° to 67.6°. At these angles, less irradiance reaches the ocean surface, and consequently less radiance from the water is available for detection by the camera. As a result, it was difficult to detect and calculate satisfactory water reflectance samples for these runs. These findings concur with other studies showing that lower sun zenith angle reduces the effects of sun glint, low solar irradiance, weak water-leaving radiance, and wave-shadowing (Garaba and Zielinski, 2013; Mobley, 1999; Salisbury, 1998; Wang, 2006).

Sky condition also affected the quality of the data. Higher quality HydroColor data were collected under clear sky conditions than under cloudy and overcast conditions,

and the cloudiness effect was significant (p <0.05, Table 3.3). This finding is consistent with existing studies recommending a clear sky condition for above-water reflectance measurement (Mobley, 1999; Hooker and Morel, 2003; Garaba and Zielinski, 2013). In principle, atmospheric light attenuation in clear sky conditions, mostly due to Raleigh scattering, is relatively more constant than in cloudy or overcast conditions (Salisbury, 1998). Further, the reflection of clouds in the water is brighter than the reflectance of blue sky; therefore, more skylight, reflected from the water surface was detected by the water-viewing sensor under cloudy conditions (Garaba and Zielinski, 2013; Mobley, 1999). Hooker and Morel (2003) showed that the cloudiness effect is not systematically detected if cloud cover is under a certain threshold. Larger and lower-level cloud cover has a higher effect on the water reflectance measurements by increasing the magnitude of radiance reflected into the sensor (Mobley, 1999).

In addition to noticing the importance of collecting data under a clear sky, we also found that Hydrocolor data acquired under overcast conditions were the poorest quality overall; in particular, a relatively large percentage (89% of bad samples) showed water reflectance values equal to zero. This is likely due to the lower sensitivity of the iPad mini 4 camera at low irradiance conditions. This result corroborates the work of Salisbury (1998), which illustrated that less measurement radiance signal can be detected under a hazy sky because of blurry shadows, and the work of Garaba and Zielinski (2013), which showed that light is more diffuse in hazy skies than clear skies.

Beyond assessing HydroColor data quality, we examined comparisons between HydroColor and the gold standard SAS solar tracker to determine how accurate HydroColor can be as an optical instrument to measure above-water reflectance. The evaluation of the accuracy of qualified (classified as perfect and good quality) HydroColor citizen data (denoted as $R_{rs,H}$) revealed the overall overestimation (mean

in Table 3.4) and higher variability (S.D in Table 3.4) for the observed three bands in relation to the SAS solar tracker data ($R_{rs,SAS}$). A possible reason for this result is that different ρ_{sky} values were used to calculate $R_{rs,H}$ and $R_{rs,SAS}$. ρ_{sky} is used to remove the proportion of the sky radiance that is reflected off the water surface and detected by the water radiance sensor (Mobley, 1999). In a flat sea surface, the reflected sky radiance can be prescribed by viewing geometry alone according to the Fresnel reflectance (Toole et al., 2000). However, the sea surface is usually wavy due to wind, and therefore the surface reflects sunlight from a range of directions other than only the viewing angle (Mobley, 1999). The ρ_{sky} value used to calculate $R_{rs,H}$ was set to a constant 0.028 in the HydroColor application (Leeuw, 2014), while the value used for $R_{rs,SAS}$ was computed by Equation (2), which is related to wind speed and sky condition. The comparisons of HydroColor data and calibrated Water Insight Spectrometer (WISP) data, the golden standard, by Leeuw (2014) does not show that HydroColor tends to overestimate R_{rs} when the same ρ_{sky} is applied to calculate the above-water reflectance.

Further, according to the results of permutation F tests, the observed differences between $R_{rs,H}$ and $R_{rs,SAS}$ for red and green bands were not statistically significant ($p > 0.05$); however, the difference was significant for the blue band ($p < 0.05$) (Table 3.5). The result of lower accuracy in the blue band is also shown in the work of Leeuw (2014), which indicated that $R_{rs,H}$ in the blue band had the highest median percent error compared to the WISP among the three bands. Interestingly, when only perfect quality $R_{rs,H}$ was evaluated with $R_{rs,SAS}$, the statistical analysis showed that the difference of blue band was not significant ($p > 0.05$, Table 3.6). These results imply that minor contaminations in the HydroColor photos still introduce errors in the accuracy of the $R_{rs,H}$ for the blue band. The high sensitivity of minor contaminations for blue band may be caused by (1) comparably weak blue signals, and therefore

subjected to more variability due to noise (low signal-to-noise) (Groetsch et al., 2017); and (2) relatively high difference and variability between the true spectral sensitivity function and the substitute one; (3) the fixed ρ_{sky} programmed in HydroColor, making it difficult to precisely correct the effect of skylight; however the skylight is mostly contributed by blue band radiance due to Rayleigh scattering (Wang, 2002). As such, for accurate retrievals of above-water reflectance at the blue spectra using HydroColor, perfect data quality is required.

The average values and variability (mean and S.D in Table 3.5) of the difference between $R_{rs,H}$ and $R_{rs,SAS}$ for band ratios were higher than individual bands. The results of permutation F tests for band ratios (Table 3.5) showed red/green ratio and red/blue ratio were not significantly different between $R_{rs,H}$ and $R_{rs,SAS}$ (p >0.05), while the difference of blue/green ratio was significant (p <0.05). The significant differences of blue/green ratio were first considered as an effect of the bad performance of the blue band with minor contaminations. However, we note that even discarding samples with contaminations; the significance was not eliminated (p <0.05, Table 3.6). Thus, a linear model with environmental factors as variables was built to correct the error (Equation (3.3.1)). Unfortunately, although this correction model was significant (p <0.05), the small adjusted R^2 (0.26) and the high percentage of correction error (124%) illustrated that it was not satisfying. Therefore the magnitudes of individual bands are recommended to be used for scientific purposes rather than band ratios, especially the blue/green ratio. The most likely explanation for this negative finding is that the blue/green band ratio above-water reflectance is highly sensitive to the differences between $R_{rs,H}$ and $R_{rs,SAS}$ measurement protocols, which may result in a slight difference in measuring location (different field of view-angle of the lens and different footprint on the water) and time. To clarify, the SAS solar tracker was installed at the same side of the ship where the citizens acquired data, but still at

approximately 15 meters distance horizontally. Also, a Hydrocolor sample usually took a minute to finish collecting the grey card, sky, and water photos; whereas SAS solar tracker collected data with three sensors simultaneously. Atmospheric properties and water optical constituents are continually changing, and therefore measured reflectances might be slightly different depending on location and time; for instance, clear blue skies with fast cirrus clouds passing in front of the sun may suddenly change the downwelling irradiance. The studies of Toole et al. (2000) and Leeuw (2014) have also mentioned that the deviations between two instruments are partly caused by imperfect matching on time and location. Interestingly, the variability on water turbidity (Figure 3.3) and chlorophyll-α concentrations (Figure 3.4) did not show any impact on the differences between HydroColor and SAS solar tracker reflectance data.

In addition, the increase of magnitude of true water reflectance raises the variability of differences between $R_{rs,SAS}$ and $R_{rs,H}$ slightly. $R_{rs,H}$ seems to depart from real water reflectance as the magnitude increases, especially the red band. We speculate this departure was a result of the difference between the true and substitute spectral sensitivity functions. If our speculation is correct, the highest average percentage difference in true and substitute spectral functions (Table 1) can account for more notable increasing variability in the red band (Figure 3.5a). We can check this speculation by repeating the sampling and analyses using iPhone 5 instead of iPad mini 4 in the future. In this way, there is no error of the spectral sensitivity functions. However, the R_{rs} collected in this study mostly distributed in relatively small magnitudes ($<0.02\ sr^{-1}$), it is still uncertain whether this increasing variability is physically significant.

4.3 Summary

Training can help to improve patience and provide the skills that avoided acquiring bad quality photos. Time of day and cloud cover affected the data quality by changing the atmospheric attenuation and sky glint. A possible explanation for the low accuracy of the blue/green ratio $R_{rs,H}$ is the high sensitivity of this ratio to the measurement protocols. The effects of the known errors (ρ_{sky} and spectral sensitivity functions) on the accuracy of the water reflectance data were assessed.

Chapter 5

Conclusion and Future Work

We have conducted the first evaluation of the accuracy of Hydrocolor samples collected by citizens aboard a ferry, based on a robust statistical analysis. We have shown that the water reflectance acquired by HydroColor can be used for data acquisition with trained citizen scientists; however, we must be careful when using this method with untrained citizen science fishers, who help with acquiring water reflectance data on the fishing boats. We suggest that oral and written instructions on the use of HydroColor are provided for the 'fisher scientists' before asking for their participation in data acquisition. In these instructions, we recommend that the data be acquired under clear skies at noon to be high-quality samples. General sources of errors are shadows in the grey card, white foam, shadows, and sun glint in water photos.

We showed that it is acceptable to use the spectral sensitivity function of the iPhone 5 as a substitute for the actual function of the iPad mini 4. Nevertheless, the difference between spectral sensitivity functions was likely a source of error. It would be better to acquire the true spectral sensitivity function of the user device if the 'fisher scientists' water reflectance samples need to be compared with ocean colour imagery reflectance in the future.

Future work would include determining the quality of water reflectance samples using HydroColor by 'fisher scientists', especially if the samples will be used as ground-truth data (Kotovirta et al., 2014). All images of HydroColor samples in this study were carefully checked for data quality. However, with large datasets, it would be impossible for an individual to check the data. Thus an automated data quality control system should be developed to enhance the reliability of 'fisher scientists' data (Alabri and Hunter, 2010). Another thing that can help to improve the quality of samples is to make some changes to the HydorColor application. For example, the compass and inclinometer can change to an easier follow guide in the photo capturing interface.

Systematic citizen science research should provide meaningful citizens involvement as well as the use of robust statistical analyses to evaluate the reliability and associated errors of the acquired data (Done et al., 2017). The engagement of citizens not only provides more data needed in natural resource management, but also closes the gap between scientists and citizens (Schläppy et al., 2017).

Bibliography

Alabri, A. and Hunter, J. (2010). Enhancing the quality and trust of citizen science data. In *e-Science (e-Science), 2010 IEEE Sixth International Conference*, pages 81–88. IEEE.

Ave, A.-S. (2017). Marine litter watch app as a tool for ecological education and awareness raising along the Romanian Black Sea coast. *Journal of Environmental Protection and Ecology*, 18(1):348–362.

Basso, D., Pesarin, F., Salmaso, L., and Solari, A. (2009). *Permutation Tests for Stochastic Ordering and ANOVA: Theory and Applications with R*. Springer Science & Business Media, New York, NY.

Bird, T. J., Bates, A. E., Lefcheck, J. S., Hill, N. A., Thomson, R. J., Edgar, G. J., Stuart-Smith, R. D., Wotherspoon, S., Krkosek, M., Stuart-Smith, J. F., et al. (2014). Statistical solutions for error and bias in global citizen science datasets. *Biological Conservation*, 173:144–154.

Blair, R. C., Troendle, J. F., and Beck, R. W. (1996). Control of familywise errors in multiple endpoint assessments via stepwise permutation tests. *Statistics in Medicine*, 15(11):1107–1121.

Blondeau-Patissier, D., Gower, J. F., Dekker, A. G., Phinn, S. R., and Brando, V. E. (2014). A review of ocean color remote sensing methods and statistical techniques

for the detection, mapping and analysis of phytoplankton blooms in coastal and open oceans. *Progress in Oceanography*, 123:123–144.

Brewin, R. J., Sathyendranath, S., Müller, D., Brockmann, C., Deschamps, P.-Y., Devred, E., Doerffer, R., Fomferra, N., Franz, B., Grant, M., et al. (2015). The ocean colour climate change initiative: III. A round-robin comparison on in-water bio-optical algorithms. *Remote Sensing of Environment*, 162:271–294.

Busch, J. A., Price, I., Jeansou, E., Zielinski, O., and van der Woerd, H. J. (2016). Citizens and satellites: Assessment of phytoplankton dynamics in a NW Mediterranean aquaculture zone. *International Journal of Applied Earth Observation and Geoinformation*, 47:40–49.

Butt, N., Slade, E., Thompson, J., Malhi, Y., and Riutta, T. (2013). Quantifying the sampling error in tree census measurements by volunteers and its effect on carbon stock estimates. *Ecological Applications*, 23(4):936–943.

Buytaert, W., Zulkafli, Z., Grainger, S., Acosta, L., Alemie, T. C., Bastiaensen, J., De Bièvre, B., Bhusal, J., Clark, J., Dewulf, A., et al. (2014). Citizen science in hydrology and water resources: Opportunities for knowledge generation, ecosystem service management, and sustainable development. *Frontiers in Earth Science*, 2:26.

Canfield Jr, D. E., Brown, C. D., Bachmann, R. W., and Hoyer, M. V. (2002). Volunteer lake monitoring: Testing the reliability of data collected by the Florida Lakewatch program. *Lake and Reservoir Management*, 18(1):1–9.

Carswell, T., Costa, M., Young, E., Komick, N., Gower, J., and Sweeting, R. (2017). Evaluation of MODIS-Aqua atmospheric correction and chlorophyll products of

western North American coastal waters based on 13 years of data. *Remote Sensing*, 9(10):1063.

Del Bel Belluz, J., Costa, M., Reid, G., and Cross, S. (2016). Bio-optical variability at a Vancouver Island aquaculture site. *Limnology and Oceanography*, 61(5):1686–1704.

Delaney, D. G., Sperling, C. D., Adams, C. S., and Leung, B. (2008). Marine invasive species: validation of citizen science and implications for national monitoring networks. *Biological Invasions*, 10(1):117–128.

Done, T., Roelfsema, C., Harvey, A., Schuller, L., Hill, J., Schläppy, M.-L., Lea, A., Bauer-Civiello, A., and Loder, J. (2017). Reliability and utility of citizen science reef monitoring data collected by reef check Australia, 2002–2015. *Marine Pollution Bulletin*, 117(1):148–155.

Dunn, O. J. (1961). Multiple comparisons among means. *Journal of the American Statistical Association*, 56(293):52–64.

Fuccillo, K. K., Crimmins, T. M., de Rivera, C. E., and Elder, T. S. (2015). Assessing accuracy in citizen science-based plant phenology monitoring. *International Journal of Biometeorology*, 59(7):917–926.

Garaba, S. and Zielinski, O. (2013). Methods in reducing surface reflected glint for shipborne above-water remote sensing. *Journal of the European Optical Society-Rapid Publications*, 8.

Garaba, S. P., Schulz, J., Wernand, M. R., and Zielinski, O. (2012). Sunglint detection for unmanned and automated platforms. *Sensors*, 12(9):12545–12561.

Gordon, H. R. (2005). Normalized water-leaving radiance: revisiting the influence of surface roughness. *Applied Optics*, 44(2):241–248.

Gordon, H. R., Brown, O. B., Evans, R. H., Brown, J. W., Smith, R. C., Baker, K. S., and Clark, D. K. (1988). A semianalytic radiance model of ocean color. *Journal of Geophysical Research: Atmospheres*, 93(D9):10909–10924.

Groetsch, P. M., Gege, P., Simis, S. G., Eleveld, M. A., and Peters, S. W. (2017). Validation of a spectral correction procedure for sun and sky reflections in above-water reflectance measurements. *Optics Express*, 25(16):A742–A761.

Han, J., Pei, J., and Kamber, M. (2011). *Data mining: Concepts and techniques*. Elsevier, Waltham, MA.

Hooker, S. B. and Morel, A. (2003). Platform and environmental effects on above-water determinations of water-leaving radiances. *Journal of Atmospheric and Oceanic Technology*, 20(1):187–205.

Johannessen, S. C., Masson, D., and Macdonald, R. W. (2006). Distribution and cycling of suspended particles inferred from transmissivity in the Strait of Georgia, Haro Strait and Juan de Fuca Strait. *Atmosphere-Ocean*, 44(1):17–27.

Kim, S., Robson, C., Zimmerman, T., Pierce, J., and Haber, E. M. (2011). Creek watch: pairing usefulness and usability for successful citizen science. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 2125–2134. ACM.

Komick, N., Costa, M., and Gower, J. (2009). Bio-optical algorithm evaluation for MODIS for western Canada coastal waters: An exploratory approach using in situ reflectance. *Remote Sensing of Environment*, 113(4):794–804.

Kotovirta, V., Toivanen, T., Järvinen, M., Lindholm, M., and Kallio, K. (2014). Participatory surface algal bloom monitoring in Finland in 2011–2013. *Environmental Systems Research*, 3(1):24.

Lane, N. D., Miluzzo, E., Lu, H., Peebles, D., Choudhury, T., and Campbell, A. T. (2010). A survey of mobile phone sensing. *IEEE Communications magazine*, 48(9):140–150.

Leeuw, T. (2014). Crowdsourcing water quality data using the iphone camera. Master's thesis, Department of Oceanography, University of Maine, Orono, Maine.

Legendre, P. and Legendre, L. F. (2012). *Numerical ecology*, volume 24. Elsevier, Laburnum Cres, UK.

Lovell, S., Hamer, M., Slotow, R., and Herbert, D. (2009). An assessment of the use of volunteers for terrestrial invertebrate biodiversity surveys. *Biodiversity and Conservation*, 18(12):3295.

Mahama, P. N.-J. (2016). Assessment of the utility of smartphones for water quality monitoring. Master's thesis, Department of Geo-information Science and Earth Observation, University of Twente, Enschede, Netherlands.

Manly, B. F. (2006). *Randomization, Bootstrap and Monte Carlo Methods in Biology*, volume 70. CRC Press, Boca Raton, FL.

Masson, D. (2006). Seasonal water mass analysis for the Straits of Juan de Fuca and Georgia. *Atmosphere-Ocean*, 44(1):1–15.

Masson, D. and Peña, A. (2009). Chlorophyll distribution in a temperate estuary: The Strait of Georgia and Juan de Fuca Strait. *Estuarine, Coastal and Shelf Science*, 82(1):19–28.

Mielke Jr, P. W. and Berry, K. J. (2007). *Permutation Methods: A Distance Function Approach*. Springer Science & Business Media, New York, NY, 2 edition.

Mobley, C. D. (1999). Estimation of the remote-sensing reflectance from above-surface measurements. *Applied Optics*, 38(36):7442–7455.

Montgomery, D. C., Peck, E. A., and Vining, G. G. (2015). *Introduction to Linear Regression Analysis*. John Wiley & Sons, Hoboken, NJ.

Neukermans, G., Loisel, H., Mériaux, X., Astoreca, R., and McKee, D. (2012). In situ variability of mass-specific beam attenuation and backscattering of marine particles with respect to particle size, density, and composition. *Limnology and Oceanography*, 57(1):124–144.

Novoa, S., Wernand, M., and Van der Woerd, H. (2014). The modern Forel-Ule scale: a do-it-yourself colour comparator for water monitoring. *Journal of the European Optical Society-Rapid Publications*, 9.

Novoa, S., Wernand, M., and van der Woerd, H. J. (2015). WACODI: A generic algorithm to derive the intrinsic color of natural waters from digital images. *Limnology and Oceanography: Methods*, 13(12):697–711.

O'Neill, J. D. and Costa, M. (2013). Mapping eelgrass (Zostera marina) in the Gulf Islands National Park Reserve of Canada using high spatial resolution satellite and airborne imagery. *Remote Sensing of Environment*, 133:152–167.

O'Reilly, J. E., Maritorena, S., Mitchell, B. G., Siegel, D. A., Carder, K. L., Garver, S. A., Kahru, M., and McClain, C. (1998). Ocean color chlorophyll algorithms for seawifs. *Journal of Geophysical Research: Oceans*, 103(C11):24937–24953.

Perry, R. I. and Masson, D. (2013). An integrated analysis of the marine social–ecological system of the Strait of Georgia, Canada, over the past four decades, and development of a regime shift index. *Progress in Oceanography*, 115:14–27.

Phillips, S. R. and Costa, M. (2017). Spatial-temporal bio-optical classification of dynamic semi-estuarine waters in western North America. *Estuarine, Coastal and Shelf Science*, 199:35–48.

Reddy, S., Estrin, D., Hansen, M., and Srivastava, M. (2010). Examining micro-payments for participatory sensing data collections. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*, pages 33–36. ACM.

Rogstadius, J., Kostakos, V., Kittur, A., Smus, B., Laredo, J., and Vukovic, M. (2011). An assessment of intrinsic and extrinsic motivation on task performance in crowdsourcing markets. *ICWSM*, 11:17–21.

RStudio Team (2016). *RStudio: Integrated Development Environment for R*. RStudio, Inc., Boston, MA.

Ruddick, K. G., De Cauwer, V., Park, Y.-J., and Moore, G. (2006). Seaborne measurements of near infrared water-leaving reflectance: The similarity spectrum for turbid waters. *Limnology and Oceanography*, 51(2):1167–1179.

Salisbury, J. W. (1998). Spectral measurements field guide. Technical report, Earth Satellite Corp Chevy Chase Md.

Schläppy, M.-L., Loder, J., Salmond, J., Lea, A., Dean, A. J., and Roelfsema, C. M. (2017). Making waves: Marine citizen science for impact. *Frontiers in Marine Science*, 4:146.

Sequeira, A. M., Roetman, P. E., Daniels, C. B., Baker, A. K., and Bradshaw, C. J. (2014). Distribution models for koalas in South Australia using citizen science-collected data. *Ecology and Evolution*, 4(11):2103–2114.

Thiel, M., Penna-Díaz, M. A., Luna-Jorquera, G., Salas, S., Sellanes, J., and Stotz, W. (2014). Citizen scientists and marine research: Volunteer participants, their

contributions, and projection for the future. *Oceanography and Marine Biology: An Annual Review*, 52:257–314.

Toivanen, T., Koponen, S., Kotovirta, V., Molinier, M., and Chengyuan, P. (2013). Water quality analysis using an inexpensive device and a mobile phone. *Environmental Systems Research*, 2(1):9.

Toole, D. A., Siegel, D. A., Menzies, D. W., Neumann, M. J., and Smith, R. C. (2000). Remote-sensing reflectance determinations in the coastal ocean environment: impact of instrumental characteristics and environmental variability. *Applied Optics*, 39(3):456–469.

Vandenberg, N., Coady, Y., Costa, M., and Agbaje, T. (2017). Pyscidon: a Python scientific framework for development of ocean network applications. *2017 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*.

Wang, M. (2002). The Rayleigh lookup tables for the seawifs data processing: Accounting for the effects of ocean surface roughness. *International Journal of Remote Sensing*, 23(13):2693–2702.

Wang, M. (2006). Effects of ocean surface reflectance variation with solar elevation on normalized water-leaving radiance. *Applied Optics*, 45(17):4122–4128.

Witten, I. H., Frank, E., Hall, M. A., and Pal, C. J. (2016). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, Cambridge, MA, 4 edition.

Zibordi, G., Mélin, F., Voss, K. J., Johnson, B. C., Franz, B. A., Kwiatkowska, E., Huot, J.-P., Wang, M., and Antoine, D. (2015). System vicarious calibration for ocean color climate change applications: Requirements for in situ data. *Remote Sensing of Environment*, 159:361–369.