Sentinel-3 Satellite Chlorophyll Concentration Validation

Department of Computer Science

University of Victoria

Gaganjot Kaur

Agenda

- Introduction
- Research Question
- Dataset Overview
- Validation Methodology
- Results and Analysis
- Conclusion and Future Work

University of Victoria

Introduction

- Remote sensing imagery affected by atmospheric effects like reflections, haze etc
- To handle these effects, atmospheric correction methods are applied
- Attempt to correlate satellite estimates from Sentinel-3 with shipboard estimates of chl-a (ground truth) in Southern Strait of Georgia.
- Duration of Data Set : July & August 2018

University of Victoria

Research Question

What is the maximum correlation between the chlorophyll concentration of the data obtained from BC Ferries and data obtained from Sentinel-3 satellite in a specific selected range of geographical area?

University of Victoria

Dataset Overview

University of Victoria

ESA Sentinel 3

BC Ferry

- Ocean and Land Color Instrument (OLCI) Sensor (remote sensing)
- Captures area underneath **for three minutes** several times a day
- Available in NetCDF format
- Multi-dimensional; Pixel Representation

- Fluorometers installed by Ocean Networks Canada (in-vivo/in-situ)
- Captures every second
- Available in CSV files
- Two-dimensional representation

University of Victoria

Image source :
1 Sentinel 3 - https://tinyurl.com/y3sgnbgf
2 BC Ferries - https://tinyurl.com/y39sab7c

Sentinel-3 Satellite Data Cleaning

Before

```
longitudes = nc.variables['longitude'][:]
longitudes
```

```
masked_array(
  data=[[-140.083383, -140.07938 , -140.075378, ..., -121.496959,
         -121.493381, -121.489804],
        [-140.083765, -140.079763, -140.075761, ..., -121.498366,
         -121.494789, -121.491212],
        [-140.084148, -140.080146, -140.076144, ..., -121.499774,
         -121.496197, -121.49262 ],
        ...,
        [-141.694965, -141.691704, -141.688442, ..., -126.332651,
         -126.329619, -126.326588],
        [-141.695371, -141.692109, -141.688848, ..., -126.333655,
         -126.330623, -126.327592],
        [-141.695777, -141.692515, -141.689254, ..., -126.334658,
         -126.331627, -126.328596]],
  mask=False,
  fill_value=1e+20)
```

Image 3: Longitude Values in Netcdf

```
latitudes = nc.variables['latitude'][:]
latitudes
```

```
masked_array(
  data=[[52.982373, 52.982151, 52.981929, ..., 50.405646, 50.404827,
         50.404009],
        [52.979813, 52.979591, 52.979369, ..., 50.403146, 50.402327,
         50.401509],
        [52.977254, 52.977031, 52.976809, ..., 50.400646, 50.399828,
         50.39901 ],
        ...,
        [42.505361, 42.505076, 42.504791, ..., 40.078929, 40.078227,
         40.077525],
        [42.502795, 42.50251 , 42.502224, ..., 40.076383, 40.075681,
         40.074979],
        [42.500229, 42.499944, 42.499658, ..., 40.073837, 40.073135,
         40.072433]],
  mask=False,
  fill_value=1e+20)
```

Image 4: Latitude Values in Netcdf

```
chl = nc.variables['logchl'][:]
chl
```

```
masked_array(
  data=[[--, --, --, ..., --, --, --],
        [--, --, --, ..., --, --, --],
        [--, --, --, ..., --, --, --],
        ...,
        [--, --, --, ..., --, --, --],
        [--, --, --, ..., --, --, --],
        [--, --, --, ..., --, --, --]],
  mask=[[ True,  True,  True, ...,  True,  True,  True],
        [ True,  True,  True, ...,  True,  True,  True],
        [ True,  True,  True, ...,  True,  True,  True],
        ...,
        [ True,  True,  True, ...,  True,  True,  True],
        [ True,  True,  True, ...,  True,  True,  True],
        [ True,  True,  True, ...,  True,  True,  True]],
  fill_value=9.96921e+36,
  dtype=float32)
```

Image 5: Logarithmic Chlorophyll Values in Netcdf

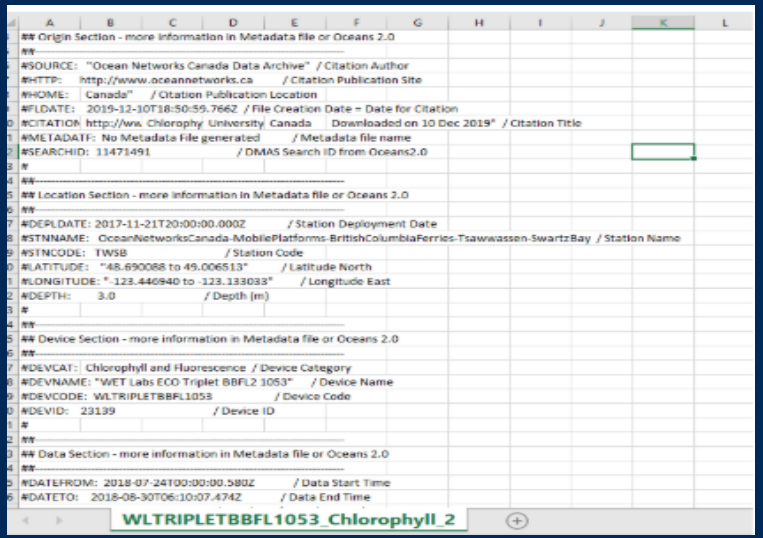
After

date	starttime	endtime	latitude	longitude	chl
2018-07-01	19:07:44	19:10:43	48.99936	-123.13	0.04713
2018-07-01	19:07:44	19:10:43	48.99936	-123.13	0.045955
2018-07-01	19:07:44	19:10:43	48.99847	-123.126	0.024866
2018-07-01	19:07:44	19:10:43	48.99758	-123.122	0.002363
2018-07-01	19:07:44	19:10:43	48.99669	-123.118	0.016258
2018-07-01	19:07:44	19:10:43	48.9995	-123.144	0.010603
2018-07-01	19:07:44	19:10:43	48.99861	-123.14	0.003041
2018-07-01	19:07:44	19:10:43	48.99773	-123.136	0.002187
2018-07-01	19:07:44	19:10:43	48.99684	-123.132	0.006078

Image 6: Sample Processed Satellite Data

BC Ferry Data Cleaning

Before



	A	B	C	D	E	F	G	H	I	J	K	L
1	## Origin Section - more information in Metadata file or Oceans 2.0											
2	##----------											
3	#SOURCE: "Ocean Networks Canada Data Archive" / Citation Author											
4	#HTTP: http://www.oceannetworks.ca / Citation Publication Site											
5	#HOME: Canada" / Citation Publication Location											
6	#FLDATE: 2019-12-10T18:50:59.766Z / File Creation Date - Date for Citation											
7	#CITATION: http://ww. Chlorophy University Canada Downloaded on 10 Dec 2019" / Citation Title											
8	#METADATF: No Metadata File generated / Metadata file name											
9	#SEARCHID: 11471491 / DMAS Search ID from Oceans2.0											
10	#											
11	##----------											
12	## Location Section - more information in Metadata file or Oceans 2.0											
13	##----------											
14	#DEPLDATE: 2017-11-21T20:00:00.000Z / Station Deployment Date											
15	#STNNAME: OceanNetworksCanada-MobilePlatforms-BritishColumbiaFerries-Tsawwassen-SwartzBay / Station Name											
16	#STNCODE: TWSB / Station Code											
17	#LATITUDE: "48.690088 to 49.006513" / Latitude North											
18	#LONGITUDE: "-123.446940 to -123.133033" / Longitude East											
19	#DEPTH: 3.0 / Depth (m)											
20	#											
21	##----------											
22	## Device Section - more information in Metadata file or Oceans 2.0											
23	##----------											
24	#DEVCAT: Chlorophyll and Fluorescence / Device Category											
25	#DEVNAME: "WET Labs ECO Triplet BBFL2 1053" / Device Name											
26	#DEVCODE: WLTRIPLETBBFL1053 / Device Code											
27	#DEVID: 23139 / Device ID											
28	#											
29	##----------											
30	## Data Section - more information in Metadata file or Oceans 2.0											
31	##----------											
32	#DATEFROM: 2018-07-24T00:00:00.580Z / Data Start Time											
33	#DATETO: 2018-08-30T06:10:07.474Z / Data End Time											

WLTRIPLETBBFL1053_Chlorophyll_2

Image 7: Sample Metadata for Ferry Data

#"Time UTC (y	"Chlorophyll (ug/	"Chlorophyll Q	"Latitude	"Latitude Q	"Longitu	"Longitude QC	"Pitch (deg	"Pitch QC	"Roll (deg	"Roll QC	"True Headin	"True Heading
## END HEADER												
2018-08-30T05	3.8552	1	48.6901	8	-123.412	8	-6.5670971	8	3.6	8	311.3541742	8
2018-08-30T05	3.8796	1	48.6901	8	-123.412	8	-5.8903904	8	3.5024024	8	311.1843243	8
2018-08-30T05	3.8918	1	48.6901	8	-123.412	8	-6.2865335	8	3.1143856	8	311.2143457	8
2018-08-30T05	3.9406	1	48.6901	8	-123.412	8	-10.7832	8	3.424	8	315.1964	8

Image 8: Sample Ferry Data

After

Date	Time	Latitude	Longitude	chl
2018-07-01	0:00:00	49.00648	-123.134	7.6616
2018-07-01	0:00:01	49.00648	-123.134	7.686
2018-07-01	0:00:02	49.00648	-123.134	7.6982
2018-07-01	0:00:03	49.00648	-123.134	7.7104
2018-07-01	0:00:04	49.00648	-123.134	7.7104
2018-07-01	0:00:05	49.00648	-123.134	7.7714
2018-07-01	0:00:06	49.00648	-123.134	7.7836

Image 9: Sample Processed Ferry Data

What we have so far?

- [x] Clean Satellite data saved in csv files for each day
- [x] Clean Ferry data saved in csv files for each day

University of Victoria

Validation Methodology

University of Victoria

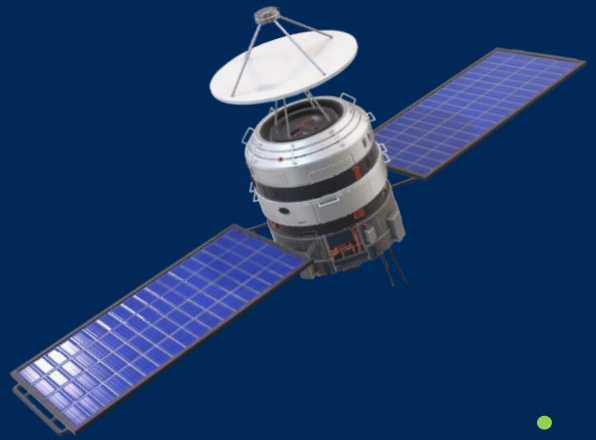
Push-broom sensors capture the area underneath for three minutes

Start-time = time at which capturing started
End-time = time at which capturing ended

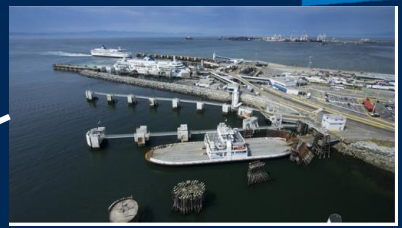
Tsawwassen Terminal

Swartz Bay

Satellite data points or pixels



Position of ferry at end-time

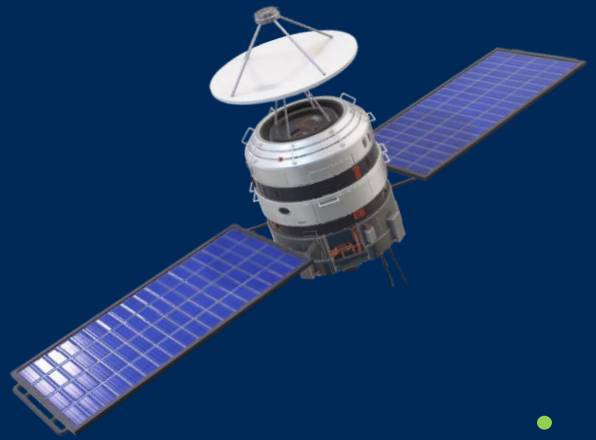


Tsawwassen Terminal

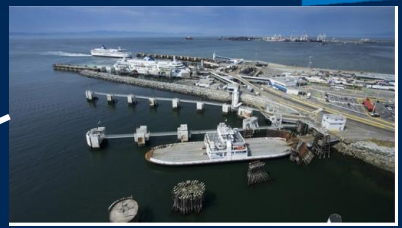
Swartz Bay

Position of ferry at start-time

Distance traversed by ferry in 3 mins



Position of ferry at end-time



Tsawwassen Terminal

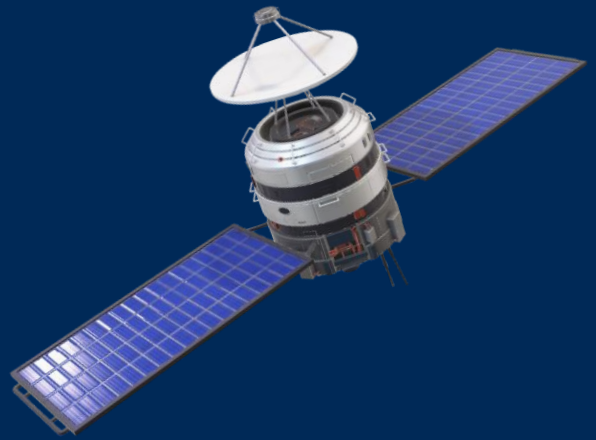
Position of ferry at mid-point for reference

Imaginary circle to narrow down the satellite points that are near to the reference ferry position

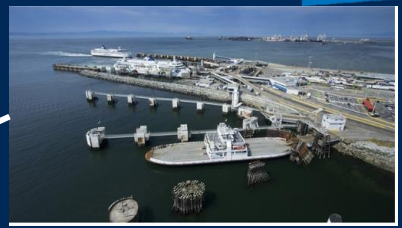
Swartz Bay

Position of ferry at start-time

Distance traversed by ferry in 3 mins



Position of ferry at end-time



Tsawwassen Terminal

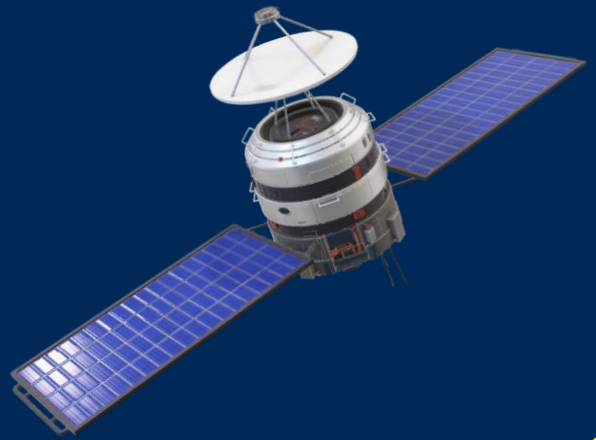
Position of ferry at mid-point for reference

Imaginary circle to narrow down the satellite points that are near to the reference ferry position

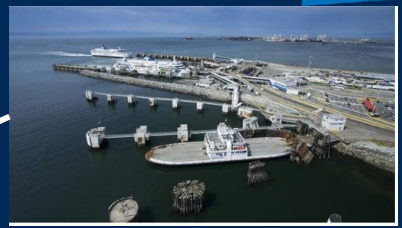
Swartz Bay

Position of ferry at start-time

Distance traversed by ferry in 3 mins



Imaginary circle with radius increased to narrow down the satellite points that are near to the reference ferry position



Tsawwassen Terminal

Position of ferry at mid-point for reference

Swartz Bay

Data extrapolation

Ferry data is gathered for 3 minutes

3 minutes = 180 seconds

Ferry captures data every second
=> There are approx. 180 ferry data points

satellite data points << # ferry data points

For correlation, # satellite data points = # ferry data points
Therefore, random over-sampling is done and Pearson Correlation is measured.

University of Victoria

Pearson Correlation Coefficient

Pearson Correlation (r) is a statistic to measue linear correlation between two variables X and Y (of equal dimensions)

Here, X : Ferry Data and Y : Satellite Data

Absolute Value of r	Strength of Relationship
$r < 0.3$	None or very weak
$0.3 < r < 0.5$	Weak
$0.5 < r < 0.7$	Moderate
$r > 0.7$	Strong

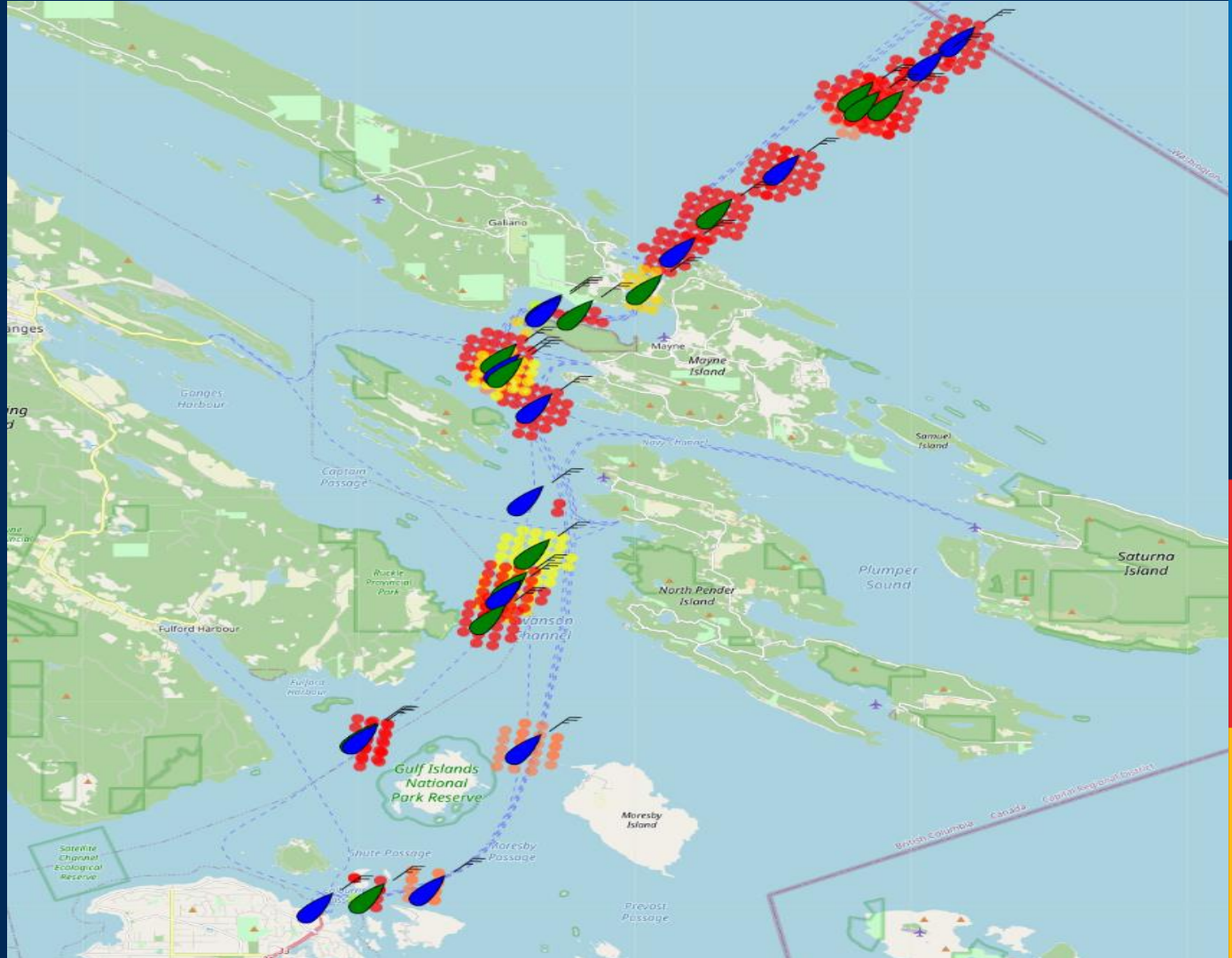
Image 10 : Pearson Correlation Coefficient representing Strength of relationship between two variables

University of Victoria

Results and Analysis

University of Victoria

Image 11 - Choropleth visualization of the BC ferry locations considered for validation over a period of two months (July, August 2018 - Blue, Green Boat markers respectively)



University of Victoria

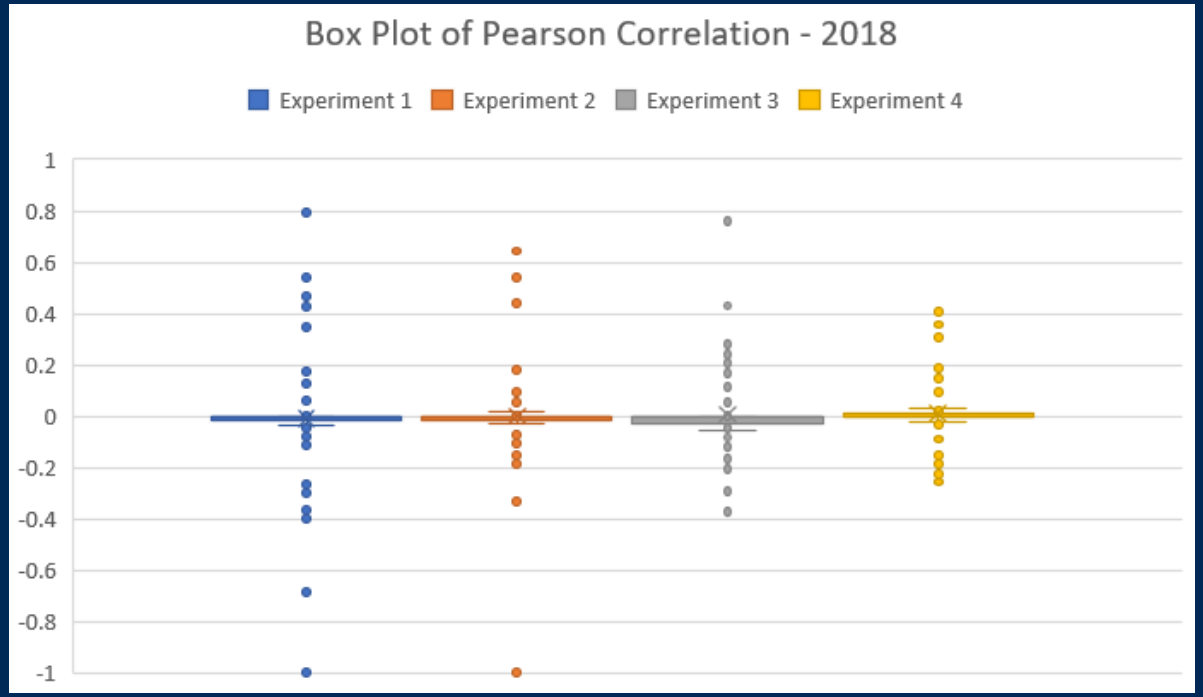


Image 12 : Box plots of Pearson Correlation Coefficients obtained by varying the radius under consideration

University of Victoria

Area b/w Mount Galiano and Mayne Island

Image 13 : Zoomed-in Choropleth Visualization for Mount Galiano and Mayne Island Area

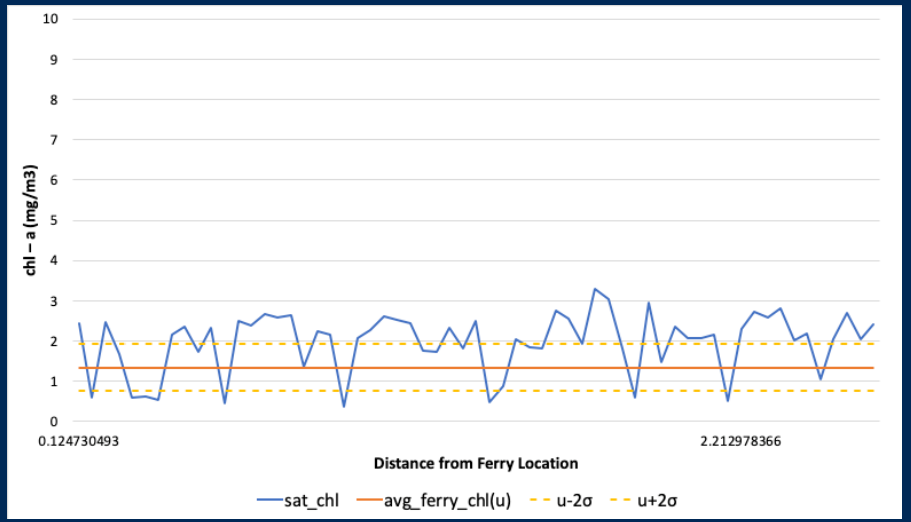


Image 14 : Plot comparing the satellite chl-a values w.r.t mean ferry chl-a values for July 25

University of Victoria

Area near Tsawwassen Terminal

Image 15: Zoomed-in Choropleth Visualization near the port of Tsawwassen and US-Canada Border

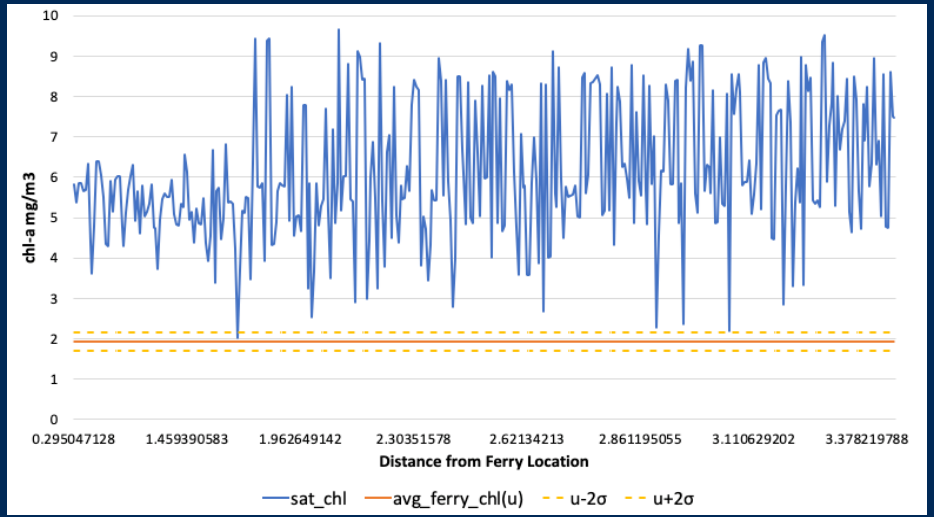


Image 16: Plot comparing the satellite chl-a values w.r.t mean ferry chl-a values for July 1

University of Victoria

Conclusions and Observations

- Few instances have strong/very strong correlation
- Radius increment does not contribute to significant change in correlation
- Removal of extreme outliers show significant increase in correlation
- Average chl-a concentration of satellite is 2-3 times higher than in case of ferry for majority of instances
- Strait of Georgia constitutes rapid waters where chl-a concentration keeps on changing almost instantaneously and is also affected by sediment flow from Fraser River

Future work

- Experiment with Spearman Correlation due to non-parametric distributions present in the data
- Test methodology on larger dataset

University of Victoria

THANK YOU!

University of Victoria