Singapore CDOM

The task is to validate satellite-derived CDOM estimates using *in situ* data in the Singapore region. First, we try that with MODIS-Aqua Level-2 IOP data that can be downloaded from NASA ocean color website.

Station list

Station list has to be a CSV file with columns: *Longitude, Latitude, Date, Time*, etc. Note the order! The original list was like:

Date time	Location	Latitude	Longitude	Salinity	temperature
1/25/2022 11:00	StJohns	1.216	103.8481	31.08	28.05
1/25/2022 13:00	Raffles	1.159	103.7398	31.11	28.15

I loaded the file into Excel and moved the columns and separated "Date time" into *Date* and *Time*. Note that this has to be in **UTC (GMT)** as satellite data are in universal time! My station list became like:

Longitude	Latitude	DateUTC	TimeUTC	Location	Salinity	temperature
103.8481	1.216	1/25/2022	11:00	StJohns	31.08	28.05
103.7398	1.159	1/25/2022	13:00	Raffles	31.11	28.15

Now I saved the Excel file into *station_list_mk.xlsx* and then into a new CSV file *station_list_mk.csv*. Note that I am keeping the original *station_list.csv* unchanged. Note that Excel has automatically changed the *Date* into the US format MM/DD/YYYY. I think that my software works also if the format is YYYY/MM/DD (need to test!). Note that the current station list has only *Salinity* and *temperature*, i.e. no measured CDOM. That means that we cannot actually "validate" the satellite data before we add CDOM data but we can extract the match-ups anyway and add CDOM later. We can evaluate the relationship between *Salinity* and *adg_443* (absorption of detritus and CDOM at 443 nm) and with *adg_s* (the slope). With *adg_443* and *adg_s* we can calculate absorption for other wavelengths.

Make a top level directory for satellite data, e.g. *C:\Sat*. Under that you specify the areas of interest, e.g. *C:\Sat\Singapore* (note – no spaces!). Each year will be a separate directory under that. Here we will process only data from year 2022; therefore we will make only one directory: *C:\Sat\Singapore\2022*. Data will be downloaded into the *tmp* directory of each year, e.g. *C:\Sat\Singapore\2022\tmp*. To process other years, you need to add the respective *tmp* directory under that year, e.g. *C:\Sat\Singapore\2023\tmp* for year 2023.

Ordering and downloading the L2 data

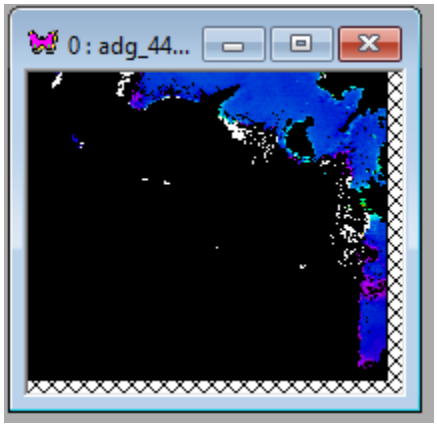
We now order the data at the Level 1-2 browser of NASA's ocean color web at https://oceancolor.gsfc.nasa.gov/cgi/browse.pl?sen=amod. I select year 2022, only MODIS-Aqua. Could also include other sensors. Set the Lat/Long range.

You must choose "**Do**" "**extract my order for me**" and **Level 2 (IOP).** You could also add **Level 2 (OC)** for Chl-a but I skip it here.

When the order is completed, you need to click on the link for the Manifest URL and download the manifest file *http_manifest.txt* to the *tmp* folder.

I download the files listed in the manifest file with *dload.bat* and extract them from tar files with *extract.bat*. Before you can use *dload.bat*, you need to edit it and use your Earth Data username, password and home directory name (where your cookies are). After that you can open command prompt in the *tmp* folder and run *dload.bat*. Before running *extract.bat* please, run the first command in extract.txt and verify that it works. Later you can run *extract.bat* that automatically deletes the tar files.

You can open the extracted *.nc* files in the *tmp* directory and realize that they have mostly invalid data due to clouds. To find files with more data, sort them by size. Bigger files have more valid (as they compress less). For example, *adg_443* in file *AQUA_MODIS.20220523T064500.L2.IOP.x.nc* looks like that:



Note that these images are **not** mapped (they are in satellite view).

Finding match-ups

We can find match-ups from Level-2 data files with a WAM command *wam_match_l2*. You can see the syntax and the many options by typing the name of the command without arguments in the command promt.

Here we extract match-ups of satellite variables *adg_443* and *adg_s*. There are many options how to exclude match-ups according to various flags. Here we use a single option of maximum time difference of 5 days (maxDiffDays=5). We open command prompt in directory *C:\Sat\Singapore* and type:

```
wam_match_l2 station_list_mk.csv 2022\tmp\AQUA_MODIS.2022*.nc adg_443 adg_s maxDiffDays=5
```

(all in 1 line). Note that the 1st argument specifies the station list file (*station_list_mk.csv*). The 2nd argument specifies the directory and the matching filenames of satellite data (*2022\tmp\AQUA_MODIS.2022*.nc*). The following arguments specify the names of the variables to be extracted (*adg_443* and *adg_s*) and the last argument specifies that we are only looking for match-ups with time difference of up to 5 days. The output is stored in a CSV file *station_list_mk_MODISA.csv*. Note that

the name is from the station list file with the name of the satellite sensor added. We have detected 5 match-ups that are acceped according to our criteria. Note that we found match-ups ony for station Raffles and not for StJohns.

The contents of the file starts with columns:

Longitude	Latitude	Date	Time	Location	Salinity	temperature	SYear	EYear	SDay
103.7398	1.159	3/22/2022	11:00	Raffles	31.4	29.03	2022	2022	8
103.7398	1.159	5/24/2022	10:00	Raffles	30.81	30.67	2022	2022	14
103.7398	1.159	7/27/2022	11:00	Raffles	29.37	30	2022	2022	20
103.7398	1.159	9/13/2022	11:30	Raffles	28.4	29.38	2022	2022	25
103.7398	1.159	11/22/2022	10:00	Raffles	30.17	29.21	2022	2022	32

After these columns there are columns like that for each variable (i.e. for both *adg_443* and *adg_s*):

VarName	Pointvalue	Nin	Nout	Min	Max	Mean	StDev	Median
adg_443	-32767	3	6	0.0815	0.0925	0.0873	0.005524	0.0879
adg_443	0.1105	7	2	0.1105	0.1501	0.127243	0.015321	0.1216
adg_443	-32767	3	6	0.0381	0.0819	0.0573	0.022394	0.0519
adg_443	-32767	2	7	0.2769	0.3641	0.3205	0.06166	0.3205
adg_443	-32767	3	6	0.2855	0.345	0.305467	0.034237	0.2859

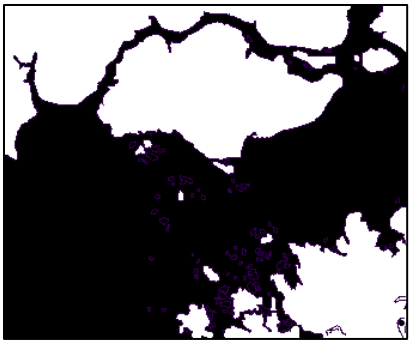
Here *Nin* is the number of valid pixels in the 3 x 3 pixel area centered at the nearest pixel. *Nout* is the number of invalid pixels in the 3 x 3 pixel area centered at the nearest pixel. Followed by statistics (*Min*, *Max*, *Mean*, *StDev*, *Median*) for all the VALID (i.e. *Nin*) pixels.

The last columns show the flags of each of the 9 pixels (i.e. 3 x 3 pixel neighnorhood), e.g. COASTZ+STRAYLIGHT+TURBIDW+PRODFAIL.

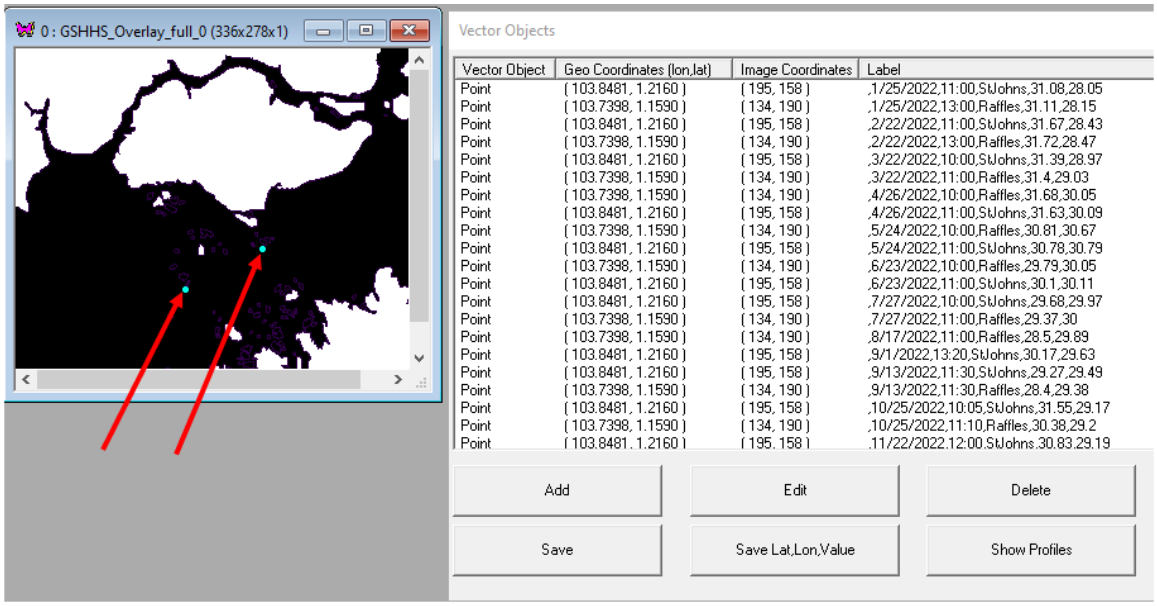
In this case the flags indicate coastal zone, straylight, turbid water and product failure for some variable (not of the included variables).

For making time series of satellite-derived CDOM, etc we need to map them to a standard map of our choice.

We have to decide the spatial domain and create a basic target map. Here we decided to use latitude 1 to 1.5 deg and longitude 103.5 to 104.1 deg. As the area is small, I created the bas map with 200 m pixel size. The base map is not needed for match-ups but will be useful later for creating time series of CDOM images. I saved the map file as sing200m.hdf:



We can also use the map to show the locations of our stations with *Geo-Get Vector Objects*:



The stations *Raffles* and *StJohns* are shown with the red arrows.