Introduction

EPSCoR ACE Southeast Test Case at University of Alaska Southeast is conducting time-series analyses of ocean-environment conditions (productivity and sedimentation) in inner coastal waters of Southeast Alaska. The project will use the MODIS satellite remote sensing sensor on the NASA Aqua and Terra Satellite MODIS as processed into the MOD09GQ dataset by the USGS Land Processes Distributed Active Archive Center (LP DAAC). MODIS bands 1 and 2 (250-m) are used to do regression analysis against on-site sedimentation observation data in order to estimate the sedimentation of inner coastal waters of Southeast Alaska.

Whole project includes the documentation of and activity of: Raw data acquisition, data process, data regression analysis, ocean metrics analysis (spatial analysis) ,and trend analysis (temporal analysis).

The project will be based on linux computing environment. IDL/ENVI, python and bash script programming languages are used in coding. MODIS Reprojection Tool Version 4.1 is used to mosaic tiles, sub-size data, and convert to geotiff files. All programs are located at $proj_home_dir=/home/jiang/projects/sanjay/prog on the computing system. They includes bash scripts, python scripts, and idl programs. The code are shared via github https://github.alaska.edu/jzhu4/epscor_modis_ocean_metrics.git.
MOD09GQ data and stacked data files are located at $proj_data_dir=/mnt/old_root/MOD09GQ on the computing system. Yearly-stacked Data (2001 to 2015) are accessible via
http://static.gina.alaska.edu/NPS_products/MODIS_OCEAN_METRICS/.

Data Acquisition

MODIS 250-m bands 1 and 2 reflectance are included in MODIS product MOD09GQ. The document for the data product can be accessed from https://lpdaac.usgs.gov/products/modis_products_table/mod09gq. The data product can be downloaded from http://e4ftl01.cr.usgs.gov/MOLT/MOD09GQ.005/. Because our study focus southeast of Alaska, only two tiles (H10V03 and H11V03) of daily data are downloaded.

Program: **download_MOD09GQ.py** is a python script used to download the source data from USGS server. Script is located in $proj_home_dir/prog/py/.

Usage:
```
download_MOD09GQ.py start_date end_date output_data_directory
# start_date and end_date are in yyyy.mm.dd format.
```

For example:

```
$proj_home_dir/prog/py/download_MOD09GQ.py 2001.01.01 \
        2001.12.31 $proj_data_dir
```

Downloaded two tile files are stored at $proj_data_dir/YYYY/MM/YYYY.MM.DD/.

Data process

Data process includes two steps. Step 1, Mosaic, sub-size, and file format conversion for daily files. Step 2, stack one year of daily files into a big stacked file.
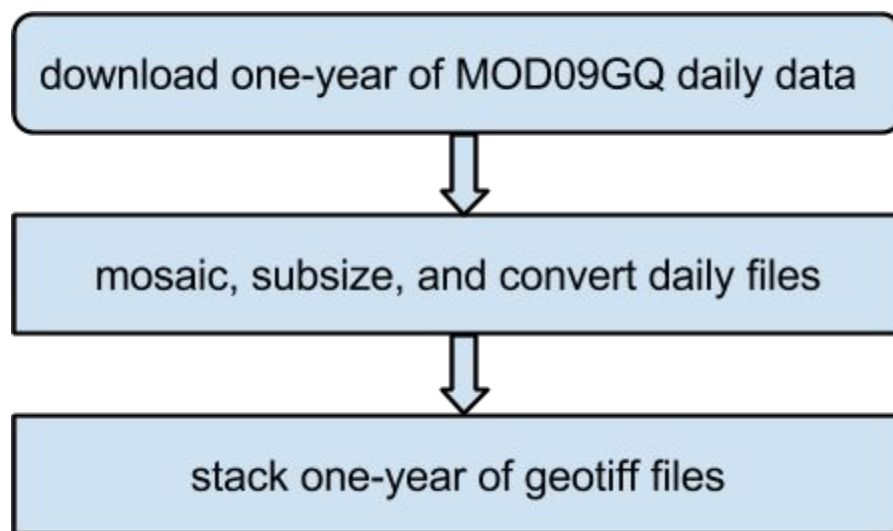

Figure 1. schematic of data acquisition and process

1. Mosaic, sub-size, and convert file type for daily files

MOD09QG dataset is composed of tile files. Two tiles H10V03 and H11V03 cover southeast of Alaska study area. In order to decrease raster size, we subsize the data according to the study area definition. Figure 2 shows the IOA of the study. The subset of data is defined in the parameter.prm file. The content of the parameter file are as following:

```
SPECTRAL_SUBSET = ( 1 1 1 1 1 )
SPATIAL_SUBSET_TYPE = INPUT_LINE_SAMPLE
SPATIAL_SUBSET_UL_CORNER = (0     1320)
SPATIAL_SUBSET_LR_CORNER = (2652  5737)
RESAMPLING_TYPE = NEAREST_NEIGHBOR
OUTPUT_PROJECTION_TYPE = SIN
```

OUTPUT_PROJECTION_PARAMETERS = (6371007.181 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0 0.0)
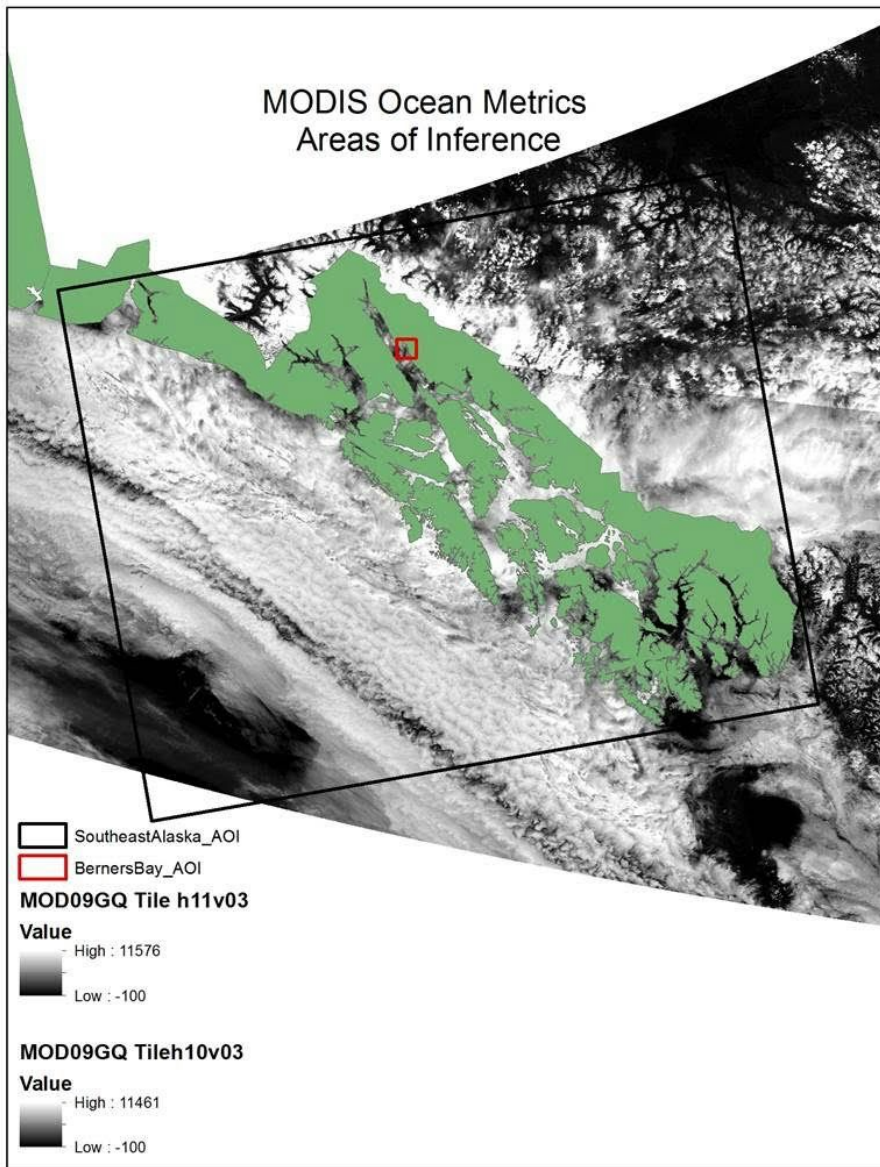DATUM = NoDatum

Figure 2. IOA of study

Each MOD09QG tile file includes 5 science dataset. We output each science dataset as a single geotiff file.

Program: `mosaic_data_range.bash` uses MODIS MRT V.4.1 to do the mosaic, subsize, and conversion for daily files. mosaic_data_range.bash can be found in project scripts directory:
`$proj_home_dir/prog/scripts/`
Usage:

```
mosaic_data_range.bash  src_data_main_dir start_date end_date
# Dates specified in the format of: yyyy.mm.dd
```

for example:

```
$proj_home_dir/prog/scripts/mosaic_date_range.bash \
         $proj_data_dir 2001.01.01 2001.01.03
```

2. Stack one-year of daily data

Each science dataset of a daily mosaiced file is output as a single geotiff file. We stack daily geotiff mosaiced files of a year into a big stacked file. Totally, we get 5 big stacked files, one file per science dataset. During stack stage, we also mask pixels of bands 1 and 2 with quality control band. Data quality control are defined in Table 1. We keep values of "good" pixels, and fill other with -9999.
"good" pixels have following characters:
00b-corrected product produced at ideal quality all bands
00b-clear
0000b-band1 data highest quality
0000b-band2 data highest quality
1b-atmospheric correction performed
0b-adjacency correction not performed
The mask value of these "good" pixels is 0b0001000000000000 or 0d4096.
After mask the data, I subtract the minimum value of the bands 1 and 2 with their minimum values, respectively. That is: band1=band1-min(band1), and band2=band2-min(band2).
stack these processed daily files together.

Table 1: MOD09GQ.005 250-meter Surface Reflectance Data QA Descriptions (16-bit)

Bit No.	Parameter Name	Bit Comb.	Sur_refl_qc_250m
0–1	MODLAND QA bits	00	corrected product produced at ideal quality all bands
		01	corrected product produced at less than ideal quality some or all bands
		10	corrected product not produced due to cloud effects all bands
		11	corrected product not produced due to other reasons some or all bands may be fill value [Note that a value of (11) overrides a value of (01)].
2–3	cloud state	00	clear
		01	cloudy
		10	mixed
		11	not set; assumed clear
4–7	band 1 data quality four bit range	0000	highest quality
		1000	dead detector; data interpolated in L1B
		1001	solar zenith >= 86 degrees
		1010	solar zenith >= 85 and < 86 degrees
		1011	missing input
		1100	internal constant used in place of climatological data for at least one atmospheric constant
		1101	correction out of bounds, pixel constrained to extreme allowable value
		1110	L1B data faulty
		1111	not processed due to deep ocean or clouds
8–11	band 2 data quality four bit range		SAME AS BAND ABOVE
12	atmospheric correction performed	1	yes
		0	no
13	adjacency correction performed	1	yes
		0	No

| 14– | spare (unused) | - | --- |
| 15 | | | |

(from)

program: $proj_home_dir/prog/scripts/stack_one_year_tif.bash

usage: stackhdf.py filename_list

For example: $proj_home_dir/prog/scripts/stack_one_year_tif.bash $proj_data_dir/2001/2001_flist_num_observations

File list for *_num_observations is produced by

program: $proj_home_dir/prog/scripts/create_flist.bash

Usage: ./create_flist.bash data_main_dir, year(yyyy)

for example: ./creat_flist.bash /mnt/old_root/MOD09GQ 2001

Data product in phase one

We produced yearly stacked data files for year 2001 to 2014. Data are available.

Data analysis

1. stacked yearly data can be used to do spatial and temporal analysis. In order to quantitatively analyze the total suspended sedimentation (TSS), reflectance data should be used to regression analysis with on-site TSS observation data.

2. spatial analysis

3. Trend analysis

Conclusions