Merging SST data

1 Introduction

Sea-surface temperature (SST) data can be obtained from different sources. Most important are polar orbiting infrared (IR) sensors (e.g. AVHRR, MODIST, MODISA), polar orbiting microwave (MW) sensors (e.g. TMI, AMSR-E), geostationary IR sensors (GOES SST) but also *in situ* data buoys and other satellite and aircraft sensors. Infrared sensors have typically best spatial resolution (~ 1 km or better) and best accuracy but are limited to cloud-free conditions that may be rare in certain areas. Microwave (MW) sensors have the advantage of "seeing" through most clouds but have lower spatial resolution (~ 10 km) and lower accuracy. A common method to deal with missing data is to composite over time, i.e. instead of daily images use 8-day composites or even monthly composites. However, sometimes even 8-day and monthly composites have no valid data and compositing over long time period may introduce significant errors due to dynamic conditions. The solution to obtain best SST data is to merge SST data from different sources, so that the result has the best resolution in time and space as well as good accuracy. Advanced methods include data merger with various models.

For example, the New Generation Sea Surface Temperature (NGSST) for Open Ocean (http://www.ocean.caos.tohoku.ac.jp/~merge/sstbinary/actvalbm.cgi?eng=1) is a merged SST product for the ocean around Japan that is being assembled from various sources by the Kawamura Lab in the Tohoku University, Japan. The NGSST dataset produces high-resolution, daily, high-accuracy SST data that is not possible from any single source. For information how to read NGSST data with WIM please read paragraph "New Generation Sea Surface Temperature for Open Ocean" in the WIM manual (WIM.pdf). A WAM program *wam_convert_ngsst* converts NGSST files to HDF files with important attributes. However, NGSST data covers only areas around Japan. If you need good SST data outside that area then you can try yourself to merge data from different sources. This exercise shows how to merge high-resolution but frequently missing (due to clouds) MODISA data with low-resolution but almost always available AMSR-E data.

2 Prerequisites

We assume that you are familiar with the basics of the command line, i.e. how to open the command window, change directory, issue a command, etc. We also assume that you are familiar and able to run WAM command line programs. If not, please check out the basic WAM and WAM exercises manuals. We also assume that you have a set of images that you are going to use. In our example we use daily MODISA 9-km images from NASA's Ocean Color website and 3-day AMSR-E SST data assembled by Remote Sensing Systems (http://www.ssmi.com/amsr/amsr_data_description.html#amsre_data) and converted to HDF with a WAM program *wam_convert_amsre* (see paragraph "AMSR-E data from Remote Sensing Systems" in WIM manual (WIM.pdf).

3 Screening the images

The first task is to screen the data and keep the best quality (quality 0 for MODIS and 7 for AVHRR Pathfinder5). The pixels with lower quality are usually contaminated by clouds). You can use *wam_screen_sst_ocpg* for screening the new MODIS SST data produced by the Ocean Color Processing Group and *wam_screen_pf* for screening AVHRR Pathfinder data. Please see the corresponding section in the WAM Manual (*WAM.pdf*).

MODISA SST data can be downloaded from ftp://oceans.gsfc.nasa.gov/MODISA/ and Remote Sensing Systems AMSR-E data from ftp://ftp.ssmi.com/amsre.

You can note that the 9-km SST file is over 27 MB and the 4-km SST file over 109 MB in size. If you want to keep those then you can safely compress them with a WAM command *wam_compress_hdf* without loosing anything. WIM/WAM use lossless internal HDF compression that is invisible to the user. After running *wam_compress_hdf* the size is reduced to about 9 MB and 33 MB, respectively (i.e. about 3 times).

We perform screening of the MODISA data with the following command:

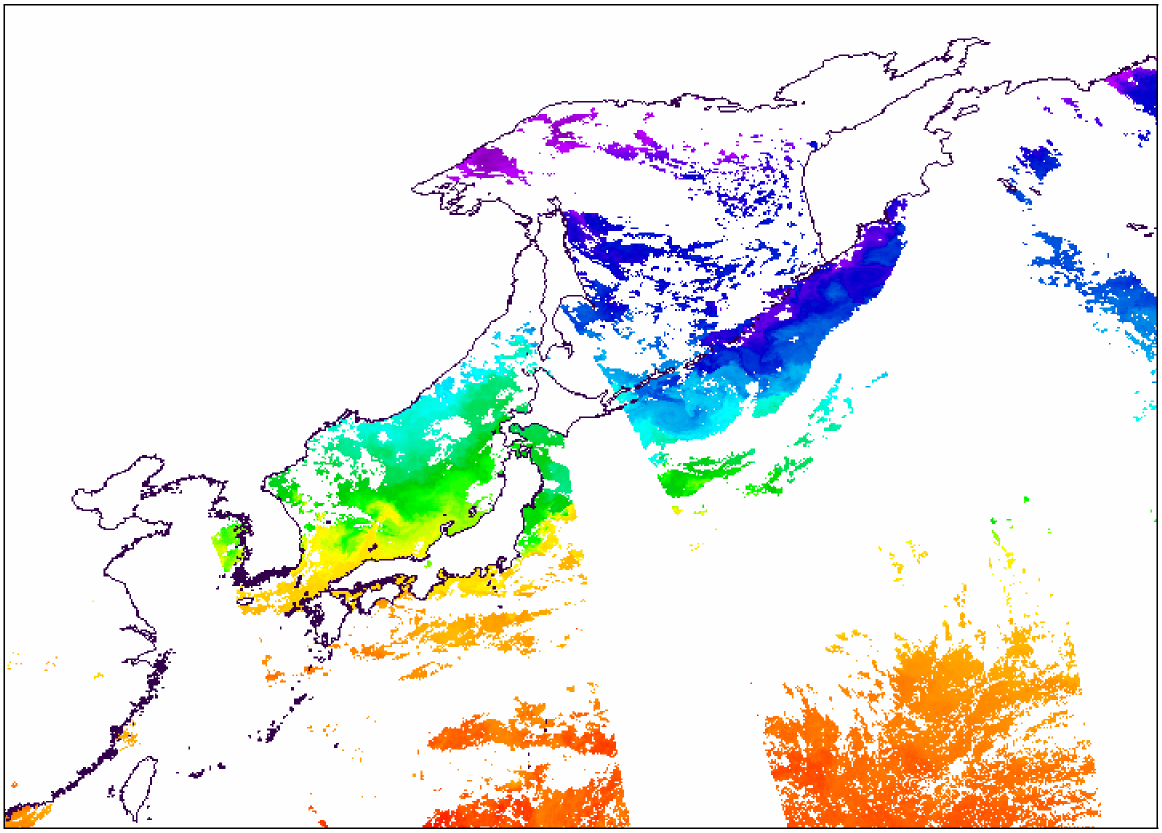
*wam_screen_sst_ocpg   A*_9*

This assumes that we have 9-km data. The same procedures can be used for 4-km MODISA data.

We convert AMSR-E binary 3-day data to HDF and extract the SST dataset (number 0) with the following command:

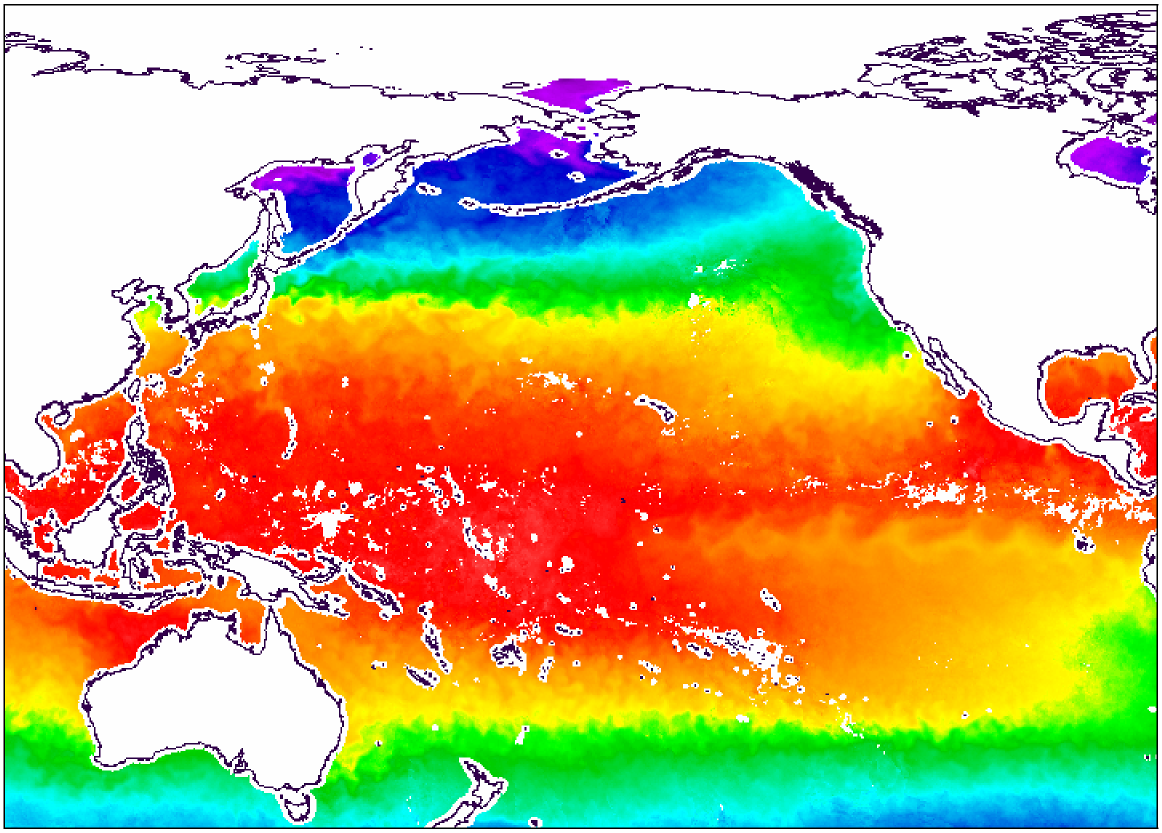
*wam_convert_amsre   *_d3d 0*

A subset of the resulting MODISA SST data for October 30, 2006 (Julian day 303) around Japan is shown below. For better visualization the missing pixels have been converted from black to white interactively with WIM with *Transf – Replace Values*, coastline was created with *Geo-Get Map Overlay* and added with *Multi – Overlay Image*.



As you can see, large areas have no data either due to clouds, sun glint or orbit characteristics.

A subset of the AMSR-E SST 3-day data for October 30, 2006 is shown below. For better visualization the coastline was created with *Geo-Get Map Overlay* and added with *Multi – Overlay Image*.



As you can see, the image is nearly complete but at much lower resolution (approximately 27 km). Data is missing near land and in areas of strong rain.

Our task is to merge the two datasets to get the best coverage.

4 Merging data

The syntax of the command is:

wam_merge_l3 Pattern1 Pattern2 [DeltaDays [PriorityOne]]

Here Pattern1 is the matching filename patterns of Set1 and Pattern2 is the matching filename pattern of Set2. We will now merge a set of daytime MODISA 9 km daily SST files with AMSR-E 3-day SST files. We assume that MODISA data are in a folder *2006* and AMSR-E data in a folder *AMSR-E* and we issue a command:

*wam_merge_l3 2006\A200630*9_screened.hdf AMSR-E\amsre_2006*v5_d3d.hdf 2 yes*

According to this command we assume that the MODISA will be the primary data and that the merged data will be in the same projection as MODISA data. Also, we set the maximum time difference for merging at 2 days {"2" as *DeltaDays*) and we will use AMSR-E data only if MODISA data are missing ("yes" for *PriorityOne*).

Using Image1 first, if IMage1 data is missing then Image2 data

Folder 1:

10 matching files found

Folder 2:

31 matching files found

======================

A2006300.L3m_DAY_SST_9_screened.hdf

amsre_20061027v5_d3d.hdf

...remapping amsre_20061027v5_d3d.hdf...

Saved M2006300.L3m_DAY_SST_9_screened.hdf

A2006301.L3m_DAY_SST_9_screened.hdf

amsre_20061028v5_d3d.hdf

...remapping amsre_20061028v5_d3d.hdf...

Saved M2006301.L3m_DAY_SST_9_screened.hdf

A2006302.L3m_DAY_SST_9_screened.hdf

amsre_20061029v5_d3d.hdf

...remapping amsre_20061029v5_d3d.hdf...

Saved M2006302.L3m_DAY_SST_9_screened.hdf

A2006303.L3m_DAY_SST_9_screened.hdf

amsre_20061030v5_d3d.hdf

...remapping amsre_20061030v5_d3d.hdf...

Saved M2006303.L3m_DAY_SST_9_screened.hdf

A2006304.L3m_DAY_SST_9_screened.hdf

amsre_20061031v5_d3d.hdf

...remapping amsre_20061031v5_d3d.hdf...

Saved M2006304.L3m_DAY_SST_9_screened.hdf

A2006305.L3m_DAY_SST_9_screened.hdf

amsre_20061031v5_d3d.hdf

...remapping amsre_20061031v5_d3d.hdf...

Saved M2006305.L3m_DAY_SST_9_screened.hdf

A2006306.L3m_DAY_SST_9_screened.hdf

amsre_20061031v5_d3d.hdf

...remapping amsre_20061031v5_d3d.hdf...

Saved M2006306.L3m_DAY_SST_9_screened.hdf

A2006307.L3m_DAY_SST_9_screened.hdf has NO match in Set 2

Copied M2006307.L3m_DAY_SST_9_screened.hdf

A2006308.L3m_DAY_SST_9_screened.hdf has NO match in Set 2

Copied M2006308.L3m_DAY_SST_9_screened.hdf

A2006309.L3m_DAY_SST_9_screened.hdf has NO match in Set 2

Copied M2006309.L3m_DAY_SST_9_screened.hdf

======================================

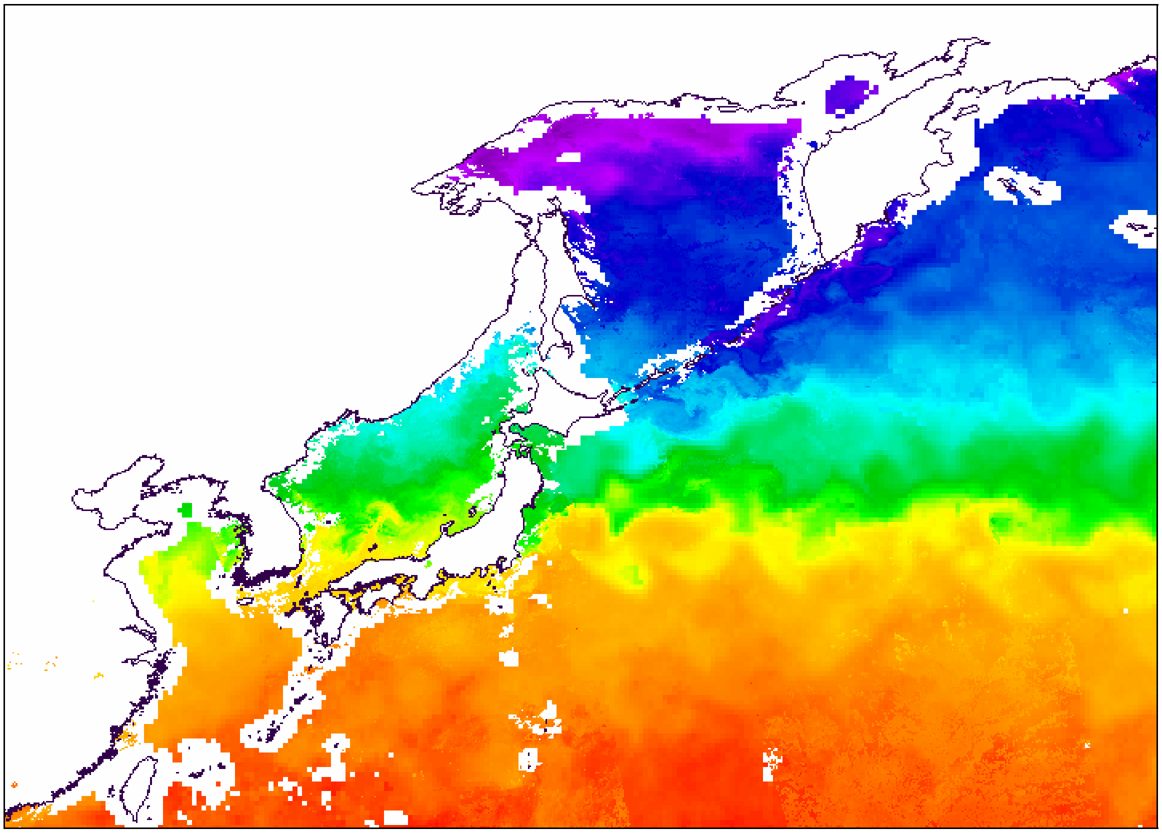
Summary:

7 merged files from 2 datasets

3 files from only Set-1

According to this, we merged 7 pairs of datasets and for the last 3 MODISA files no matching AMSR-E files were found within the 2 day limit difference. These last 3 MODISA files were just copied to the merged data directory.

A subset of the merged MODISA/AMSR-E SST for October 30, 2006 (Julian day 303) is shown below. For better visualization the coastline was created with *Geo-Get Map Overlay* and added with *Multi – Overlay Image*.



As you can see, the merged image has much better coverage and has the resolution of the MODISA image. Data is still missing near coasts where microwave data is not working. A few spots over the open ocean are also missing - they were probably affected by rain or sun glint in AMSR-E data.

A corresponding subset from a corresponding NGSST dataset (October 30, 2006) is shown below. As you can see, the NGSST has high quality data up to the coasts but the data is missing west of 166 E and west of 116 E (lower left corner). Of course, simple merger of 2 datasets is no substitute for the optimum interpolation of multiple datasets using modeling and complex quality control procedures, however, with this simple exercise we were able to fill most of the missing pixels and produce a merged and superior dataset that is useful for many applications.