Anomalies, EOF/PC analysis with WAM

1 Introduction

Calculating anomalies is a powerful method of change detection in time series. Empirical Orthogonal Function (EOF) analysis and the related Principal Components Analysis (PCA) are another set of powerful methods to extract information from large datasets, e.g. series of satellite images. The terminology of these methods is rather confusing as similar methods are used under different names. In this exercise we will calculate anomalies, EOFs and principal components of a series of chlorophyll images. In doing so, we also generate a set of 12 monthly mean images.

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 standard mapped monthly chlorophyll images from OCTS and SeaWiFS.

3 Selecting the standard map and area of interest

In this exercise we will just use subsets of the standard 9-km SMI (standard mapped) images of Chl-a from two sensors - OCTS and SeaWiFS. You can also use the merged SeaWiFS/MODIS-Aqua data (starting from July, 2002) but that time series is shorter that is a drawback. We use OCTS as it covers the period before the major El Nino event of 1997/1998. Here we pick the Japan East Sea (JES) as our area of interest. You can choose another area and/or a different map projection.

In this exercise we will be generating hundreds of different files. It is **VERY IMPORTANT** to keep all files types **WELL ORGANIZED** in separate folders. You can use your own folder names but I recommend the following. I create a top level folder C:\sat\JES for this project. Under that folder I have subfolders ***chlo*** (for chlorophyll-a image subsets), ***chlo_anomaly*** (for the chl-a anomalies), ***chlo_means*** (for the monthly means), ***chlo_eof*** (forf the EOFs and PCs). I recommend using the same structure:

- C:\sat\JES
    - chlo
    - chlo_anomaly
   - chlo_means
    - chlo_eof

Before you start running WAM programs you need to pick your area of interest and a suitable target map projection. If you have your area of interest and target map from a previous exercise then you can use that. Here I generate a new target projection with the WIM *File-New* menu selection (left panel below):
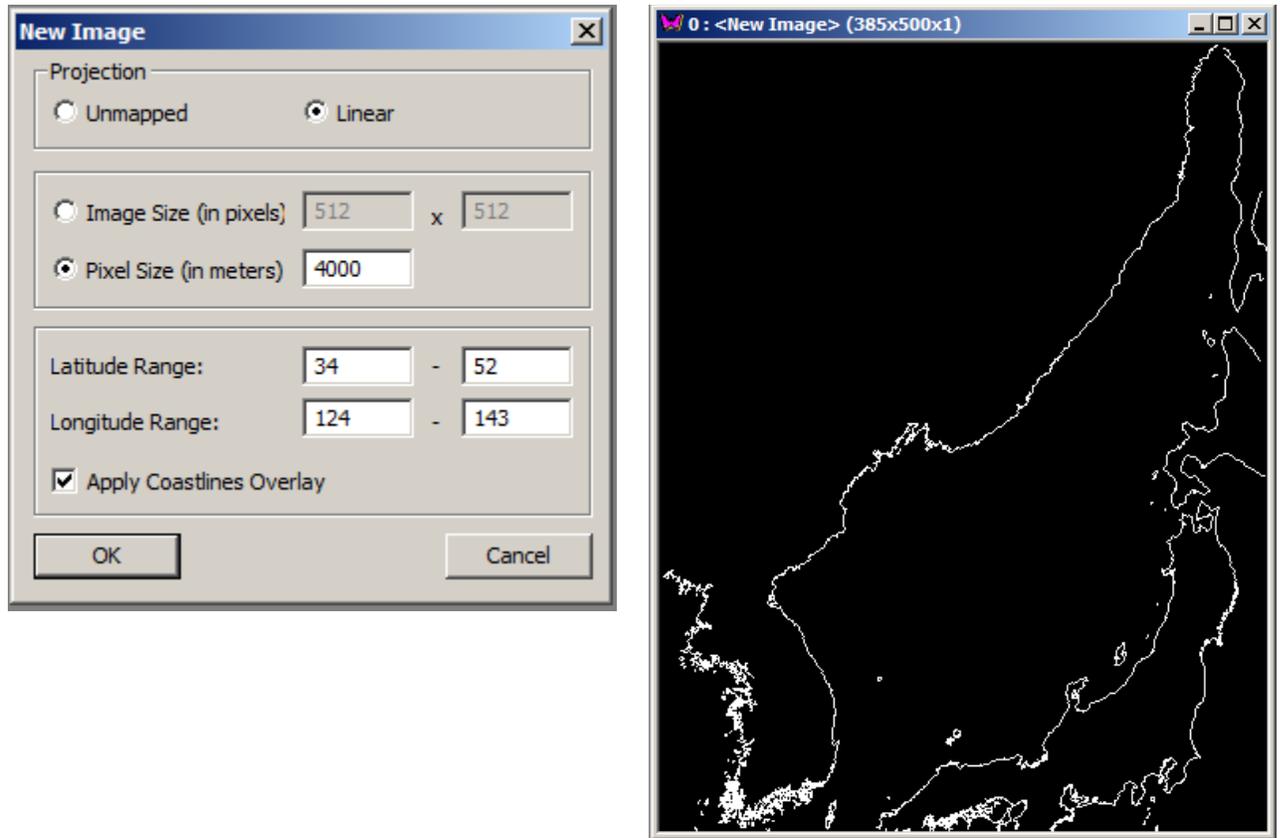
Note that I have picked latitude range of 34-52˚ and longitude range 124-143˚ (degrees). The resulting image of size 385 x 500 pixels with coastlines is shown on the right panel above. I will save this image in the *C:\sat\JES* folder as *JES_4km.hdf*. Note that I am using a *Linear* projection and pixel size of 4 km. The source images that I am going to use are of 9 km resolution and therefore they will need to be remapped to my 4 km projection. I could also generate the target map with wam_series by cutting out the same latitude-longitude range and keeping the original 9 km image resolution with *wam_series* using the following settings:
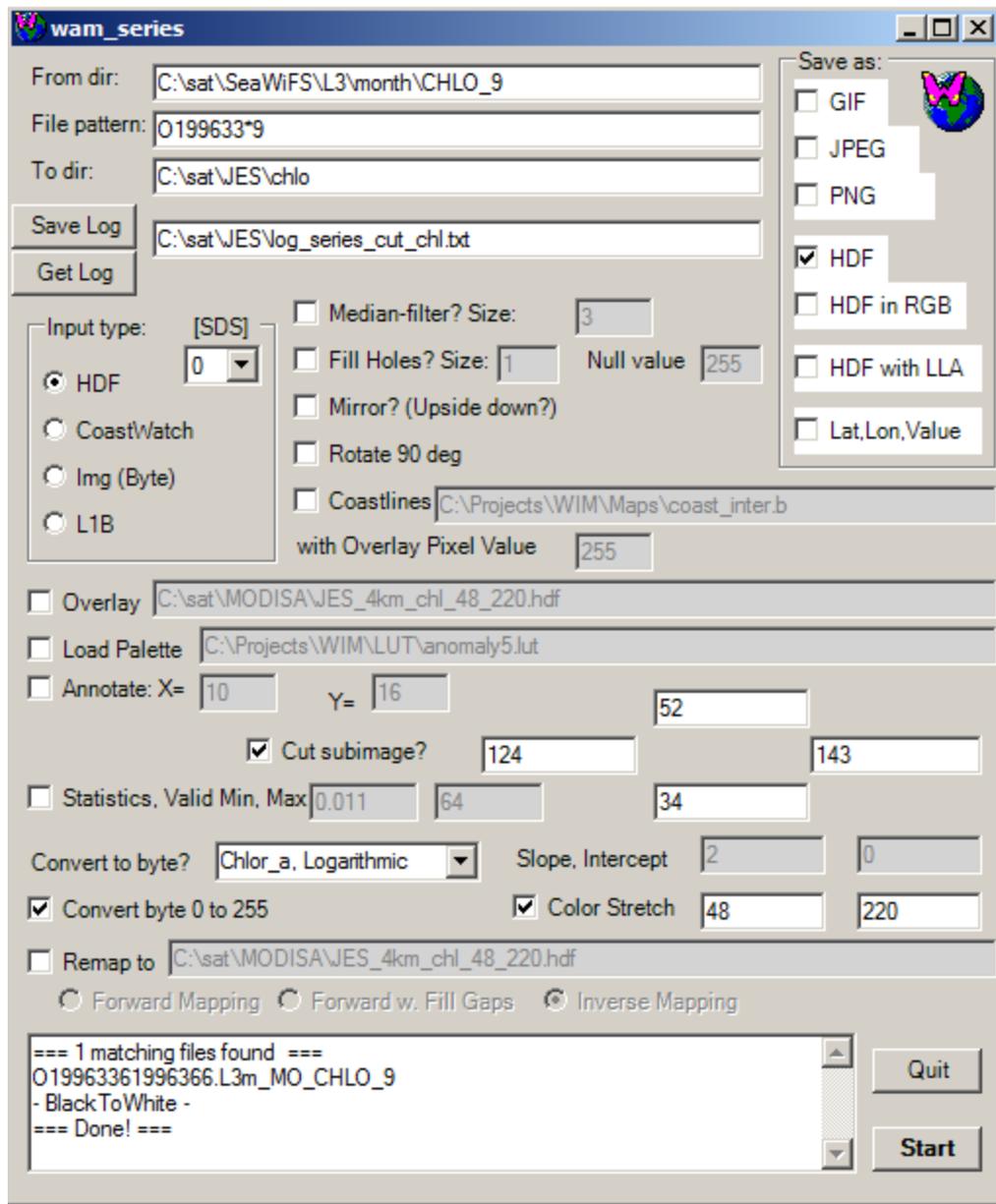
Note that I have cut out the same latitude-longitude range from a single matching OCTS chlo image, saved this as HDF, stretched the colors to 48-220 (i.e. 0.05-20 mg m^{-3} of chl-a). It is not important here but I have also changed the pixels of value 0 to 255 (black to white), and made sure that the output is in *Log-Chl* scaling (this source file was already in this scaling). In the figure below the left image is the Cut image and the coastlines image that I created for it the right is shown in the right image.
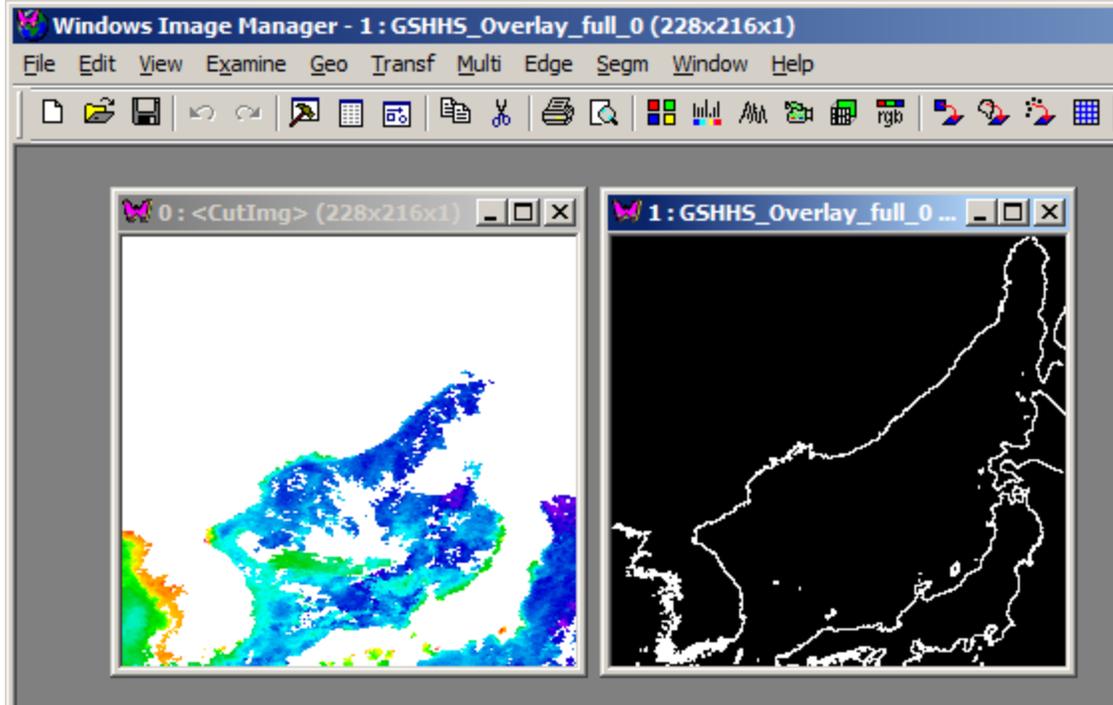
Note that the image size is 228 x 216 pixels. I will save the coastlines image in the C:\sat\JES folder as *JES_9km.hdf*. In the following analysis you can use either the 4 km or the 9 km target map. I will use the 4 km map.

The next step is to create the land mask of the target image. I will create the coastlines again with pixel value 255 with the *Geo-Get Map Overlay* command in WIM but with the *coast_full.b* file. I will then fill the gaps between coastlines and image edges manually with *Edit-Draw* and the brush tool in order to prevent the paint to leak out. I will then fill the land areas with pixel value 255 (white). If the paint leaks out I will use

the *Undo* command on the toolbar (see the arrows   ) and try to close the coastlines manually and then try again. The resulting image will look like below left and I will save it as *C:\sat\JES\JES_4km_land.hdf*.

I will now create an overlay with Chl-a color bar. To do that I will set scaling to *Log-Chl* (with the *Settings* dialog) and stretch colors to 48-220 ((i.e. 0.05-20 mg m^{-3}). I then select a rectangular area on land for the color bar and pick *Custom* tick marks 0.05, 0.2, 0.5, 1, 2, 5, 10, 20 and text color 1 (black). Make sure that there is some space above the color bar to put one line of annotation. Now create coastlines with pixel value 1 (black). Don't worry if you will not see it right away – it is because you have stretched the colors from 48 and pixel value 1 will be black on black. Overlay the coastlines (with pixel value 1) on top of the image with Chl-a color bar. The result looks like the right image below. Now save the result as *C:\sat\JES\JES_4km_chl_48_220.hdf*.
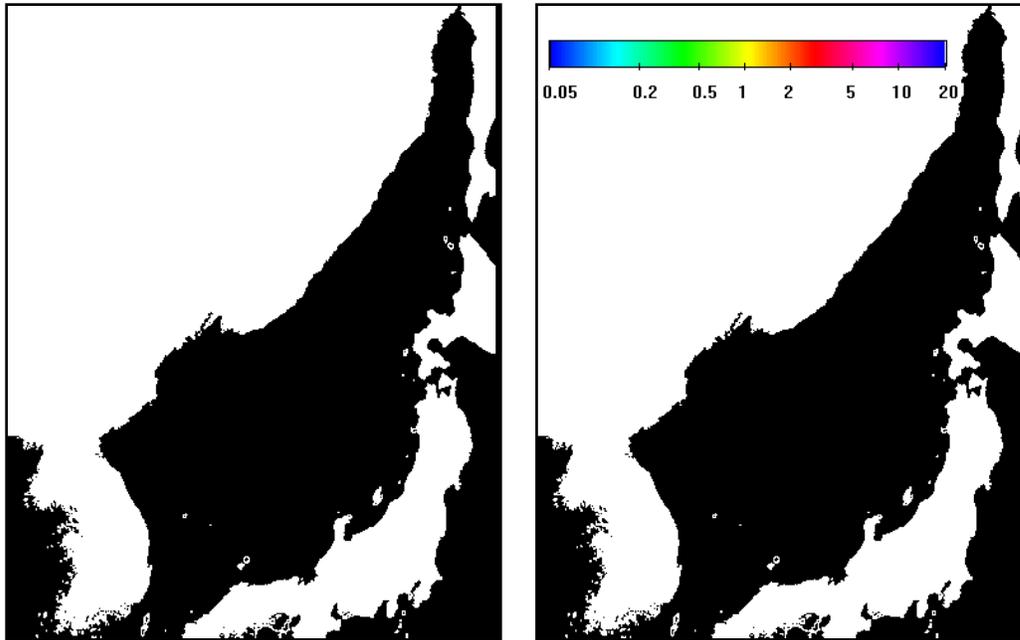
We will now create two masks: "land mask" and "sea mask". I create the land mask by using the land image that I saved as *C:\sat\JES\JES_4km_land.hdf*. I will separate the Japan East Sea (my area of interest) from the other ocean areas that I want to exclude by manually drawing on the map with *Edit-Draw* and filling the unwanted ocean areas with pixel value 255. I will then expand the land area with Edge-Dilate operation and save as *C:\sat\JES\land_mask.hdf*. In the land mask image land (and other unwanted areas) have pixel value 255 and the ocean (our area of interest) has pixel value 0. I now convert the land mask to sea mask by using *Transf-Replace Values* operation and replace values 0 to 0 with value 1 and then values 255 to 255 with value 0. The resulting mask images are shown below. Remember, land mask has land with pixel value 255 and ocean with value 0, sea mask has ocean with pixel value 1 and land with value 0. The idea of the land mask is that pixel values 0 will be used in *wam_anomaly*. The next program, *wam_eof*, on the contrary, will use mask file *sea_mask.hdf* and pixels with values different from 0. It would be possible to rewrite the *wam_anomaly* and *wam_eof* programs, so that they would both use either land mask or sea mask but for historical reasons they use different masks.

4 Make the series of subset images

To create the subset dataset we use *wam_series* to remap the standard mapped images (SMI) to our projection, and we perform several operations on the way. Remember, we can also use the *Cut* operation in *wam_series* but then we are using the same resolution and mapping as the source image. A screenshot of *wam_series* with the *Remap* option is shown below. Note that here we have matched only the December, 1999 OCTS file and done a number of operations (*Median filter* with 3 x 3 window, *Fill Holes*, etc) on the way. Note that we have also put the annotation on top of the color bar at x coordinate 10 (pixels from left) and y coordinate 20 (pixels from top).

Attention! Please start by matching only one or few files to see that the output is what you want. After you are satisfied with the output and only then – edit the matching File pattern in *wam_series* and process ALL the required images.

For example, to match all OCTS images you haves to modify the "*File pattern*" dialog and match with O*_9. To match all of the SeaWiFS images you haves to modify the "*File pattern*" dialog box and S*_9.
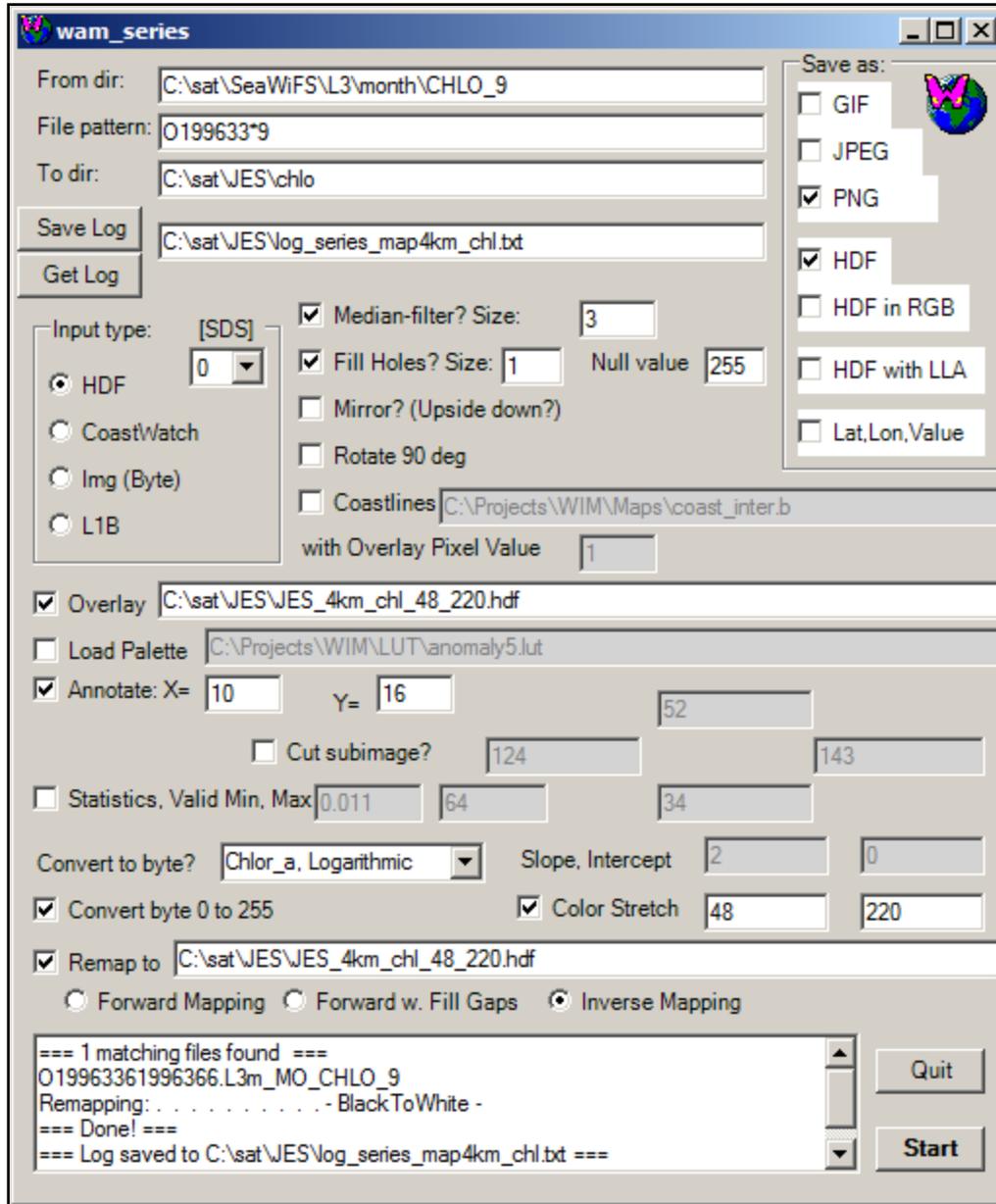
5 Run *wam_anomaly*

wam_anomaly is a command line WAM program. Open the Command Prompt and change directory to *C:\sat\JES*:

    cd C:\sat\JES

Now run *wam_anomaly* with the following arguments:

    wam_anomaly chlo*.hdf 12 anomaly5.lut false land_mask.hdf no 10 16

To understand the command syntax you can run *wam_anomaly* without arguments. In short, we use 12 to create monthly anomalies (i.e. 12 months), LUT file *anomaly5.lut*, *false* means that we will create missing data as no anomaly. We use *lands_mask.hdf* as

our land mask and put annotation at x=10 and y=16. When the command finishes, it has created *anomaly.hdf* and *anomaly.jpg* files for each input file. It also created *_Means.hdf and *_ValidCounts.hdf files. Sort the *chlo* folder by *Time* and move all the *anomaly.* files into a separate folder *chlo_anomaly* at the same level (not under *chlo*). Browse though the *anomaly* files and explain the patterns. For example, the April, 1999 (La Nina) image (below left) shows strong negative Chl-a anomaly along the continental coast.
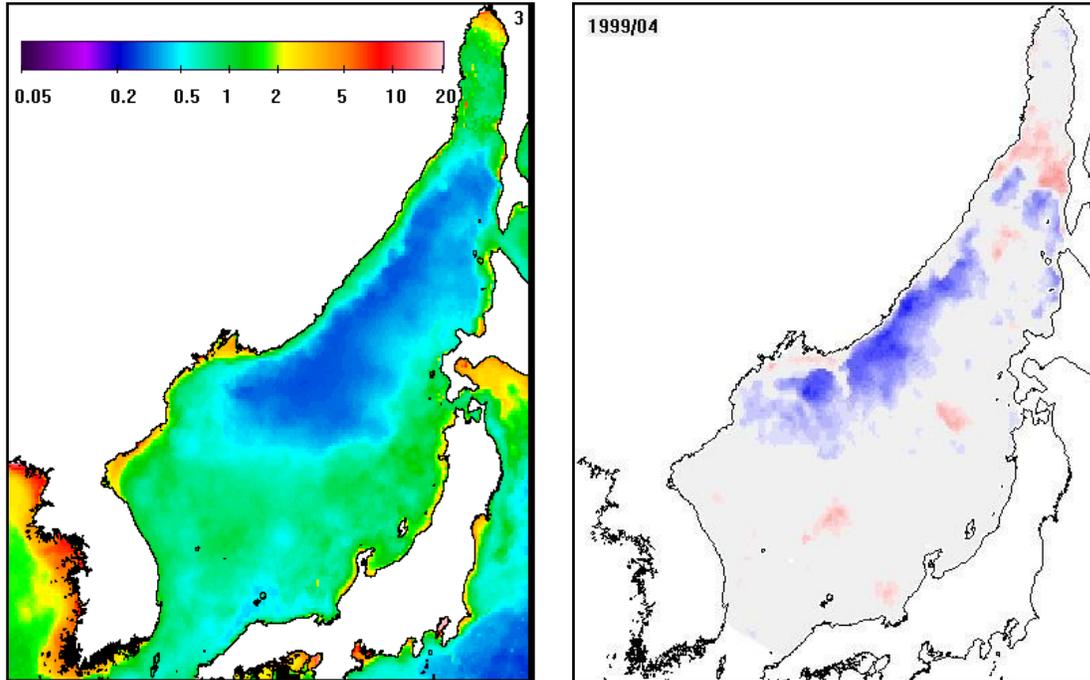
Move the *Means.hdf* and *ValidCounts.hdf* files into the *chlo_means* folder also at the same level. Rename them into *Means.hdsf* and *ValidCounts.hdf*, respectively. Open the *Means.hdsf* and *ValidCounts.hdf* files with WIM and explain their contents. We can easily generate nice annotated images with *wam_overlay*. We will use palette (LUT) file *chl1_white_end.lut* for the mean images. Of course, we could create overlays manually with WIM but we have to do that 12 times with the individual monthly images. It is much faster with *wam_overlay*.

wam_overlay chlo_means\Means.hdf JES_4km_chl_48_220.hdf chl1_white_end.lut 48 220

All this command is in one line, of course. The numbers 48 and 220 mean color stretching from pixel value 48 to pixel value 220.

If you are familiar with Babarosa GIF Animator then you can quickly generate an animated GIF from the monthly images. You can also browse through them. Notice the month number in the upper right corner. A sample March image is shown below right.

6 Run *wam_eof*

The EOF and principal component analysis is run with the following command:

    wam_eof chlo_anomaly*anomaly.hdf sea_mask.hdf

After the command finishes, you will have 3 new files\:

    PC__anomaly.hdf.hdf

    _anomaly.hdf_Eigenvalues.csv

    _anomaly.hdf_Eigenvectors.csv

Create a new folder C:\sat\JES\chlo_eof and move those 3 files there. It is convenient to annotate all principal component images with the *wam_overlay* command:

    wam_overlay chlo_eof\PC__anomaly.hdf.hdf JES_4km_land.hdf anomaly5.lut

All the PCs have linear scaling and are stretched from min to max. The exact numerical values are not important as they can be multiplied by any constant, e.g. -1 in order to mirror it. The first two principal components are shown below. The red areas have positive and the blue areas negative values.
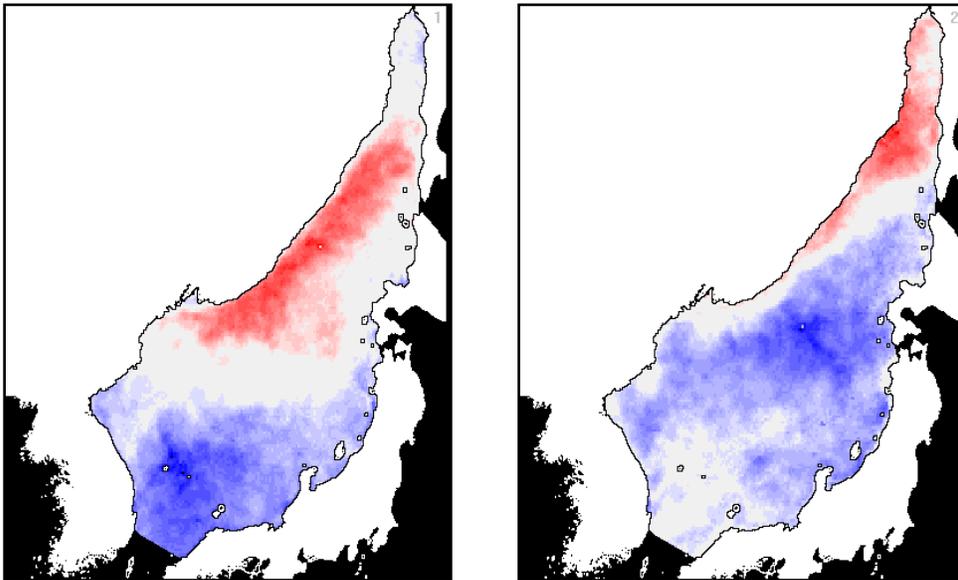

Now look at the *Eigenvalues.csv* file. The top part of the file is shown below.

#	Mean	SD	EVal	EVal%	EvalCum%
1	124.606	9.236	13.072	10.5	10.5
2	122.359	9.596	7.752	6.2	16.7
3	118.18	16.516	6.43	5.1	21.8
4	107.799	16.525	5.251	4.2	26
5	113.312	18.073	4.363	3.5	29.5
6	127.053	11.707	4.131	3.3	32.8

The numbers here mean that the 1st principal component explains 10.5% of total variance; the first 6 components together describe nearly 1/3 (32.8%) of the total variance. This shows that the spatio-temporal distribution of chl-a is quite complex and no single factor can explain more than 10.5% of the variance. Even using all the best 6 components we still have 2/3 of the variance unexplained.

The *Eigenvectors.csv* file shows the temporal evolution (strength or score) of the different components. They are called EOFs. Make a new decimal year time axis and plot some of the EOFs in time. If you can find some ENSO indices then you may find that some of the EOFs are correlated with ENSO.

You can compare your analysis with similar analysis of the California Current at http://www.wimsoft.com/manual.htm and at http://spg.ucsd.edu/Satellite_Projects/NPac_variability/Npac_variability.htm.
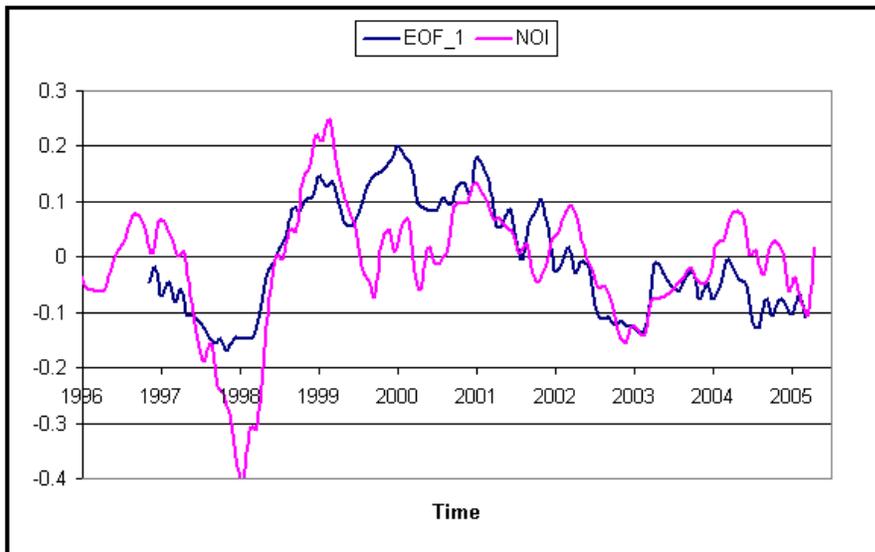
The plot above shows that in the California Current the spatial structure of the 1st principal component is similar to the anomaly structure during the peak of the El Niño in January, 1998 (Kahru and Mitchell, 2000). The temporal evolution of the first principal component (EOF1) shows close resemblance any of the El Niño-ENSO indices, e.g. the Extratropical Northern Oscillation index (NOI, Schwing et al., 2002). Note the strong El Nino event of 1997-1998, and the weak events of 2003 and 2005 corresponding to the minima.

References

Kahru, M., B.G. Mitchell (2000), Influence of the 1997-98 El Niño on the surface chlorophyll in the California Current, *Geophysical Research Letters*, 27, 2937-2940.

Schwing, F.B., T. Murphree, P.M. Green (202), The Northern Oscillation Index (NOI): a new climate index for the northeast Pacific, *Progress in Oceanography*, 53, 115-139.