**Exercise 8 - Visualizing and analyzing MODIS data**

Prerequisite

* Python: <http://www.python.org/doc/>
* GDAL: <http://www.gdal.org/>
* Geospatial Data Abstraction Library
* Raster data access
* Used by commercial software like ArcGIS
* C++ library, but Python bindings exist
* Matplotlib: <http://matplotlib.org/>

Note: Outputs (e.g., figure, statistics, metainfo) might be different by input

We will focus at trying to determine the effects of the 2003 drought in Europe on the distribution of LAI values over Europe by comparing the 2002 season (relatively wet) and the 2003 season (extremely dry and warm). In this exercise you will use some of the advanced visualization capabilities of Python/GDAL/matplotlib.

1. **Computing LAI statistics over Europe**

In this step we will first compare the overall distribution of LAI values over Europe by generating and comparing the histograms of the LAI values in 2002 and 2003. Carry out the following steps:

* Load the 2002 final LAI values (MOD15A2H.A2002225.h18v04\_h18v03\_mosaic\_Lai.tif) as well as the masks that we created for both datasets. We need the mask for statistics calculation in order to avoid that the no-data values influence the statistics.
* Convert both the LAI and mask layers as Numpy array.
* Generate Numpy mask array for the statistics calculation.
* Apply the mask on 2002 LAI layer to remove all the pixels that not wanted.
* Print out statistics

import gdal

import matplotlib.pyplot as plt

import numpy as np

import os

## set your work directory

os.chdir('/Users/Shared/GE529-2020/data/')

# open 2002 dataset

lai2002 = gdal.Open('lai\_mosaic\_2002.tif')

laiArray2002 = lai2002.ReadAsArray()

qc2002 = gdal.Open('qc\_mosaic\_2002.tif')

qcArray2002 = qc2002.ReadAsArray()

qcArray2002 = qcArray2002 & 224

qcArray2002 = qcArray2002 >>5

np.unique(qcArray2002)

mask2002 = qcArray2002==0

laiArray2002 = laiArray2002 \* 0.1

laim2002 = np.ma.array(laiArray2002, mask=qcArray2002)

print("[ STATS ] = Minimum={0}, Maximum={1}, Mean={2}, StdDev={3}".format(laim2002.min(),laim2002.max(),laim2002.mean(),laim2002.std()))

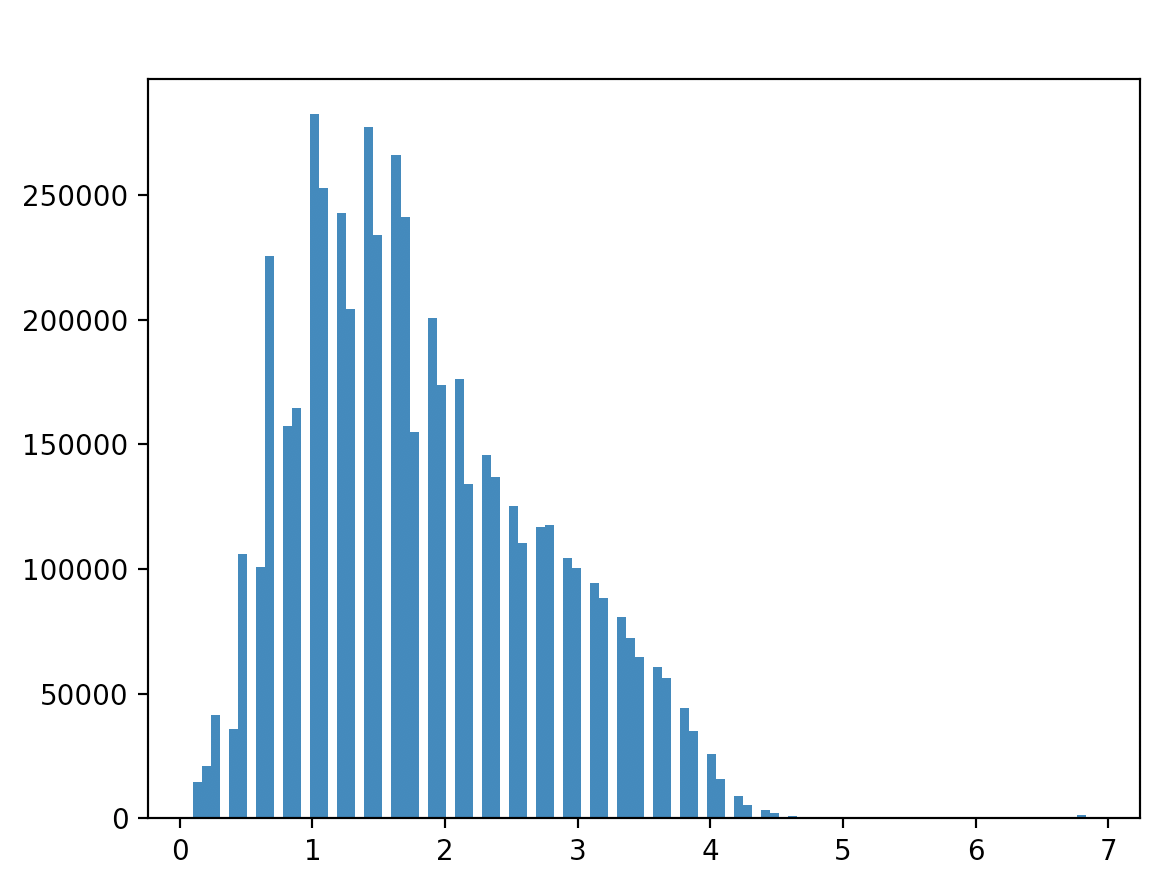
* Print out the basic statistics of the LAI in 2002.
* Calculate the statistics for the LAI of 2003 in the same way.
* Plot the histogram of LAI 2002

# compute 2002 histogram

laim2002 = laim2002.compressed()

plt.hist(laim2002, bins=100)

plt.show()



**Figure 1.** Histogram for LAI 2002

1. **Compare and visualize the LAI statistics**

* Print out the statistical parameter of the LAI in 2002 & 2003. What do you conclude from the mean LAI value?
* Now look at the distributions, this can be down more easily by having the two distributions in the same plot window. We will use matplotlib putting two

# open 2003 dataset

lai2003 = gdal.Open('lai\_mosaic\_2003.tif')

laiArray2003 = lai2003.ReadAsArray()

qc2003 = gdal.Open('qc\_mosaic\_2003.tif')

qcArray2003 = qc2003.ReadAsArray()

qcArray2003 = qcArray2003 & 224

qcArray2003 = qcArray2003 >>5

np.unique(qcArray2003)

mask2003 = qcArray2003==0

laiArray2003 = laiArray2003 \* 0.1

laim2003 = np.ma.array(laiArray2003, mask=qcArray2003)

# compute histogram

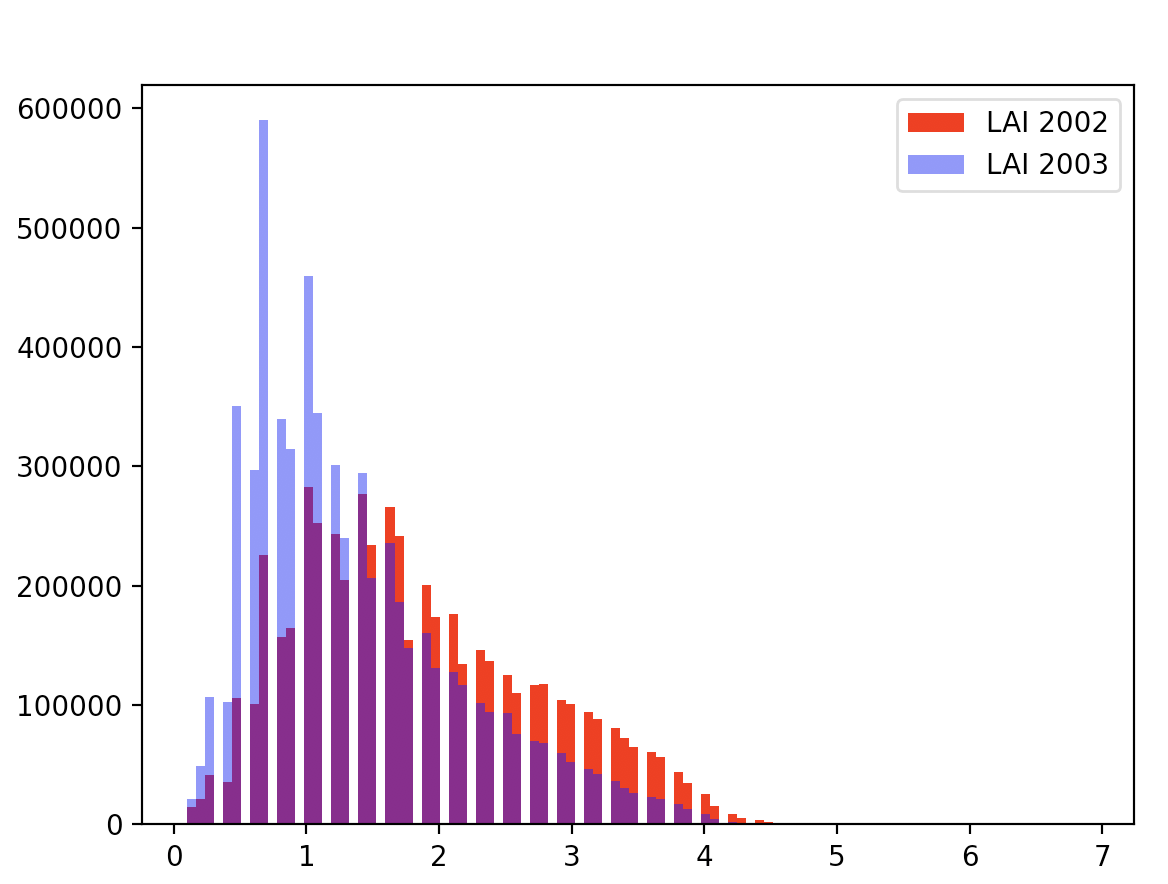
laim2003 = laim2003.compressed()

plt.hist(laim2002, bins=100, color='r', label='LAI 2002')

plt.hist(laim2003, bins=100, color='b', alpha=0.5, label='LAI 2003')

plt.legend()

plt.show()



**Figure 2.** Histogram comparison for LAI 2002 & 2003

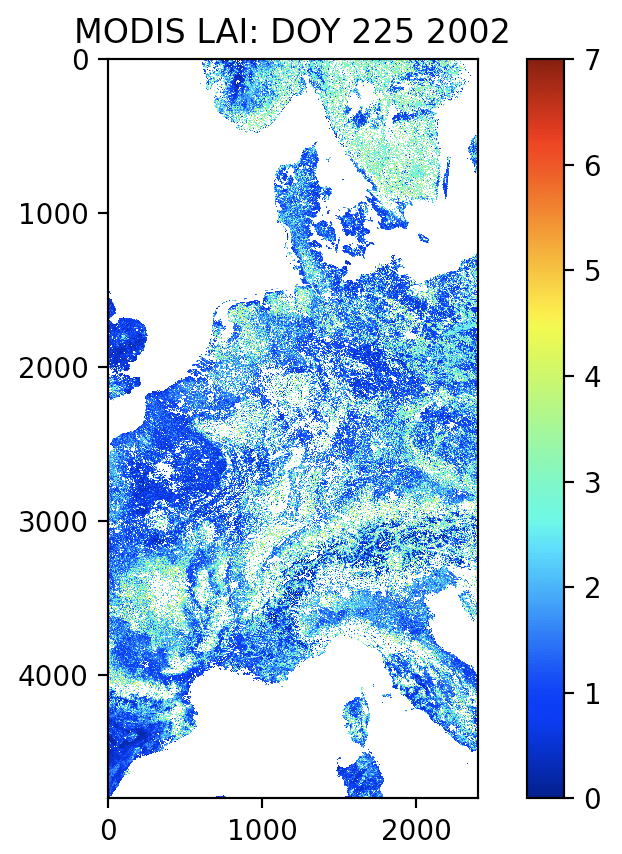
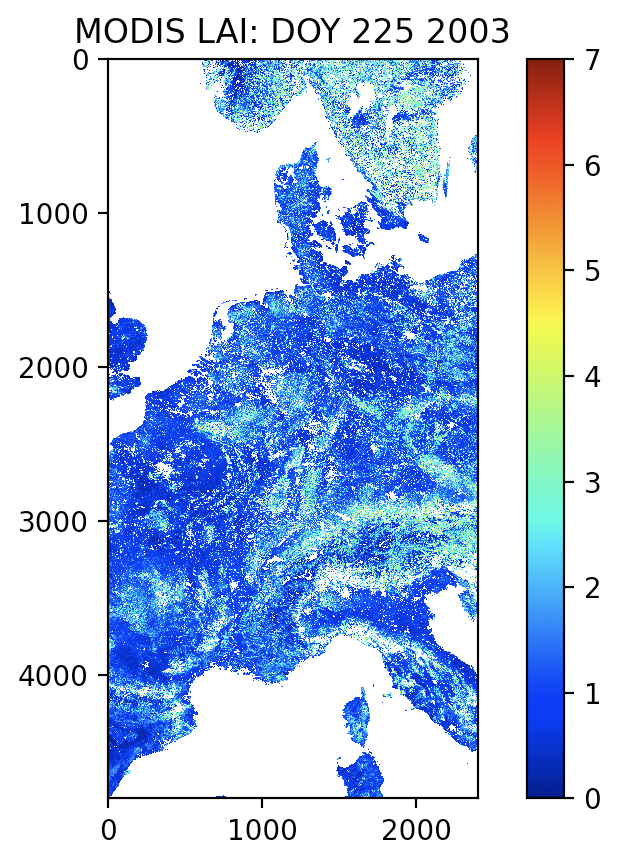
Question:

* What do you conclude about the LAI 2002 & 2003 from the main distribution on the left side of the plot window? What about the small peak on the right side?

1. **Regional inspection of the drought influence on LAI**

By looking at the image statistics and the distribution of the LAI values we get a feel for the overall situation in Europe. However, it does not give spatial information in order to locate the areas have suffered badly from drought. Therefore, we have to take a look at the image data itself.

* Open the final LAI datasets for 2002 and 2003 and load them using GDAL.
* Make plot for Europe LAI 2002 &2003.

**** ****

**Figure 3.** Compare Europe LAI 2002 & 2003

1. **Showing drought and heat affected pixels**

For the current tutorial we assume that all areas that are seriously affected by the 2003 drought and heat, are characterized by a 25% smaller LAI in 2003 compared to 2002. The boundary of 25% deviation can be visualized by masking out all the LAI pixels with less than 25% deviation. We use Numpy equation to apply this purpose.

We use equation below to remove pixels unaffected by the 2003 drought.

mask=laim2003>laim2002\*0.75

All pixels that fall within this drought category are affected by drought, according to our definition.

Drought affected pixels can be plotted easily by matplotlib from Python.

# show drought affected pixels

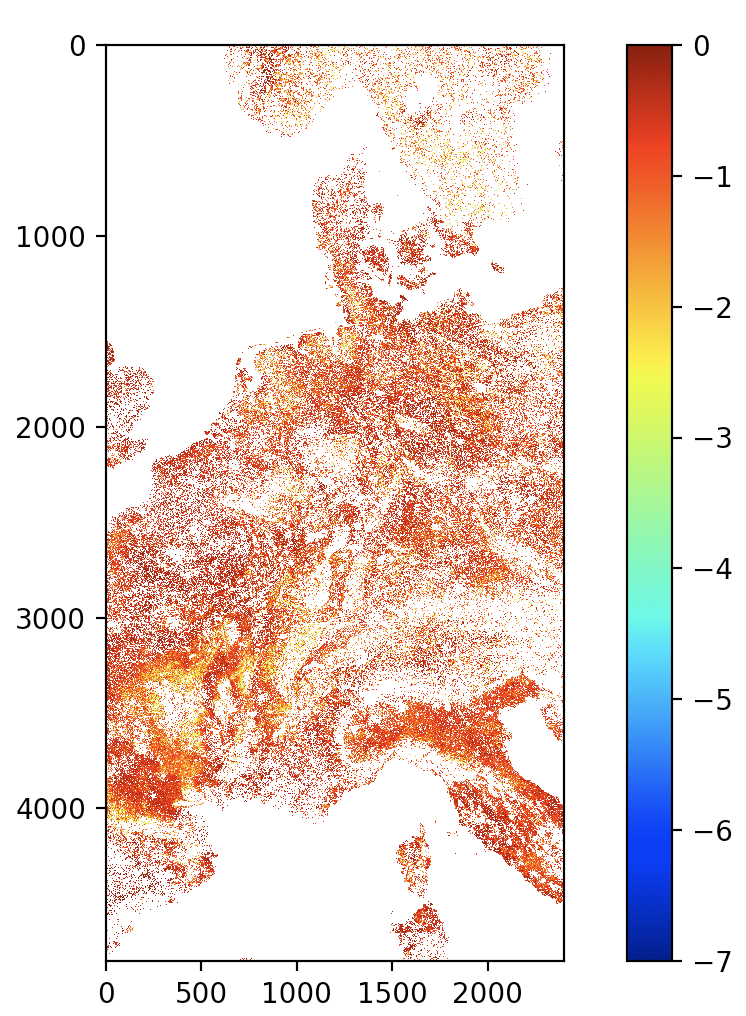
drought=np.ma.array((laim2003-laim2002), mask=laim2003>laim2002\*0.75)

cmap=plt.cm.jet

plt.imshow(drought,cmap=cmap,interpolation='nearest',vmin=-7, vmax=0)

plt.colorbar()

plt.show()



**Figure 4.** Drought affected areas

**End of Exercise 8**