

实验 1 水色遥感数据的获取与可视化

编写人：陈玥舟（报告撰写与代码编写）、蒋浩宇（内容设计与完善）

一、实验目的

1. 了解如何获取海洋数据；
2. 掌握用 Python 对 NetCDF (.nc) 格式数据的信息查询与读写；
3. 掌握用 Python 对网格化数据的进行可视化的基本操作；

二、实验任务

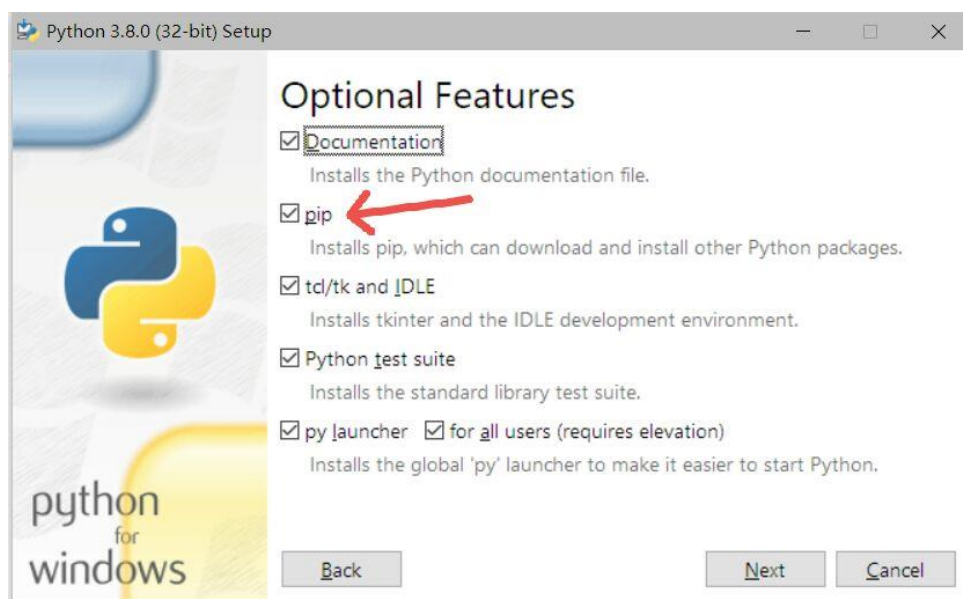
1. 在 NASA 的 Ocean Color 网站 <https://oceancolor.gsfc.nasa.gov/> 下载 MODIS-Aqua 的 chlor_a (叶绿素 a) 浓度某个月的数据 (9km 分辨率)；
2. 用 Python 或其它 nc 查看软件(如 HDFView)对下载数据的信息进行查询；
3. 用 Python 或其它数据可视化软件 (如 Surfer) 对数据进行可视化。

三、实验步骤

1. Python 环境搭建

搭建 Python 环境的常见模式是安装一个 Python 的 IDE (集成开发环境), 再使用 pip 或 conda 等命令安装特别扩展的第三方包, 常见的 Python IDE/平台有 Python 自带的 IDE (IDLE)、Anaconda、Pycharm、VScode 和 Jupyter Notebook 等。其中, Jupyter Notebook 是一个基于 Web 网页形式的交互式笔记本应用, 能可视化展示并保存运行结果。本报告将以 Jupyter Notebook 作为基本编程工具进行实验。

(1) 在 Python 官方网址 www.python.org/downloads/ 下载相应的 Python3.x 版本 (建议下载 3.6 以上版本) (Python 2.7.9 + 或 Python 3.4+ 以上版本, 如下图可在安装时勾选 pip 选项完成配置。Python3.x 版本可以通过 cmd 下输入命令 `pip3 --version` 来判断是否已安装。如未安装, 则可打开网址 <https://bootstrap.pypa.io/get-pip.py> 下载安装相对应脚本, 接着 cmd 下输入命令 `python3 get-pip.py` 运行安装脚本)；



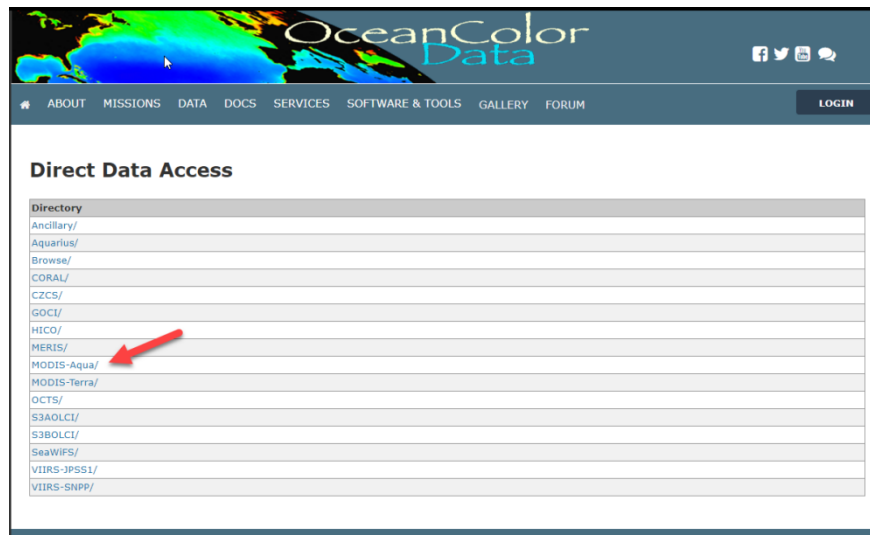
(2) 打开 cmd，输入命令：`pip install jupyter notebook` 安装 Jupyter Notebook；

(3) 安装一个 Python 的专门处理 nc 文件的库：`netCDF4`。这个库可以通过打开 cmd，直接输入命令 `pip install netCDF4` 来安装。（除了 `netCDF4` 这个库以外，还会用到其他库，如 `numpy`、`matplotlib`，同样可以通过在 cmd 输入命令 `pip install “ ”` 需要安装的库名”来进行安装）

(4) 安装 Basemap 工具包。这是 python 可视化库 Matplotlib 下的一个工具包，主要功能是绘制二维地图，对空间数据的可视化非常重要。打开网址“<http://www.lfd.uci.edu/~gohlke/pythonlibs/>”，根据自己安装的 python 版本手动下载相应的 .Whl 文件（可在页面上按 Ctrl+F，输入 basemap 快速定位），如“`basemap-1.2.2-cp36-cp36m-win_amd64.whl`”；把文件下载到一个路径下，如“`C:\Users\famil\basemap-1.2.2-cp36-cp36m-win_amd64.whl`”；在 cmd 下定位到文件所在文件夹，然后输入命令“`pip install ‘.whl 文件名’`”完成安装，如“`pip install basemap-1.2.2-cp36-cp36m-win_amd64.whl`”。

2. 下载数据

- 1、登录 <https://oceancolor.gsfc.nasa.gov/> 网站，浏览网站信息。
- 2、找到数据获取页面，按照数据目录找到需要的数据。



The screenshot shows the OceanColor Data website with the breadcrumb trail: OceanData Home > MODIS-Aqua > Mapped > Monthly > 9km > chlor_a. Below the breadcrumb is a table listing files.

Filename	Last Modified	Size
A20021822002212.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:05:28	12982521
A20022132002243.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:05:27	14418192
A20022442002273.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:08:51	15654483
A20022742002304.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:07:34	15294212
A20023052002334.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:07:16	14660174
A20023352002365.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:06:46	14566814
A20030012003031.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:08:19	15258390
A20030322003059.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:07:42	15777362
A20030602003090.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:09:09	16023085
A20030912003120.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:08:52	14627629
A20031212003151.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:08:06	13078859
A20031522003181.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:05:43	12609562
A20031822003212.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:04:55	13227413
A20032132003243.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:07:36	14694419
A20032442003273.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:08:46	15521435
A20032742003304.L3m_MO_CHL_chlor_a_9km.nc	2019-12-20 06:46:42	15475834
A20033052003334.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:08:08	14690678
A20033352003365.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:06:40	14606190
A20040012004031.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:08:18	15268882
A20040322004060.L3m_MO_CHL_chlor_a_9km.nc	2018-01-03 18:05:37	15912056

下载文件可用本账号登录：

账号：yz0325 密码：Yz0325..

也可以自行注册账号。

3. 读取数据信息

在 cmd 下输入命令：jupyter notebook，打开 Jupyter Notebook 编辑器。

Jupyter 窗口输入以下内容：

```
import netCDF4 as nc          #导入库
#使用 netCDF4 读取 nc 文件时，所有路径要用英文命名，路径中的“\”要改成“\\”
file = nc.Dataset('A20171212017151.L3m_MO_CHL_chlor_a_9km.nc')
print(file)                  #查看 nc 文件信息
```

可以看到数据的位置、类型、全局属性、变量以及文件夹等信息，部分如下图所示：

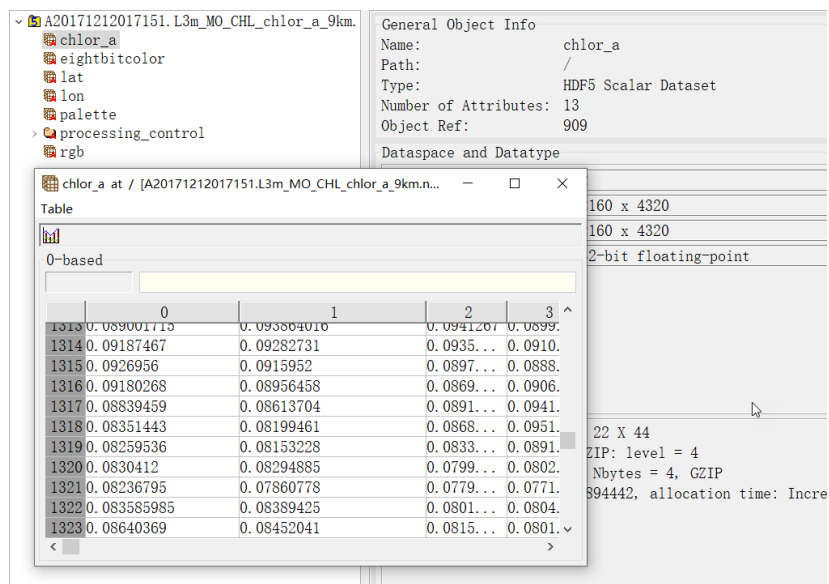
```

<class 'netCDF4._netCDF4.Dataset'>
root group (NETCDF4 data model, file format HDF5):
  product_name: A20171212017151.L3m_MO_CHL_chlor_a_9km.nc
  instrument: MODIS
  title: HMODISA Level-3 Standard Mapped Image
  project: Ocean Biology Processing Group (NASA/GSFC/OBPG)
  platform: Aqua
  temporal_range: month
  processing_version: 2018.0
  date_created: 2018-01-03T23:00:34.000Z
  history: l3mapgen par=A20171212017151.L3m_MO_CHL_chlor_a_9km.nc.param
  l2_flag_names: ATMFALL, LAND, HILT, HISATZEN, STRAYLIGHT, CLDICE, COCCOLITH, LOWLW, CHLWARN, CHLFAIL, NAVWARN, MAXAERITER, ATMWARN, HISOLZEN, NAVFAIL,
  FILTER, HIGLINT
  time_coverage_start: 2017-05-01T01:00:00.000Z
  time_coverage_end: 2017-06-01T02:29:59.000Z
  start_orbit_number: 79741
  end_orbit_number: 80194
  map_projection: Equidistant Cylindrical
  latitude_units: degrees_north
  longitude_units: degrees_east

data_bins: 3855363
data_minimum: 0.015458127
data_maximum: 97.845245
dimensions(sizes): lat(2160), lon(4320), rgb(3), eightbitcolor(256)
variables(dimensions): float32 chlor_a(lat, lon), float32 lat(lat), float32 lon(lon), uint8 palette(rgb, eightbitcolor)
groups: processing_control

```

也可以利用 HDFView 对数据进行查看。



4. 数据读取

结合数据信息，将经纬度和叶绿素浓度信息读入数组中。

```

chlor = (file.variables['chlor_a'][:]) #叶绿素浓度
lons = (file.variables['lon'][:]) #经度
lats = (file.variables['lat'][:]) #纬度

```

5. 数据可视化

在 python 进行地图绘制方法众多，basemap 模块包是一个很常用的插件。

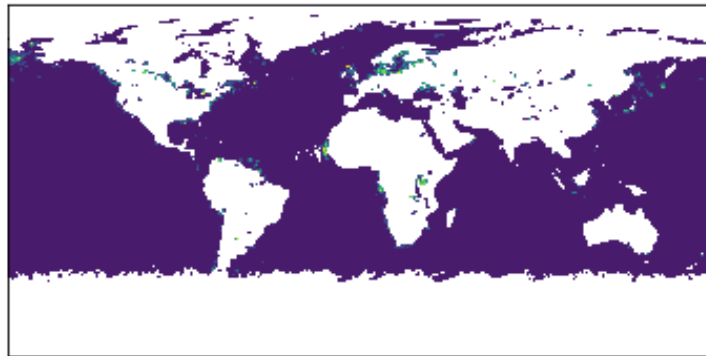
在绘制地图之前，常常需要对经度、纬度的网格进行编织。

以下是绘图代码：

```

import numpy as np
from mpl_toolkits.basemap import Basemap
import matplotlib.pyplot as plt
#对经纬度进行编织
lon, lat = np.meshgrid(lons, lats)
#画图
m = Basemap(llcrnrlon=-180,llcrnrlat=-80,urcrnrlon=180,
             urcrnrlat=80,projection='mill',resolution='l')
m.contourf(lon,lat,chlor) #绘制浓度分布
plt.show()

```



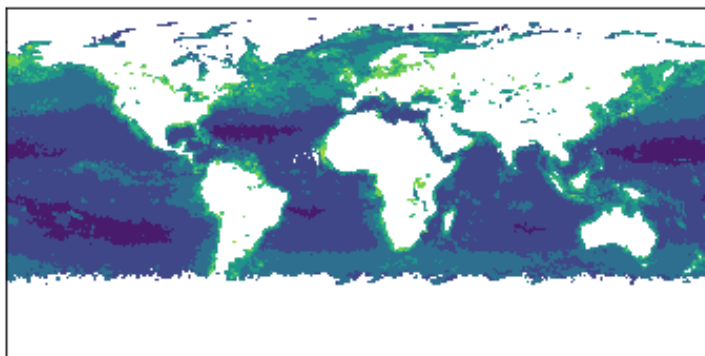
叶绿素高浓度区域和低浓度区域相差几十倍，所以直接对结果进行可视化往往差别不明显（仔细观察也能看出图中部分差异）。为解决这一问题，通常在绘制叶绿素浓度的时候，常常对进行取对数操作。

取对数以后再用上述代码绘制结果如下：

```

#对叶绿素浓度取对数
chl = np.log(chlor)
#画图
m.contourf(lon,lat,chl) #绘制浓度分布
plt.show()

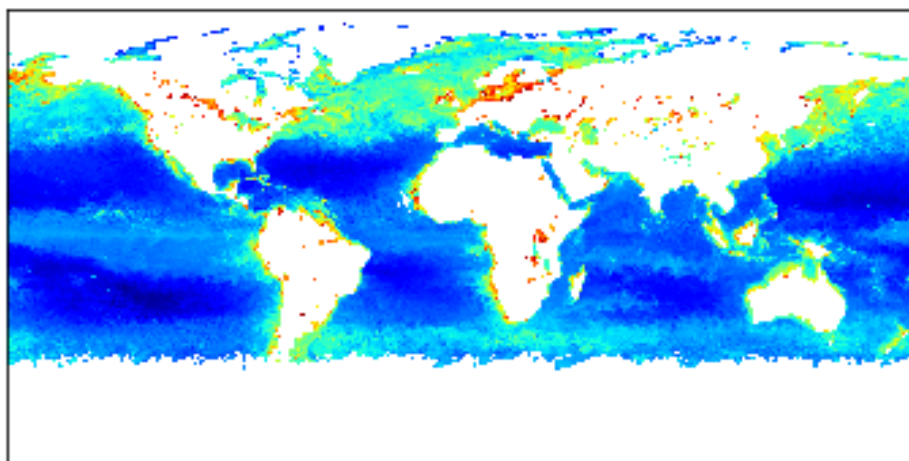
```



更换颜色搭配可以用 `cmap` 参数。此外，对高分辨率的数据，用 `pcolor` 比用 `contourf` 更合适一些。

```
m.pcolor(lon,lat,ch1,cmap='jet') #绘制浓度分布，常用的颜色映射有 jet、hsv、gray  
plt.show()
```

更换后的结果如下：

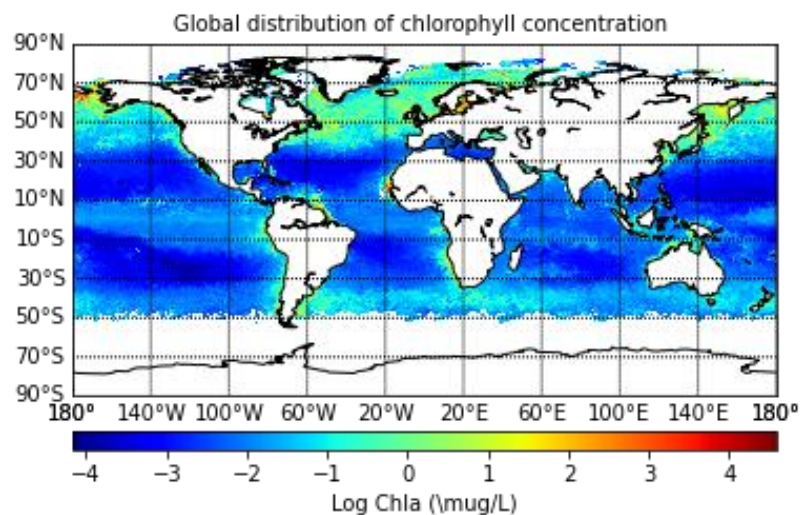


添加其它制图要素（网格线、colorbar、标题等）：

```

m = Basemap(llcrnrlon=-180,llcrnrlat=-80,urcrnrlon=180,
            urcrnrlat=80,projection='mill',resolution='l')
#为地图添加上海岸线和经纬网格
m.drawparallels(np.arange(-90., 91., 20.), labels = [1,0,0,0], fontsize = 10)
#纬线
m.drawmeridians(np.arange(-180., 181., 40.), labels = [0,0,0,1], fontsize = 10)
#经线
m.drawcoastlines() #海岸线
m.fillcontinents(color = 'white') #为陆地填上白色
cs = m.pcolor(lon,lat,ch1,cmap='jet') #绘制浓度分布
cbar = m.colorbar(cs, location = 'bottom',label = 'Log Chla (\mug/L)', pad =
"10%") #显示颜色图例
plt.title('Global distribution of chlorophyll concentration ', fontsize = '10')
#添加标题
plt.show()

```

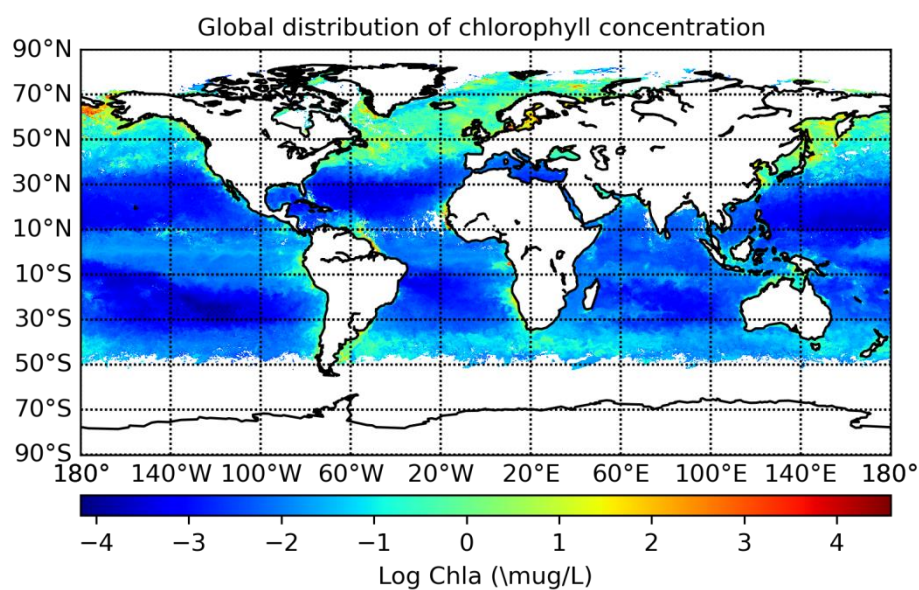


最后，对图片进行存储时，用高 dpi 的结果存储，可以形成更清晰的图像。

```

#重画一遍图片
plt.savefig("C:\\Users\\DELL\\Desktop\\1.1.png", dpi = 300, bbox_inches =
'tight') #保存图片,去掉图片周围空白
plt.show()

```

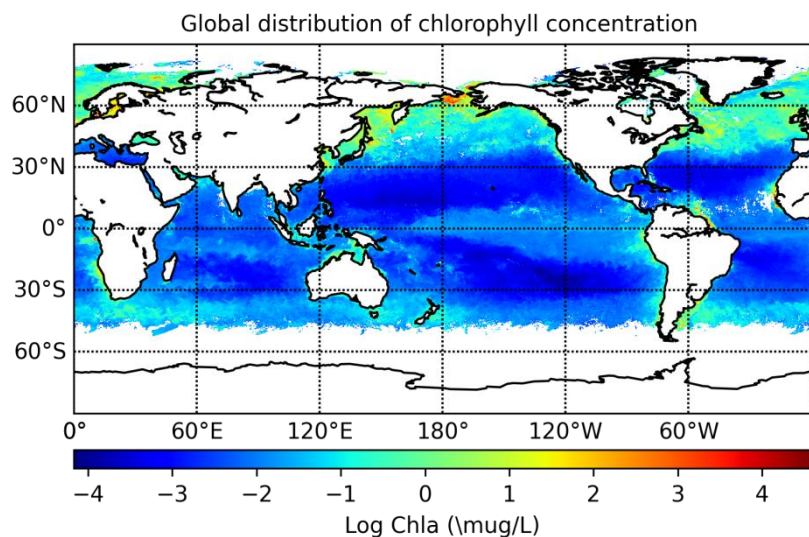
四、课堂/课后作业：

- 1、完成所有报告上的练习，理解每一句代码的含义。
- 2、用 Matlab 或其它 nc 查看软件对所下载 chlor_a 数据的信息进行查询，回答如下问题：

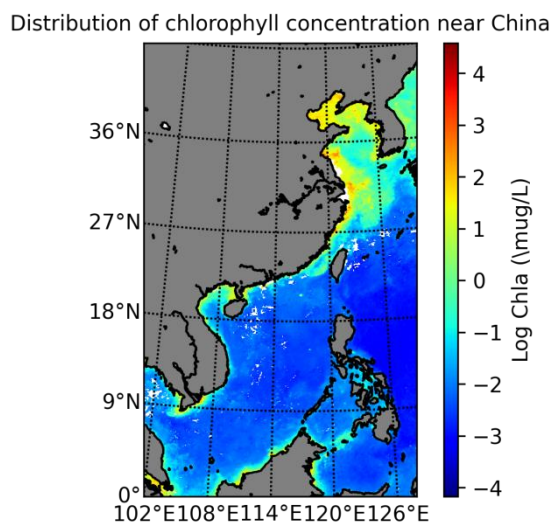
- (1) 文件的生成日期是什么时候？
- (2) 文件数据对应的具体时间段是什么？
- (3) chlor_a 变量在此文件中的全称是什么？
- (4) 文件中 chlor_a、经度、纬度数据所采用的单位分别是什么？

3、完成以下作图：

- (1) 注意到我们上面所有的作图都是以大西洋 0° 经线为中心的，这是因为我们在选择投影方式和坐标的过程中经度的范围是：[-180 180]。如何得到如下图所示以太平洋 180° 经线为中心的全平面投影图(miller 投影('mill'))？（注意与上图的所有不同之处，纬度范围、投影方式、colorbar 位置、底图颜色等，尽量做到一致）



(2) 采用 lambert 投影 ('lcc'), 作出类似下图的中国海区域的叶绿素浓度分布。



报告提交内容: (1) 数据、(2) 所有的 .m/.py 的完整程序源代码、
(3) 输出的图片、(4) 报告正文, 打包放在一个文件夹中。要求运行代码即可完成结果的输出。

附录：

5.

(1)

```
import numpy as np
import netCDF4 as nc
import matplotlib.pyplot as plt
from mpl_toolkits.basemap import Basemap

file=nc.Dataset('A20171212017151.L3m_MO_CHL_chlor_a_9km.nc')
lons=(file.variables['lon'][:])
lats=(file.variables['lat'][:])
chlor=(file.variables['chlor_a'][:])
lons=np.concatenate((lons[2160:], lons[:2160]+360)) #将数据前一半和数据后一半互换
chlor=np.concatenate((chlor[:, 2160:], chlor[:, :2160]), axis=1)
lon,lat=np.meshgrid(lons,lats)
chl = np.log(chlor)

m= Basemap(lat_0=0, lon_0=180)
m.drawparallels(np.arange(-60., 61., 30.), labels=[1,0,0,0], fontsize=10)
m.drawmeridians(np.arange(-180., 181., 60.), labels=[0,0,0,1], fontsize=10)
m.drawcoastlines()
m.fillcontinents(color='white')
cs=m.pcolor(lon,lat,chl,cmap='jet')
cbar = m.colorbar(cs, location='bottom', pad="10%",label='Log Chla (\mug/L)')
plt.title('Global distribution of chlorophyll concentration ',FontSize='10') #添加标题
plt.show()
```

(2)

```
import numpy as np
import netCDF4 as nc
from mpl_toolkits.basemap import Basemap
import matplotlib.pyplot as plt
file=nc.Dataset('A20171212017151.L3m_MO_CHL_chlor_a_9km.nc')
chlor=(file.variables['chlor_a'][:])
lons=(file.variables['lon'][:])
lats=(file.variables['lat'][:])
lon,lat=np.meshgrid(lons,lats)
chl = np.log(chlor)
m = Basemap(projection='lcc',lat_0 = 22., lon_0 =
117.,llcrnrlat=0,urcnrlat=45,llcrnrlon=102,urcnrlon=132,resolution='i')
m.drawparallels(np.arange(-90., 91., 9.), labels=[1,0,0,0], fontsize=10) #为地图添加经纬线
```

```

m.drawmeridians(np.arange(-180., 181., 6.), labels=[0,0,0,1], fontsize=10)
m.drawcoastlines()
m.fillcontinents(color='gray')
x,y=m(lon,lat)
cs=m.pcolor(x,y,chl,cmap='jet')
cbar = m.colorbar(cs, location='right', pad="10%",label='Log Chla (\mug/L)') #显示颜色图例
plt.title('Distribution of chlorophyll concentration near China',Fontsize='10') #添加标题
plt.show()

```

补充：

若想让浓度分布在图中对比更鲜明，可缩小绘制范围

(1)

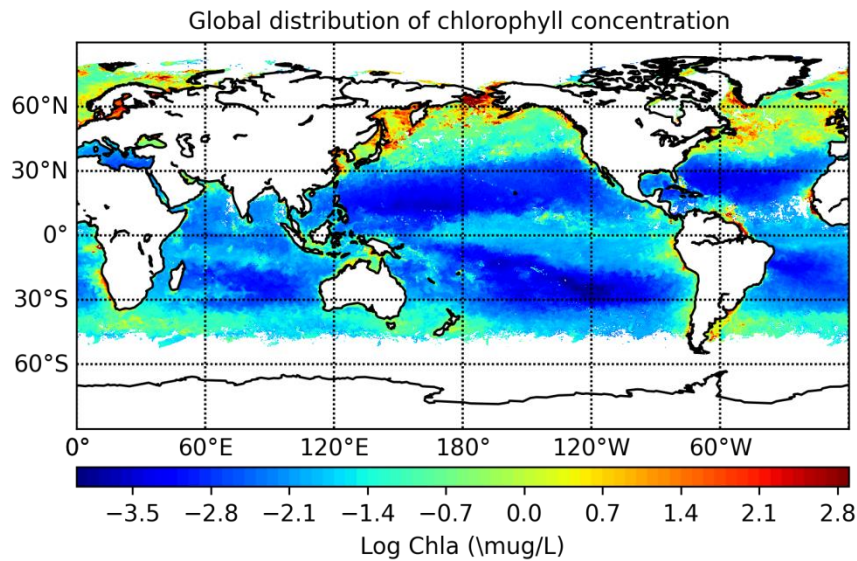
```

import numpy as np
import netCDF4 as nc
import matplotlib.pyplot as plt
from mpl_toolkits.basemap import Basemap

file=nc.Dataset('A20171212017151.L3m_MO_CHL_chlor_a_9km.nc')
lons=(file.variables['lon'][:])
lats=(file.variables['lat'][:])
chl=(file.variables['chlor_a'][:])
lons=np.concatenate((lons[2160:], lons[:2160]+360)) #将数据前一半和数据后一半互换
chl=np.concatenate((chl[:, 2160:], chl[:, :2160]), axis=1)
lon,lat=np.meshgrid(lons,lats)
chl = np.log(chlor)

m= Basemap(lat_0=0, lon_0=180)
m.drawparallels(np.arange(-60., 61., 30.), labels=[1,0,0,0], fontsize=10)
m.drawmeridians(np.arange(-180., 181., 60.), labels=[0,0,0,1], fontsize=10)
m.drawcoastlines()
m.fillcontinents(color='white')
cs=m.contourf(lon,lat,chl,np.arange(-4,3,0.1),cmap='jet')
cbar = m.colorbar(cs, location='bottom', pad="10%",label='Log Chla (\mug/L)')
plt.title('Global distribution of chlorophyll concentration ',Fontsize='10') #添加标题
plt.show()

```



(2)

```
import numpy as np
import netCDF4 as nc
from mpl_toolkits.basemap import Basemap
import matplotlib.pyplot as plt
file=nc.Dataset('A20171212017151.L3m_MO_CHL_chlor_a_9km.nc')
chlor=(file.variables['chlor_a'][:])
lons=(file.variables['lon'][:])
lats=(file.variables['lat'][:])
lon,lat=np.meshgrid(lons,lats)
chl = np.log(chlor)
m=Basemap(projection='lcc',lat_0=22.,
lon_0=117.,llcrnrlat=0,urcrnrlat=45,llcrnrlon=102,urcrnrlon=132,resolution='i')
m.drawparallels(np.arange(-90., 91., 9.), labels=[1,0,0,0], fontsize=10) #为地图添加经纬线
m.drawmeridians(np.arange(-180., 181., 6.), labels=[0,0,0,1], fontsize=10)
m.drawcoastlines()
m.fillcontinents(color='gray')
x,y=m(lon,lat)
cs=m.contourf(x,y,chl,np.arange(-3,2.5,0.1),cmap='jet')
cbar = m.colorbar(cs, location='right', pad="10%",label='Log Chla (\mu\text{g/L})') #显示颜色图例
plt.title('Distribution of chlorophyll concentration near China',FontSize='10') #添加标题
plt.show()
```

Distribution of chlorophyll concentration near China

