# Mini project: Analyzing $NO_2$ air quality using space-based and ground-based observations

**Developed by Xiaomeng Jin (Columbia University) for the NASA HAQAST6 Meeting**
**July 12, 2019**

## Part I: Make maps of $NO_2$ using Giovanni.

1. Go to http://giovanni.sci.gsfc.nasa.gov .
2. You will need to register an account at NASA EARTHDATA to have full access (otherwise you'll be limited to 4 time steps only). If you don't have an account, follow the instructions to register for an account.
3. Once logged in, you'll be automatically directed to Giovanni website with full access.
4. Click on 'NO2' under the 'Measurements'.
5. Plot annual average tropospheric $NO_2$ columns for 2018 (2018-01-01 to 2018-12-31) over the United States under the 'Maps' option. Save the figure as a .png file and the data as a NetCDF file (other options include KMZ and GEOTIFF). Customize the map if you wish.
6. Make a similar map but for 2005. Save the figure as a .png file and the data as a NetCDF file.

## Part II: Make time series of $NO_2$

7. Go back to Data Selection. For the city with the highest $NO_2$, select a 1x1 degree box centered around the city. Plot the time series of area average $NO_2$ in 2018 for this city (under the 'Time Series' tab, select 'Area Average'). Download the plot as a .png and the accompanying data as a .csv file.
8. Load the data in Excel. Remove missing values (-1e30).

## Part III: Compare satellite observations with ground-based measurements

9. Go to the EPA AQS website: https://www.epa.gov/outdoor-air-quality-data
10. Go to the 'Download Daily Data' page.

11. Download ground-based $NO_2$ data for the city you selected to download tropospheric column $NO_2$ data from in Giovanni, again for 2018. Select "NO2" under 'Pollutant'. Select the same city and pick a single site. Download the data as a .csv file.

12. Load this file in excel. Combine the ground-based $NO_2$ data with satellite-based $NO_2$ data by dates. Note there may be some mismatch in dates.

13. Assume tropospheric $NO_2$ only exists in the boundary layer (height = 1 km), and is well mixed within the boundary layer. The number density of air is $2.53\times10^{25}$ molecules/$m^3$ at sea level for P = 1013 hPa, T = 20˚C. Convert the satellite-based $NO_2$ to the same units as the ground-based $NO_2$.

14. Make a scatter plot between the ground-based $NO_2$ vs. satellite-based $NO_2$. Fit a linear regression. Calculate the R value. Do they agree with each other?

15. Consider exploring different plotting options on Giovanni and datasets for your other applications.

16. (BONUS) Giovanni does not support comparison between two time periods, but we can use other tools like Panoply to calculate the difference in $NO_2$. Download Panoply at https://www.giss.nasa.gov/tools/panoply/ , and install the software. Under Panoply, open the two NetCDF files you saved at steps 5 and 6. Double click the variable to be plotted, and select 'georeferenced longitude-latitude plot'. Once one file is plotted, drag the same variable of the other file to the plot; hold until you see a green plus sign. Once overlaid on the first layer, Panoply will automatically create a map of the difference between the two variables. Panoply also allows other calculations. Change the scale of the data under 'Scale'.

---

## Other Materials:

1. Official Giovanni User Guide: https://giovanni.gsfc.nasa.gov/giovanni/doc/UsersManualworkingdocument.docx.html
2. Giovanni tutorial created by the NASA HAQAST Communications Team: https://haqast.wiscweb.wisc.edu/nasa-giovanni-tutorial/
3. Daniel Tong, Suhung Shen, Jennifer Wei (2018), How-to use the NASA Giovanni portal to create time series plots of satellite $NO_2$: https://www.ldeo.columbia.edu/~amfiore/haqast_TT_files/Giovanni_how_to_create_longterm_timeseries_plots_20180813.docx
4. YouTube channel of NASA Goddard Earth Science Data and Information Services Center, Giovanni Analysis Options Tutorial Series: https://www.youtube.com/user/NASAGESDISC