Session 2: Questions & Answers

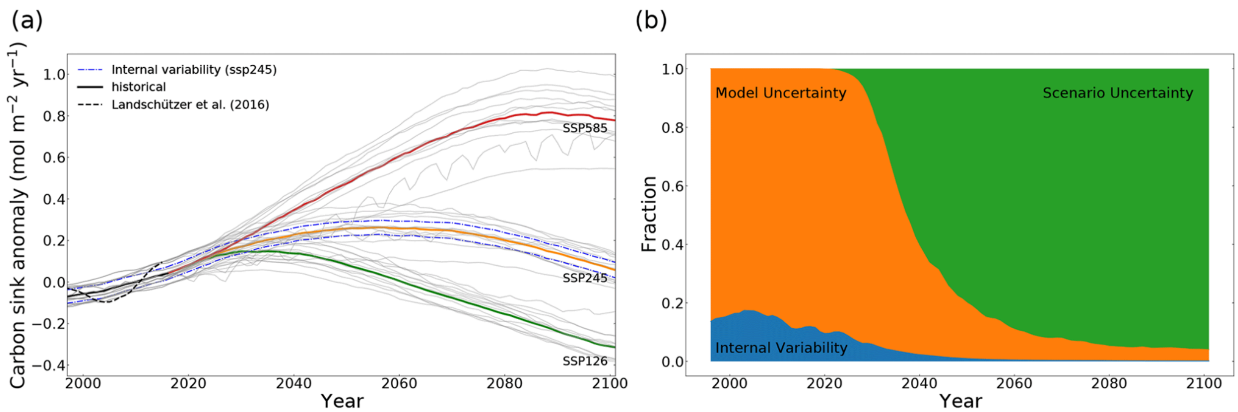
Please type your questions in the Question Box. We will try our best to answer all your questions. If we don't, feel free to email Junjie Liu (junjie.liu@jpl.nasa.gov), Vivienne Payne, (vivienne.h.payne@jpl.nasa.gov), Karen Yuen (karen.yuen@jpl.nasa.gov), David Moroni (david.f.moroni@jpl.nasa.gov), Abhishek Chatterjee (abhishek.chatterjee@jpl.nasa.gov) or Erika Podest (erika.podest@jpl.nasa.gov)

Question 1: Could you explain the concept of integration and its significance in the CO_2 formula? And also state what it means, with regards to what it translates to in our natural language for the sake of understanding?
Answer 1: In the CO_2 formula, "integration" means summation of the surface fluxes over a specific time period.

Question 2: How can we just ignore water vapor when discussing concentrations and fluxes? Everything is with respect to "dry air", but any column of atmosphere has some level of moisture.
Answer 2: You are right that the atmosphere always has some level of water vapor. So when we calculate the column CO_2 concentration, we need to first remove the water vapor mass from the total air mass and then calculate the number of dry air molecules using the conversion factor of 28.96 g/mol. The reason that we use dry air is because the total dry air mass in the atmosphere is constant with time, while the natural air with water vapor could have different total mass at different time periods, which could confuse the change in atmospheric water vapor with changes in atmospheric CO_2 concentration.

Question 3: Could the variation of atmospheric concentration of CO_2 during Spring and Summer in each hemisphere be possibly due to monsoonal activities?
Answer 3: The monsoonal activities contribute to the seasonal variation of atmospheric CO_2 concentration through two mechanisms. One is by its impact on atmospheric transport patterns, and the other is through its impact on the surface carbon cycle by changing climate states.

Question 4: What are the predictions of when the ocean carbon flux will stop being a significant sink?
Answer 4: When the ocean carbon flux will stop being a significant sink depends on the future fossil fuel emission scenario (e.g., Gooya et al., 2023). We have included a reference below:
The figure below shows different emission scenarios. The higher emission scenario, SSP585, shows that the ocean will remain a significant carbon sink. However, the lower emission scenario, SSP126, shows that the ocean will stop being a carbon sink around 2060s.



Gooya, P., Swart, N. C., and Hamme, R. C.: Time-varying changes and uncertainties in the CMIP6 ocean carbon sink from global to local scale, Earth Syst. Dynam., 14, 383–398, https://doi.org/10.5194/esd-14-383-2023, 2023.

Question 5: The Sahara desert is devoid of vegetation, for the most part, and is less of a CO_2 source from fossil fuels. Because of this, one would expect that this reflects on the simulation of atmospheric CO_2, but it doesn't. Why is that so?
Answer 5: The changes in atmospheric CO_2 over the Sahara desert are largely due to atmospheric transport, since the magnitude of surface carbon fluxes is very weak.

Question 6: Can a city, land parcel, or a designated area for agroforestry be tracked if it has become a carbon sink? Can a village or agricultural land parcel be tracked for carbon credits? Who would like to see a step by step guide or a process document on how this can be implemented? Do you have a case study to share with the data?
Answer 6: This is a great question. There are several methods that you can use to estimate carbon sinks. One is to use atmospheric CO_2 concentrations combined with modeling. Our team has estimates of country-scale carbon budgets (Byrne et al., 2023).

You can calculate the carbon sink for the area you are interested in with the dataset our team generated. However, since our primary data constraint is atmospheric CO_2 concentration, we are more confident in the carbon sink estimates over a large region, probably not over a village.
Another way to track carbon sink is through biomass estimates (Xu et al., 2021) that are based on satellite and ground observations. This method has higher spatial resolution but coarser temporal resolution (annual).
The United Nations Frameworks on Climate Change (UNFCCC) has published a field manual on how to assess stock changes:
https://unfccc.int/resource/docs/publications/cdm_afforestation_field-manual_web.pdf

Byrne, B., et al. National CO_2 budgets (2015–2020) inferred from atmospheric CO_2 observations in support of the global stocktake, Earth Syst. Sci. Data, 15, 963–1004,

Liang Xu *et al.* ,Changes in global terrestrial live biomass over the 21st century.*Sci. Adv.*7,eabe9829(2021).DOI:10.1126/sciadv.abe9829

Question 7: Can you provide an intuitive sense of the rate of lateral transport compared to local sinks and sources.
Answer 7: It depends on the region - for a large point source or a megacity, the CO2 enhancement due to local fluxes could be a few ppm on some days, such as in Los Angeles or a large power plant. Remember that the magnitude of CO2 concentration is more than 410 ppm.
On an annual scale, the global CO2 increase is about 2.5 ppm, which is the sum of the contributions of all the surface carbon fluxes. When you plot CO2 in your area, you could still see about 2.5 ppm increase. However, we need to keep in mind that the contribution of local fluxes to this 2.5 ppm depends on the magnitude of surface sources and sinks. This 2.5ppm increase largely comes from lateral transport.

Question 8: Why install xarray and netcdf4 library? Doesn't xarray support NC files?
Answer 8: While xarray does support netCDF files, the notebooks for Part 2 only feature the netcdf4 library. We provide xarray in our environment.yml file as an optional package for users who prefer the extended features of xarray. Another helpful resource is the GES DISC Github tutorials repository: https://github.com/nasa/gesdisc-tutorials. In this GES DISC repository, you'll find many other examples of how to access, extract

and do various types of analysis on various datasets available at GES DISC, including examples featuring xarray.

Question 9: Can I download the CO_2 data for a specific region? Or just clip the NetCDF?
Answer 9: That can be done using OPeNDAP. The Part 3 demo explains how to do this using OPeNDAP. We would have to be cautious when we describe what we mean by region. OCO-2 and OCO-3 are not "mappers" per say, so the data products available are bound by the instrument views. We will show in activity 3, for OCO-3 how we can get data for specific areas based on lat/lon; however, this will still be an averaged quantity at a given time. Please check out the Snapshot Area Maps (SAM) available at: https://ocov3.jpl.nasa.gov/sams/ and join us for the third demo in Session 3. Dr. Chatterjee will provide more science context as well.

Question 10: If we load these notebooks in Google Colab, will the download script work there or will we face any authentication errors?
Answer 10: We have not tested these notebooks in Google Colab. The download script uses HTTPS-based download, so in theory it should work in any Jupyter-based environment/platform.

Question 11: Is there any training data related to carbon dioxide measurements oriented to test Machine Learning algorithms (data cubes, datasets, etc.)?
Answer 11: Not to our knowledge, but there is generic ML training provided by ARSET with a github repo: https://github.com/NASAARSET/ARSET_ML_Fundamentals. Refer to our previous Machine Learning training and Question 12.

Question 12: Is it possible to use artificial intelligence and machine learning in inferring CO_2 measurement from ensemble models and multi sensors (AI, ML prediction, multi sensor data fusion)?
Answer 12: This is an active research area. There are several studies that use machine learning to infer global CO_2 concentration with OCO-2 observations.
https://dl.acm.org/doi/abs/10.1145/3570991.3571062
https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2023GL107536
https://essd.copernicus.org/preprints/essd-2023-465/
https://www.sciencedirect.com/science/article/pii/S0048969722066876

Question 13: Is there a way to connect and get the data through Google Earth Engine?
Answer 13: GEE is not in scope for this training and it's up to the user to incorporate their own preferred datasets into GEE, which supports certain raster file formats, such as GeoTIFF. As OCO-2 and OCO-3 data is natively NetCDF, the user would have to perform their own file format conversion to work with outside tools and services such as GEE.

Question 14: The data has a coarse spatial resolution. Are there any methods, such as downscaling, that can increase this resolution? What parameters (such as environmental factors) should be considered when choosing a method to enhance spatial resolution?
Answer 14: In the OCO-2 notebook, we are simply doing nearest neighbor binning at a relatively coarse 5 by 5 degree lat/lon resolution. Rescaling to higher resolution was not considered in the scope of this exercise.
We suggest participants review these two papers (and other references therein) for recent attempts to downscale the OCO-2 datasets (XCO2 or SIF) to higher resolutions. These references also outline the various factors that need to be considered when applying AI or ML based downscaling methods.

1. For XCO2 - https://ieeexplore.ieee.org/document/10475183
2. For SIF - https://www.sciencedirect.com/science/article/pii/S0168169922005737?via%3Dihub

Question 15: Are all data you are using in the demonstration available online please?
Answer 15: Yes. The landing pages for each dataset are hyperlinked in the top markdown cell of each notebook. Below are the direct dataset links for what was presented in today's Part 2 notebook demonstration:
https://daac.gsfc.nasa.gov/datasets/OCO2_L2_Lite_FP_11.1r/summary
https://disc.gsfc.nasa.gov/datasets/CMSFluxNBE_3/summary
https://disc.gsfc.nasa.gov/datasets/CMSFluxOcean_3/summary
https://disc.gsfc.nasa.gov/datasets/CMSFluxFossilFuelPrior_3/summary

Question 16: The white areas mean no flux? Why?
Answer 16: Depending on the flux product, it could be due to missing data over land, which would be the case for the ocean flux product, or data not available over the oceans, which would be the case for the NBE product. In areas where we expect to see fluxes, the white area in the color bar would account for near-zero fluxes.

Question 17: Please advise the best methodology to estimate above ground carbon stock for cropland and grassland.
Answer 17: GEDI ecosystem LiDAR on the space station and other radar and lidar data can be used to estimate above ground carbon stock. Here is a reference that calculates global above ground and below ground biomass using Radar, LiDAR, and ground measurements.
Liang Xu *et al.* ,Changes in global terrestrial live biomass over the 21st century.*Sci. Adv.*7,eabe9829(2021).DOI:10.1126/sciadv.abe9829

Question 18: Between NPP and GPP, which one is okay to estimate carbon stock?
Answer 18: GPP is gross primary product, which represents the amount of gross carbon fixed through photosynthesis, while NPP is the GPP minus the rate of energy loss to metabolism and maintenance. The changes in live biomass carbon stock can be calculated from the integration of NPP over a specific time period.

Question 19: Why the large carbon sink over land in 2022?
Answer 19: It could be related to the La Nina condition in 2022. During La Nina years, the tropical land receives more rain and has lower temperature, which in general leads to more vegetation growth and more carbon uptake. However, the exact reason needs more detailed analysis. WIth the flux data that we shared during this course, you can investigate what processes cause the large carbon sink in 2022.
https://www.aces.edu/blog/topics/crop-production/el-nino-southern-oscillation-and-its-impact-on-alabamas-climate/

Question 20: Do you know of any artificial intelligence applications that do this? I could see AI being taught to do this 24/7.
Answer 20: Machine learning approaches have been used to estimate CO_2 concentrations. See the answer for question 12. There is no published work yet on how to estimate carbon fluxes with machine learning, but it is an active research area.

Question 21: Can we export the data to a TIF format, similar to JPG, and analyze the changes in carbon dioxide during a specific time period?
Answer 21: While we didn't develop the code to have a TIF or GeoTIFF conversion capability, there are open source solutions in Python that can be incorporated into a Jupyter notebook. Here is an example (note: this is not an endorsement, so please use caution): https://corteva.github.io/rioxarray/stable/examples/convert_to_raster.html.

Question 22: Can we use the analyses outlined today to gain insight on areas that may be on the verge of desertification?
Answer 22: If an area is on the verge of desertification, you can see the carbon fluxes related to that event. You can also use this information to make predictions of future events.

Question 23: What should be considered when comparing and interpreting OCO-2/3 XCO_2 data with ground-based, remote sensing, and model observations?
Answer 23: When you compare the columns, you should take in account the averaging kernels and pressure weighting functions.

Question 24: Data is retrieved from a website. There are several mathematical formulas that show how to perform calculations. Are there any bands, such as RGB, NIR, or SWIR, that can be downloaded and processed using software? Does the OCO-2 satellite have spectral bands similar to other satellites?
Answer 24: The OCO-2 and OCO-3 share the same type of instrument, which is a grating spectrometer that focuses on 3 spectral bands in the NIR and SWIR. The OCO-2 Level 1B radiance spectra are publicly available at the GES DISC and the Level 2 algorithm software has been released as open source:
https://software.nasa.gov/software/NPO-49044-1

Question 25: Is there a plan to update data for the months of May and June in 2024 and subsequent months? The availability is limited until April.
Answer 25: Data will become available after processing, so yes, May and June data will be forthcoming and available.

Question 26: Nowadays the historically dry areas received good and heavy rainfall. Is there any linkage with recent increases in CO_2 concentration?

Answer 26: Increase in rainfall in dry areas often promotes photosynthesis and net carbon uptake from the atmosphere, which leads to a reduction in CO_2 concentration, not an increase in CO_2 concentration. The annual increase of CO_2 concentration is largely due to anthropogenic emissions. On a long-term average, about half of the anthropogenic emissions remain in the atmosphere, and the rest has been absorbed by the natural carbon cycle over land and ocean.

Question 27: Are there any studies done by this team regarding the present condition of the Euphrates river and solutions?
Answer 27: Our team has not done studies regarding the present condition of the Euphrates river. We encourage you to use our data to explore this question.

Question 28: Is there any spatial autocorrelation between CO_2 concentration and Urban Heat Islands in urban areas?
Answer 28: This is an interesting question. I encourage you to explore this question with our data. The urban CO_2 concentration changes are due to both urban anthropogenic emissions and urban biosphere carbon fluxes. The impact of the urban heat island effect on the urban biosphere and on urban anthropogenic emissions could lead to changes in CO_2 concentration.

Question 29: Are these CO_2 emissions responsible for monsoonal variation?
Answer 29: The annual CO_2 emissions cause an increase in atmospheric CO_2 concentration, which is the primary driver for global climate change.

Question 30: What is the absorption capacity of the gasses being studied?
Answer 30: Not sure what is meant by this question. The molecular optical depths vary with gas and with spectral position.