Part 2 Questions & Answers Session

Please type your questions in the Question Box. We will try our best to get to all your questions. If we don't, feel free to email Amita Mehta (amita.v.mehta@nasa.gov).

Question 1: Are there more resources we can access to learn more about PACE data access and processing?
Answer 1: Part 3 of this ARSET training – Access and Visualization of PACE/OCI Data using Python/Jupyter Notebook Software – will provide 2 Jupyter Notebook-based tutorials on PACE data access, visualization, and analysis.

We also have just launched a new help page, The Help Hub, that groups a wealth of resources including videos and Jupyter Notebooks in one place. We will keep adding to it over time, so keep checking back:
https://oceancolor.gsfc.nasa.gov/resources/docs/tutorials/

Question 2: Do you provide consumable data API for PACE?
Answer 2: All NASA Earthdata is programmatically searchable (i.e., has an application programming interface, API) provided by the Common Metadata Repository (CMR). There is a Python wrapper for the API called earthaccess, and an R wrapper called earthdatalogin. We will learn about using earthaccess in Part 3. In the future, we also expect to provide services like subsetting and downloading as different file formats (e.g. GeoTIFF or Zarr) through the Harmony API.

Question 3: Can someone from outside the USA apply to join the PACE Community of Practice?
Answer 3: Yes. Anyone interested in staying up-to-date on the PACE mission, data, and applications activities can sign up. Instructions below.

1. Send an email to *with 'join' in the subject line* to pace-community-join@lists.nasa.gov.
2. Look for a confirmation email, and then confirm!

Question 4: In what ways can hyperspectral chlorophyll-a data contribute to fisheries health assessments?
Answer 4: The hyperspectral capabilities of OCI will allow for a wide range of products that were not possible to get with multispectral sensors. Phytoplankton community composition algorithms, for example, can inform us on the types of plankton and some of them can potentially be linked to fish populations.

Question 5: What advancements in management practices can result from improved understanding of phytoplankton community composition?
Answer 5: A better understanding of phytoplankton community composition could, for example, help us differentiate between toxic and beneficial algae from space. This means we could, in the future, know more about the factors that lead to more toxic algal blooms and help managers work on those factors.

Question 6: Is there a link for HyperCoast?
Answer 6: HyperCoast Data Visualization Tool: https://hypercoast.org/
PACE-specific HyperCoast Notebook: https://hypercoast.org/examples/pace_oci_l2/

Question 7: How can I successfully install and configure the SeaDAS software on my Windows system, and what are the key challenges or requirements during the installation process?
Answer 7: The installation instructions and requirements for SeaDAS can be found on this page: https://seadas.gsfc.nasa.gov/downloads/.

Question 8: Can OCI separate and quantify non-algal particles and phytoplankton contamination? If we use polarization data from HARP2 and SPEXone, can those be more accurate?
Answer 8: Please see the inherent optical property (IOP) related data products listed here for OCI capabilities, in particular the spectral non algal particle absorption coefficient and spectral phytoplankton absorption coefficient relate to your question (but are not yet available): https://pace.oceansciences.org/data_table.htm#category2.

Question 9: How do you differentiate reflectance from algae/plankton and reflectance from sediments in shallow environments?

Answer 9: This depends on what concentration of sediment (and what type) is present compared to the concentration of phytoplankton. The more sediment there is relative to phytoplankton, the harder it can be to differentiate phytoplankton from sediment using PACE-OCI data products. Sediment is highly reflective, providing a larger signal compared to phytoplankton, so enough sediment can overwhelm the phytoplankton signal. In shallow environments this is also true for bottom reflectance of sediment, if present and detectable by the satellite.

Question 10: Can you expand upon what steps will be taken to validate the Pace Std. Data?
Answer 10: Validation data are being collected all over the world by the PACE Validation Science Team currently and over the next couple of years. Learn more here about these efforts at https://pace.oceansciences.org/pvstdoi.htm.

The PACE Mission also just completed a month-long calibration and validation effort in Southern California for OCI and polarimetry products called PACE-PAX. This effort collected coordinated ship, airplane, and ground station data collection. Read more about it, as well as other PACE-related calibration and validation campaigns, here: https://pace.oceansciences.org/campaigns.htm.

Question 11: Could you elaborate on the application of PACE data products for application of waterborne disease like Cholera?
Answer 11: You can read more about Dr. Julta's work here: https://pace.oceansciences.org/people_ea.htm?id=68.

Question 12: How accurately do the PACE data detect a potential fishing zone?
Answer 12: This depends on what type of fish you are looking for and what region you are in. If, for example, the fish of interest are attracted to zones with a lot of phytoplankton, PACE data can estimate where the largest concentrations of phytoplankton are in the water, which indicates a higher probability of those fish being there.

Question 13: Will changing the format of the data affect the quality and projection of the data? For example, I usually convert them either into .tif files or arrays (.txt).

Answer 13: Depends on how you convert the data. Converting a “Level-3 Mapped” product, which comes in a NetCDF4 file format, to a GeoTIFF without changing the projection (from Plate Carree) can be done without affecting the data values. Any reprojection will change the data values, but any geospatial data processing software (including SeaDAS) can do this without reducing data quality.

Question 14: Where can I find detailed information about the algorithms used in HyperCoast to generate chl-a products from PACE imagery?
Answer 14: HyperCoast is an open-sourced software for visualizing and analyzing data products created by researchers outside of NASA. It does not include algorithms for chlorophyll-a (chl-a) products. Rather, it currently relies on the chl-a products distributed by the NASA OB.DAAC which use the standard NASA chl-a algorithm. You can explore the HyperCoast Github Repository to find out more on their software. https://github.com/opengeos/HyperCoast

Question 15: What is the spatial resolution of the OCI level 1C data with the common grid? When projecting the L1b to a common grid, do I lose spatial resolution, since the resolution varies cross-track and changes with scan angle?
Answer 15: The user guide for PACE L1C data products tells you all about how OCI, HARP2, and SPEXOne data are combined for the L1C products. It gives the spatial resolution as 5.2 km by 5.2 km, with the reduction in spatial resolution relative to OCI’s 1 km resolution due to the coarser resolutions of the two polarimeters.

Question 16: Data suite is a great concept. If I want to do something similar from L1B to L2, what do you suggest? I see the SeaDAS command l1bextract_oci only subsets region not wavelengths. Is it difficult to subset wavelengths?
Answer 16: If you go from L1B to L2 using L2gen, you can choose the specific products you want to work with. You can choose as many or as few wavelengths as you want! Refer to the previous SeaDAS training.

Question 17: How do you interpolate the missing values due to cloud coverage? Are there any algorithms that NASA scientists usually use?
Answer 17: In PACE-OCI ocean color data products, no interpolation algorithms are utilized. Pixels covered by clouds are masked out, which means no data values are present for cloud pixels as PACE-OCI cannot “see” through the highly reflective clouds.

Temporally-composited Level 3 OCI data products can help in cloudy areas. These *average* satellite observations in each pixel over a certain time frame, such as 8-day or monthly. Averaging the data over time effectively reduces the number of pixels masked by clouds; the trade off is it also averages out changes in the observed data over that time.

Question 18: Are there any options for image segmentation in the Earthdata portal?
Answer 18: If “image segmentation” is referring to isolating an area of interest, the Earthdata portal does have options for spatial and temporal subsetting of data.

Question 19: When does the PACE data trace back? How about data before 2024?
Answer 19: There is no PACE data before 2024. PACE was launched in 2024 and data products are available starting in March 2024. Previous satellite-based ocean color instruments such as NASA’s MODIS-Aqua and SeaWiFS have collected global, **multi-spectral** data nearly daily since 1997. No satellites previous to PACE have collected global **hyperspectral** observations every 1-2 days.

Question 20: Could it be possible to have a GitHub tutorial on how to upload our code, maybe with a really simple example, that is available permanently, so that we can follow step by step? That would be very helpful for me, for example, that have never used GitHub but would be happy to share my codes. It could also have some best practices, so that we get used to doing it in the best way, from the first time. Thank you very much in advance.
Answer 20: Our team is actively creating tutorials using Jupyter Notebooks and are open to suggestions, we will add this suggestion to the list of topics to potentially consider. Also there are introductory GitHub tutorials you can find by searching online that will teach you the basics of starting and maintaining your own code repository.

Question 21: With Level 2 (L2) data, do we still need to do the atmospheric correction based on local atmospheric conditions to get more accurate values?
Answer 21: L2 data downloaded from Earthdata is atmospherically corrected already. If someone processes L1B to L2 data using L2gen, then the atmospheric correction can be turned off, but is included by default.

Question 22: I'm an educator for the GLOBE Program and I'm taking AOT measurements. I'm very happy to learn how I can compare the AOT satellite data for a specific location to compare with ground based data.
Answer 22: This tutorial Jupyter Notebook walks through steps that are similar to what you are requesting, please take a look. Even if you are not a Python coder, reading through the steps for matching satellite and ground data will still be informative. https://pacehackweek.github.io/pace-2024/presentations/hackweek/satellite_insitu_matchups.html

Question 23: To what level will the phytoplankton composition analysis be? Species or genus or lower?
Answer 23: This is addressed in the Part 1 Q&A #24. “With PACE OCI, our aim is to retrieve taxonomic groups of phytoplankton such as diatoms, dinoflagellates, pico-eukaryotes, etc. There are some phytoplankton groups that OCI can quantify at the genus level such as Prochlorococcus, Synechococcus, “armored” Coccolithophorids such as Emilania huxleyii, Mycrocystis, and hopefully, a few other HABs. See material in the training 1 module on the MOANA algorithm developed for the Atlantic Ocean that provides cell concentrations for Prochlorococcus, Synechococcus, and pico-eukaryotes.”

Question 24: What is the difference between types in OBDAAC download parameters?
Answer 24: Unsure what “types” and “download parameters” mean here. If “types” is referring to “data status” that was discussed in today’s presentation, please see the table in the presentation that describes what each status means. Then, later in the presentation we present how data status relates to where you can find and download data, including the OB.DAAC level 3&4 browser and search feature.

Question 25: How can I extract time series data from a NetCDF file for a specific geo-location, such as chlorophyll-a or band values from level-2 products, for a period of up to 10 years and convert it into a CSV file for data analysis or model development? The second issue is I have personally been facing challenges installing SeaDAS software, maybe a video demo could help.
Answer 25: You will need to download many NetCDF data files to derive a 10 year time series, one file per your data step requirements (e.g. daily, 8-day, monthly). Challenges

in installing SeaDAS are known. A demo/training module is in development - likely to be posted at https://seadas.gsfc.nasa.gov/tutorials/. You can ask for advice via the help forum https://seadas.gsfc.nasa.gov/issue-reporting/. Keep a lookout for both in-person (e.g., upcoming Ocean Optics 2024 meeting) and virtual SeaDAS training events.

Question 26: How far away from release are the phytoplankton ID algorithms?
Answer 26: Initial products will become available in a few months.

Question 27: Can we analyze plastic pollution using hyS?
Answer 27: Yes, scientists such as Dr. Heidi Dierssen are working on this using PACE data.
https://cce.nasa.gov/ocean_biology_biogeochemistry/details.html?itemType=project&itemProgID=8&itemID=4380&projType=project&projID=4380&progID=8

Question 28: Are any specific tools for atmospheric correction available in SeaDAS for PACE processing?
Answer 28: L2gen is a SeaDAS processing tool that can be used to perform atmospheric correction on L1B files.

Question 29: How can I assess the fish health using that data in the freshwater that is close to some industries?
Answer 29: You may be able to use PACE water quality data products to determine the water quality for your sites of interest including presence of Harmful Algal Blooms that may affect fish health. CYAN products are coming to PACE soon. These would be a starting point for water quality in freshwater systems.

Question 30: Thank you for the excellent presentations. How can we do analyses on PFTs using the PACE data? Thank you!
Answer 30: PACE data products for phytoplankton community composition will become available shortly that you may download and use for your particular scientific and applications work.

Question 31: Is there much data validation effort for optically complex coastal waters? For example, Chesapeake Bay outflow or Northern Gulf of Mexico Shelf?

Answer 31: NASA has supported a PACE Validation Science Team that is primarily focused on validating global data products. There is some work in coastal waters such as the nearshore Gulf of Maine, Santa Barbara channel off California and San Diego. There are other PACE validation activities that will be working on US coastal waters including the northern Gulf of Mexico.

Question 32: Should we install Processors together with SeaDAS for PACE?
Answer 32: SeaDAS v.9.0.1 can be used to visualize and process PACE data. If "Processors" in this question refers to the SeaDAS Processors available to install in SeaDAS and found in the SeaDAS-Toolbox pulldown menu, whether or not you install them depends on how you will use SeaDAS. If you are primarily visualizing the data you may not need extra processors installed, but if you plan to use SeaDAS to process your data files you would need the appropriate processors. An example is if you need to process a Level 2 file to a Level 3 file in SeaDAS, this would require the appropriate processors be installed. Please visit the following link to explore the SeaDAS help documentation – look at the topics found under "SeaDAS-Toolbox/OCSSW" on this webpage to learn more about the different processors, what they do, and what your SeaDAS toolbox may need: https://seadas.gsfc.nasa.gov/help/