Part 1 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) or William Wainwright (william.wainwright@nasa.gov).

Question 1: Is it possible to replicate this on another continent? Also, what processing scale is available?
Answer 1: The models used to produce daily water quality maps for STREAM have been trained with the intention of being globally-applicable to inland and coastal water bodies. The models do have varying performance depending on the amount of in-situ samples we've been able to train on from that region, but as the models are open source they are free to be used on your own data as Ryan will outline in Part 2 of the training.

Question 2: What are the countries covered by STREAM ?
Answer 2: STREAM covers the continental US (CONUS), Hawaii, Alaska, and then many select water bodies/satellite tiles around the world including but not limited to India, South Korea, South Africa, Benin, Ghana, Chile, Uruguay, Peru, Colombia, Mexico, and Cuba.

Question 3: Is there any way to monitor groundwater data with SAR data? Is there any way to monitor seawater and groundwater intrusion rates in coastal areas like Sundarbans? Currently, in situ data is being collected in very limited areas in the Sundarbans, but due to the complexity of the region it's not very effective. Is there any way to do water quality monitoring with satellite data here?
Answer 3: ARSET will have a training on groundwater in April where we can address this question. This training is focused on water quality.

Question 4: Does the tool work well with touch screen-based devices like iPads and so on?
Answer 4: STREAM does not currently work very well on mobile devices unless you enable desktop mode. That's something I'm working on improving.

Question 5: What is the spatial coverage of products under STREAM?

Answer 5: Landsat 8/9-derived products retain the native spatial resolution of 30m. Sentinel-2-derived products are delivered at 20m resolution after the bands are resampled. HSL combined products do not contain all the bands we need to derive water quality. Landsat 8 and S2 have global coverage but the WQ products in STREAM are not global.

Question 6: If I select a Sentinel-2 tile, does that mean the data for that tile is estimated from Sentinel-2? If that is the case, why not data fusion using both Sentinel 2 and Landsat?
Answer 6: Yes, you see search results and product maps for the satellite you select. We do not use the harmonized Landsat-Sentinel product because it does not have all of the bands necessary for water quality estimation.

Question 7: Is PACE data on STREAM Web?
Answer 7: PACE is not currently supported for the STREAM website, but the models behind STREAM also work for PACE and you can learn more about them in Part 2 of the training next week.

Question 8: Does it extract data from S2 or Landsat as soon as they are available? Or are the WQ data uploaded to STREAM at a set period, when a considerable amount of products are generated?
Answer 8: We automatically download scenes in bulk from Copernicus (S2) and USGS (L8/9) as they become available at Level-1, typically at the end of the day of observation. Given the 600-800 tiles processed per day, it can typically take until the next day to fully download and process all tiles. There is not a particular order of tiles we process, just as they become available.

Question 9: Can I collect Karenia brevis data (ch-a) in the Gulf of America?
Answer 9: Coastal regions of the Gulf are covered natively by STREAM. If you find that your site is not covered by STREAM, you may apply the models yourself and/or reach out to us about potentially including your region of interest.

Question 10: STREAM gives modeled estimates. How do we get the standard errors of those estimates? How do we get the standard errors for the modeled estimates of chla, TSS, and Secchi depth?
Answer 10: We're working to add uncertainty as an additional product to the archive, but you can learn a lot more about the training and validation of the models in Part 2 of the training next week.

Question 11: Is there any way that I can improve the temporal resolution of the data? I mean if I need daily data, how can I get it?
Answer 11: One of the tradeoffs of the increased spatial resolution of Landsat and Sentinel-2 relative to other water quality satellites is the decreased temporal resolution. In the future, STREAM will begin to cover additional satellites with an aim of increasing temporal resolution.

Question 12: Is this calibrated? How accurate are the results? Is there any way to calibrate the data?
Answer 12: The models behind STREAM are heavily calibrated and validated using datasets like GLORIA and AERONET-OC for training/testing. The accuracy of the results can vary depending on how many samples we have to train against in a given region. The best way to calibrate the data for STREAM as a whole would be to help us expand our in-situ dataset. You can also run the models locally to assess performance in your regions of interest.

Question 13: Are these values of different parameters only for surface water? or do they represent sub surface water in the water bodies?
Answer 13: Most of the in situ data is taken within the first ~0.5-1 meters, representing surface waters. Most of the ocean color signal comes from the surface waters in optically complex eutrophic waters.

Question 14: What is the model evaluation for the model used for estimating chl-a and TSS for sentinel and that of Landsat? Is there a reference that can be cited for that information?
Answer 14: The name for the model is “Aquaverse”. The publications which evaluate Landsat MDNs are available here: https://doi.org/10.1016/j.rse.2021.112860 and https://doi.org/10.1016/j.rse.2023.113889

Question 15: Does STREAM work in seasonal bodies of water around the world?
Answer 15: Yes, I believe if the water body is visible by Sentinel-2/Landsat 8 & 9 (resolveable by 20m-30m resolution) then you should be able to monitor it via STREAM. We don’t apply any fixed land masking so if water appears in a scene it will be displayed on STREAM.

Question 16: Is it possible to have a quick look at the data used in the estimations of Chla and other predictors?

Answer 16: The GLORIA and AERONET-OC datasets are publicly available, but you can learn more about the data and products used in next week's session.

Question 17: Can we get information about other water quality parameters? For example, if I need data regarding nutrients, can I get that?
Answer 17: I don't believe nutrients are optically active (interact substantially with the visible and near-infrared light). However, I believe nutrients are often correlated with TSS, though that is not my area of expertise.

Question 18: Please clarify the date ranges available for these analyses. What is the earliest data available to analyze?
Answer 18: It is from June 2024 onwards for the online webtool (STREAM). Some tiles have been processed going back further but that was done on an ad hoc basis.

Question 19: When analyzing TDS from satellite data, do we focus on the highest value, the lowest value, or a different metric (like the average)?
Answer 19: Range and average should both be looked at.

Question 20: If we apply STREAM to a region like Jordan that is not explicitly listed, what type of local data would most improve model calibration and confidence in the results?
Answer 20: Any in-situ data you're able to provide would likely help increase the performance of models in that region. Obviously the more unique or under-represented a water system is, the more likely new data is to improve accuracy, but generally speaking any new data helps.

Question 21: Also, does the model only use Landsat and Sentinel or are there other hydrological and meteorological parameters used?
Answer 21: Wind speed, water vapor, ozone and NO_2 concentrations. Aerosol is taken care of by the MDN but the aforementioned data are used in atmospheric correction.

Question 22: Is it possible to use the Himawari Geostationary satellite with STREAM?
Answer 22: We would have to make sure the models will support this.

Question 23: For the nutrient level analysis, since these are not optically inherent properties of water, indirect methods such as empirical methods, ML methods are the only way?

Answer 23: If they do not affect the optical characteristics of the water directly then you would have to correlate them somehow to a property that does affect the optical characteristics of the water in order to be able to use remote sensing.

Question 24: How soon would the rest of the world be covered with this model, especially the North Africa region?
Answer 24: Full global coverage is pending in STREAM due to the processing and storage. This is why the model is open source which you can apply in your region.

Question 25: Can we use STREAM for monitoring water quality in rivers?
Answer 25: Yes, depending on the width of the river (it must be >~60 m width). Our dataset comes from a wide range of inland and coastal water bodies, representing a wide range of optical conditions.

Question 26: Can the bacterial count like e coli and fecal coliform also be found out through STREAM?
Answer 26: Similar to nutrients, if you can trace it to WQ parameters, possibly.

Question 27: What are the top 5 features for the model from the feature importance analysis? Are they the same for Landsat and Sentinel?
Answer 27: We will address this in the next session.

Question 28: Is there any training for applying the model for individual locations?
Answer 28: We will address the model in Part 2 of this training.

Question 29: When STREAM is applied in regions without extensive in-situ data, what level of confidence is considered sufficient for decision-making or policy applications?
Answer 29: This will depend upon the decisions being made. Remote sensing derived water quality will never be as accurate as in situ but it can depend whether that increased accuracy is required or not. STREAM may be enough for you to make decisions off directly or it may better serve as a guide to apply limited in-situ sampling resources, for example.

Question 30: Can STREAM be used for a river that is less than 30 m wide but very long?
Answer 30: It can be used but you may see a reduction in accuracy (how much land was averaged in, etc). If less than 30 m wide (you need at least pixels to average).

Question 31: Do you use any of the water processing softwares such as ACOLITE, Sen2Corr, lib2gen, or C2RCC for converting TOA surface level reflectance to water leaving reflectance for the STREAM platform?
Answer 31: We regularly compare our processing to those atmospheric correction engines, but we use our own combination of models for STREAM.

Question 32: What model was used? Machine learning or physical models? Which particular ones were evaluated before the best was chosen?
Answer 32: We use physical models for Rayleigh scattering and gaseous absorption corrections as we found them to be more performant than MDN. We use MDN for aerosol contributions and for water quality parameters as those models are much more performant.

Question 33: What are the most common misinterpretations or misuse of STREAM outputs that users should be aware of?
Answer 33: Cloud or land pixels may show up and will naturally throw off the results. We're working to improve masking and reduce the frequency of these events, but the RGB true color maps should provide some clarity in cases of doubt.

Question 34: Can it monitor turbidity in lakes?
Answer 34: Turbidity correlates to Secchi depth.

Question 35: Is this service going to be available worldwide? Is the availability of STREAM data layers contingent upon availability of in situ data? Are there any plans to expand the area covered by the tool to places outside the USA?
Answer 35: The models are trained to be applicable globally using a wide variety of in-situ samples, so being able to apply them is not contingent upon in-situ data being available at that location. Coverage of daily processing by STREAM though is not currently able to be available worldwide due to the steep processing requirements.

Question 36: I have used the blue band on Landsat 8 OLI but I'm not sure about outputs or which bands are best to evaluate the water/are there specific indices?
Answer 36: STREAM takes level-1 TOA data that you may be familiar with like the blue band and estimates water quality parameters like Chla, TSS, and Secchi depth. How those parameters can be used to evaluate the water depends largely on your use cases.

Question 37: For the website and the API example script, is there any filtering for cloud coverage?
Answer 37: We don't currently provide a means for filtering cloud cover.

Question 38: I'm curious what in-situ training data sets were used for Secchi depth. Did you use The North American Lake Management Society's National Secchi Dip in? GLOBE also has a Secchi dataset with 62K measurements that we are currently curating.
Answer 38: I'll have to check with my colleagues if that dataset is included in our training data but that sounds very exciting!

Question 39: My research broadly involves assessing water quality in lakes (~100,000 lakes) across the United States. Processing the TIF files from STREAM seems like intensive work. I am curious if any attempts were made to summarize the water quality data (say chlorophyll a) at the national scale for different time periods?
Answer 39: STREAM attempts to reduce the friction in accessing water quality maps by stitching together the different TIF files into a connected map. In the near future, we will be developing time series tools built into STREAM to make analyzing trends at any scale as easy as possible.

Question 40: Complete beginner here, what do they mean by tile list?
Answer 40: The tiles referred to on STREAM are the individual squares in the satellite grid drawn over Earth's surface. You often request data for a water body based on which grid tile it is contained within, or sometimes multiple tiles if your water body spans multiple.

Question 41: I'm looking for STREAM products for Colombia, but I can't find them. In which areas of the world is STREAM available?
Answer 41: STREAM covers the Continental US (CONUS), Hawaii, Alaska, and then many select water bodies/satellite tiles around the world including but not limited to India, South Korea, South Africa, Benin, Ghana, Chile, Uruguay, Peru, Colombia, Mexico, and Cuba. One tool I would like to implement soon is a quick visual reference for tiles which are covered or not on the global grid. So when you enable the grid it would clearly highlight whether a tile is covered by STREAM's daily processing or not.

Question 42: Is it a must to use QGIS or can I use ArcGIS Pro?

Answer 42: Any GIS tool should suffice. The TIF files are geo-referenced so you can interact with them in a variety of ways including GDAL

Question 43: When we extract the water quality parameters for a given ROI, is that an average value of the ROI?
Answer 43: Yes.

Question 44: Is the MDN model used to produce the STREAM images publicly available?
Answer 44: It is, it's called Aquaverse, and will be the subject of the next ARSET training!

Question 45: Could the ML algorithms used in STREAM be applied in Google Earth Engine?
Answer 45: We do not have plans to support GEE.

Question 46: Is there a plan to have archival data before 2024 available in STREAM?
Answer 46: There definitely is, and it would be necessary to make full use of the planned time series tools. The timeline for filling in the backlog of data is unknown though as the catalog of products and sensors covered is still growing.

Question 47: There is another app, cYan, which also models water quality across the USA. What is the difference/relationship between these initiatives?
Answer 47: STREAM currently covers Landsat 8/9 and Sentinel-2 while CyAN uses Sentinel-3 and MERIS data.

Question 48: Did you develop the MDN atmospheric correction method specially for this project?
Answer 48: Yes and no. Our team developed the models to improve the accuracy of remote sensing water quality for inland and coastal waters. One envisioned application of those models was STREAM, but the models are not solely intended to be used for STREAM.

Question 49: Is there any procedure to calculate Odor Indices of Inland Wetland?
Answer 49: No, we do not plan to support Odor Indices of Inland Wetlands.

Question 50: Can it be used for water flow?

Answer 50: wNo, but you may want to check out SWOT: https://swot.jpl.nasa.gov/

Question 51: Is there a procedure for identifying Mercury content on land and in rivers affected by gold mining?
Answer 51: We are not aware of a procedure for that, and if there were one it would have to be traceable to water quality parameters we provide on STREAM in order for it to be relevant in that application.

Question 52: Are there any benchmarking studies to identify accuracies of Landsat vs. Sentinel based estimates for the same spatial extent?
Answer 52: We have multiple publications comparing our available models and algorithms. One such publication is: https://doi.org/10.1016/j.rse.2021.112860

Question 53: Is it scientifically valid to compute a comprehensive wetland water quality index by taking the simple mean of Chlorophyll-a, TSS, and turbidity, given their different units, scales, and ecological significance?
Answer 53: Unsure at this time, but probably not since the products have different units, scales, and significance.

Question 54: Do you have any suggestions for monitoring very small water bodies?
Answer 54: We would recommend using Planet Superdove with our MDN: https://doi.org/10.3389/frsen.2025.1624783

Question 55: In case of shallow waters, is there any correction for the bottom effect?
Answer 55: Not yet to apply to coastal applications

Question 56: Is it possible to analyze organic pollutant concentration as well in water like PFAS, Organic dyes, trace quantities of drugs?
Answer 56: No, we do not support that.

Question 57: Cyanobacteria are responsible for producing a large proportion of the oxygen on Earth. Can satellite images of chlorophyll levels be used to identify areas where oxygen production is high?
Answer 57: Not that we are aware.

Question 58: To what extent can surface water monitoring data (flow, water quality, satellite imagery) within STREAM be used to identify groundwater recharge zones and pathways, and what are the limitations compared with field measurements and hydrogeological data?
Question 58: Flow is not available through STREAM.

Question 59: Can we measure parameters for COD (BOD)?
Question 59: Not with STREAM.

Question 60: Can the Python example script in the STREAM API documentation be used to trigger MDN processing for tiles outside the current 'supported countries' list, specifically for the Niger River in Nigeria?
Question 60: It does not trigger any new processing.