

## Discussion points regarding the retention of MODIS products from earlier collections

The ESDIS Project manager has asked me to confirm from the perspective of the MODIS Science Team that deleting an older version of a granule 6 months after it has been reprocessed will not have significant impact on the science communities that order MODIS data products. Based on discussions with the MODIS science disciplines, I sent an email to that effect in May. I've agreed to revisit this question, since the ESDIS Project is willing to hold the entire previous collection until 6 months after the current reprocessing has completed. Holding the entire previous collection until 6 months after the latest reprocessing has been completed will add significantly more volume to the archive. If we'd like to exercise this option, the ESDIS Project would like us to request this in an email and cite reasons for storing the older collection beyond the 6 months specified in the current requirements.

There are several possible options to consider:

- Option 1a. Stay within current Functional and Performance Requirements Specifications (F&PRS) for the EODIS Core System. The requirement is a 6 calendar month rolling delete of products from the older collection. The reprocessing rate is assumed to be 2X for each mission. If we reprocess faster than 2X per mission, the DAACs would hold products from the old collection for a proportionally shorter time, i.e. at 4X per mission the products from the older collection would be held for 3 calendar months before being purged.
- Option 1b. Implement a 6 calendar month rolling delete of products that assumes a higher reprocessing rate than in the F&PRS. For example, if we run at 4X then the DAACs would hold 24 data months per mission of the older collection before deleting the granules.
- Option 2. Hold all of the previous collection until 6 months after the current reprocessing has finished.
- Option 3a. Implement a rolling delete of most products as described in Option 1a but select some Level 3 products to save until 6 months after the current reprocessing has finished.
- Option 3b. Implement a 6 calendar month rolling delete of most products but use a higher reprocessing rate than 2X per mission but select some Level 3 products to save until 6 months after the current reprocessing has finished.

In discussions of these options at the PIP meeting, there was a general consensus on several points.

- First, for all Level 2 and Level 2g products, the granule from the previous collection can be deleted from the archive as soon as the same granule for the new collection is archived.

- Second, discipline representatives stated that there are some products for which it may be useful to the science community archive both the versions from the current and previous collection since they are key to research in global change.
  - Bill Ridgway suggested retaining all Level 3 Atmosphere products from the previous Collection, roughly 0.2TB/mission-year.
  - Wayne Esaias suggested retaining Level 3 Ocean chlorophyll parameters, at TBD TB/year. Storing half of all the Ocean L3 maps for Ocean Color would be an additional volume of 2.4TB/ mission-year.
  - Robert Wolfe suggested that there may be some Level 3 land products at 8-day and higher compositing periods that the Land Science Team will want to see archived. Volume is TBD TB/mission-year based on products selected. Storing all 8-day and higher products would amount to an additional 12.2TB/year.
- Third, with respect to selecting products to be archived for Collections prior to the current one, Headquarters program managers for Oceans, Atmospheres and Land should be involved in the discussion as representatives of the respective disciplines.

As the requirements currently stand, the responsibility for saving data products from older collections that have been used in published research lies with researchers who have ordered the data products. From the perspective of the MODIS Science Team, products in a new collection are of improved quality over those in a prior collection. Therefore, long term trends found in earlier collections should persist into later ones if they are real. It is on this basis that the deletion of all products from the prior collection from the archive was deemed to be OK when this first came up a couple of months ago.

Another point we should consider is the time to retain Level 1B, Cloud Mask and Profiles products from the previous Collection once new versions of granules are made. If we are suggesting that Level 2 products could be deleted from the archive as soon as the newest Level 2 granule is produced, then do we also endorse deleting the Level 1B product from the earlier collection as soon as the newest Level 1B granule is produced?

If we delete the older collection for L1B-Daily products as soon as the latest version of the products are inserted into the archives instead of a 6 month rolling delete, the savings would be 195.8TB. The potential savings by eliminating the amount buffered for Level 1 through Level 2 for the 6 month rolling delete at the DAACs are as follows: 78.1TB at GES DAAC, 2.3TB at NSIDC and 114.3TB at the LP DAAC. The savings should allow the disciplines to store key products between the current and previous reprocessing collections within the current requirements carried in the F&PRS.

At the Technical Team meeting on 8/7 it was suggested that storing all products from a couple of data months in each reprocessing campaign would enable interested science users to replicate the results of the reprocessing. Science users would also be able to compare any product between collections in these intervals.