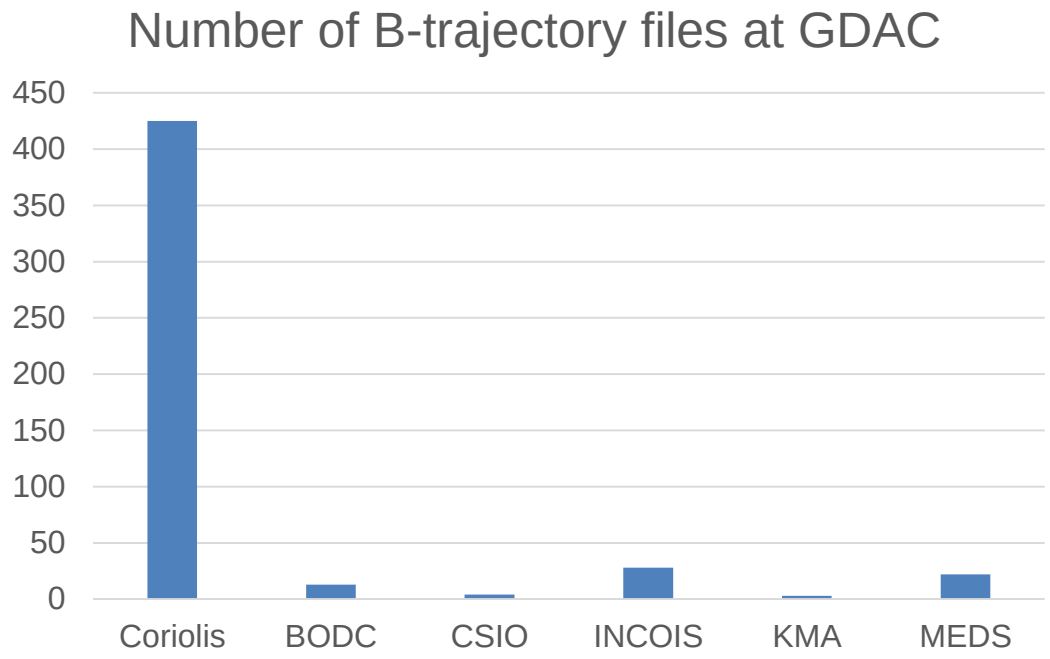
Combining c- and b-trajectory files

H. Bittig, T. Maurer, M. Scanderbeg, A. Wong, J. Nair and C. Schmid

Background

- Trajectory files were split into core and BGC following profile split
- Reasons for splitting profile files do not necessarily apply to trajectory files
  - File size
  - Stability of parameters
  - C- and B- file connection
- Working group formed after ADMT-19 has examined the issues and suggests that trajectory files NOT be split into c- and b- trajectory files, but keep one combined trajectory file



Total number of b-trajectory files: 495
Total number of BGC-Argo floats: 1429

File size

- For profile files, the difference in vertical sampling schemes between core and BGC sampling means that the profile files often have N_PROF > 1, with lots of white spaces to fill up the rectangular (N_PROF, N_LEVELS) netCDF structure. Separating into c- and b-profile files reduces the file size to some degree and makes it easier for users who only want CTD data
- This reasoning does not apply to traj files because there is mainly the N_MEASUREMENT dimension. So for all fields, there is only one way to grow which is along N_MEASUREMENT which does not create a lot of whitespace like N_PROF. The only exception is the i-parameter UV_INTENSITY_NITRATE which has an extra dimension (typically N_VALUES42) to accommodate spectral information. This N_VALUESxxx dimension is confined to UV_INTENSITY_NITRATE, is not used in any other fields and does not affect the size of the traj file.

Stability of parameters

- Network-wide reprocessing of raw <BGC PARAM>s have happened once for CHLA and once for backscattering. In future, reprocessing of raw <BGC PARAM>s can still be expected, but the period of stability should increase from hereon. Hence the raw parameters in a combined c- and b- traj file should remain fairly stable.
- The < BGC PARAM>_ADJUSTED fields could be updated more frequently than the <core PARAM>_ADJUSTED fields in a combined c- and b- traj file. This could be problematic. See later point on version control of combined Dtraj.

C- and B- file connection

- Because the N_MEASUREMENT dimension is the unambiguous link between c- and b- traj files, DACs need to be able to create a trajectory file that contains measurement codes for ALL types of events: core AND BGC.
- Therefore, to comply with v3.1, DACs need to create the basic structure of a combined trajectory file minus the BGC parameters.

Proposal to combine c- and b- trajectory files

- ***File size*** is not a reason for trajectory files to be split into c- and b-traj files due to the N_MEASUREMENT array structure
- ***Parameter stability*** is not a big factor for the Rtraj files since processing of raw <BGCPARAM> should remain fairly stable going forward.
- ***C- and B- file connection:*** DACs need to understand what MCs to include in the trajectory file even if kept as c- and b-traj files
- Allows users to look at one combined trajectory file to find all information – may still need to look at D and R trajectory files, but two is better than three or four if not combined
- Less files to keep track of, host, and store at GDACs

Suggest calling this combined traj file Version 3.2

- Change TRAJECTORY_PARAMETERS(N_PARAM,~~STRING16~~STRING64) (as in B-traj)
- Add TRAJECTORY_PARAMETER_DATA_MODE (N_MEASUREMENT, N_PARAM) – include P/T/S
- Add SCIENTIFIC_CALIB (N_PARAM, N_CALIB, STRING256)
- If approved, opportunity to add other desirable variables to traj file
  - Add JULD_DATA_MODE(N_MEASUREMENT) to make JULD adjustments clearer
    - JULD is different than other params in Argo because there can be FillValue in JULD in real time, but data in JULD_ADJUSTED for the same event to indicate a time is estimated and does not come directly from the float
    - Can make it hard to tell if JULD_ADJUSTED has actually been adjusted or is an estimate when file is in A/D mode.
  - Position error?

Difficulties

Version control for combined Dtraj

- Coordination would be needed between core and BGC groups since only one Dtraj file would exist
- Same as within BGC groups

For combined Rtraj? Please think about it and tell us now!

If endorsed, how to execute transition?

1. Finalize combined trajectory file format Version 3.2.
2. Seek comment and approval from AST in March 2020.
3. Update User Manual and begin creating combined trajectory files.
4. Coordinate with GDAC file checker for Version 3.2
5. Begin educating users on trajectory file changes.
6. There needs to be a transition period for BGC floats when both Version 3.1 (separate c- and b-traj) and 3.2 (combined traj) can exist.
   Floats without BGC parameters can stay in Version 3.1.