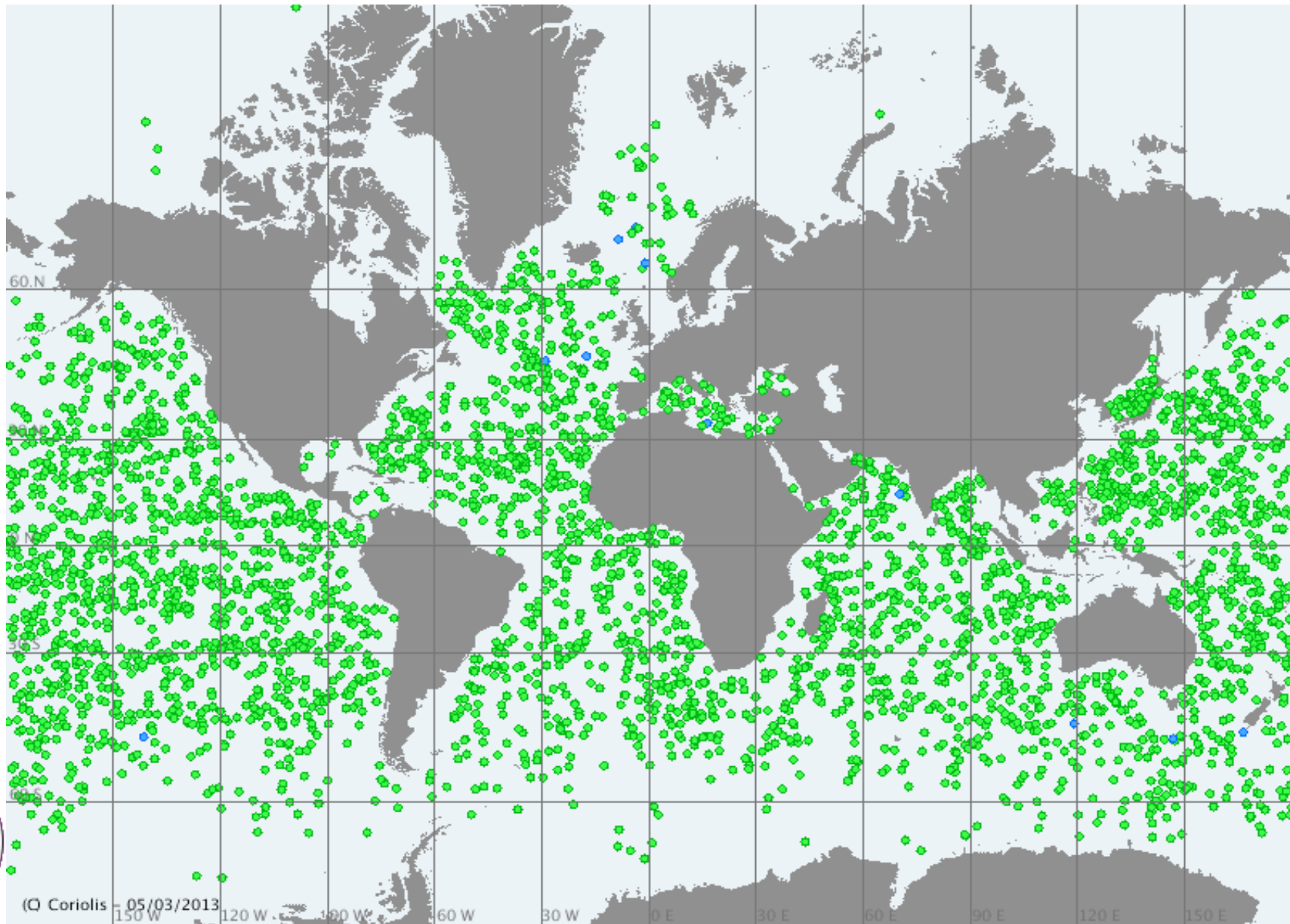


Feedback from ADMT13

S Pouliquen / Ann Gronell-Thresher
AST14

14-17 March 2011

14th Argo Steering Team meeting 2013



Updates at ADMT

- Change in ADMT Chairs : A Gronell Thresher/CSIRO is now co-chair of ADMT together with S Pouliquen/IFREMER and has spent the year settling into the role
- Note: All formats have been updated to CF compliance and the new version number for all 4 file types is 3.0
- We are continuing phone hook-ups with DACS between plenary meetings:
 - A lot of the actions are still finalized just before AST or ADMT, often in a rush
 - An increased number of deadlines often helps DACs focus on the actions
 - ADMT-Chairs have organized 2 phone meetings last year and another this year to review the action status. These were held in January and June 2012 and March 2013
 - March meeting focused on the actions due to AST14
 - Most are on track but CSIO/NMDIS and KMA have not responded



- Current ADMT efforts focus on
 - Further reducing the delays
 - Delayed-mode QC
 - Improving data consistency and completeness
 - Audited technical files, Density inversion tests, Pressure offset corrections,
 - Monitoring data quality through OA and altimetry analyses
 - Detection and correction of systematic errors
 - Improving trajectory formats and data quality
 - Helping Bio-Argo work towards compatibility with the Argo core data stream



Real Time Data Stream

- Data management infrastructure is in flux with Coriolis functional and GODAE suspended
 - USGDAC was understaffed – Mark was given additional duties and no longer had the time to devote to the Argo data server
 - Further, Mark had a personal tragedy and,
 - During the summer, Coriolis accidentally began sending multi-profile files to GODAE and other DACs were waiting to do the same
 - GODAE's format checker had not been updated so these files were rejected.
 - GODAE was no longer in sync with Coriolis – users could not get the complete dataset from this source.
 - Thierry offered to restore delivery of the original version files but got no answer
 - Thierry then sent an update to the format checker that would allow the multi-profile files but it was not installed





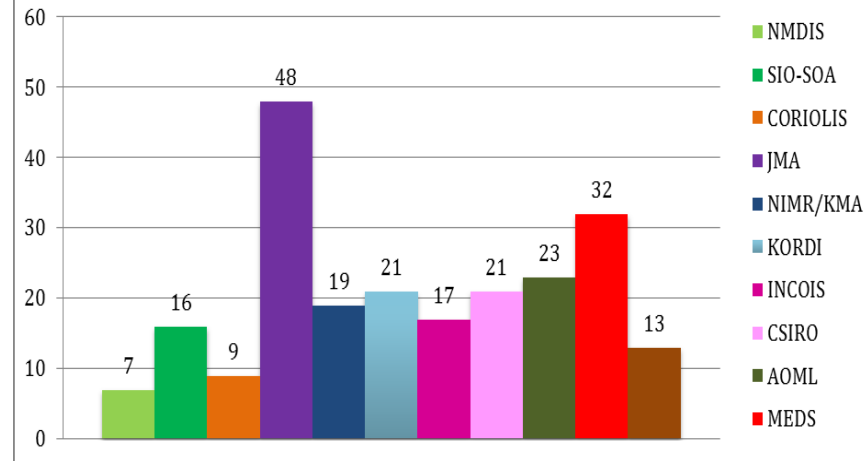
- With no backup at GODAE, and staff from Coriolis on leave, neither GDAC could react to the situation as it developed. By the time we realized there was a major issue, it was broken beyond immediate repair.
- ADMT decided that DACs should stop sending their data to GODAE until this was resolved.
- ADMT also decided that GODAE should turn off their format checker and accept data only through synchronization from Coriolis, while Coriolis would not allow synchronization in the reverse direction
 - This has now been done and the Coriolis format checker installed
 - File removals are still not working at GODAE
- This would maintain the functionality of GODAE while protecting users from the consequences of GODAE rejecting large numbers of profile files.
- GODAE was provided with a list of steps necessary to return to full function

Real Time Data Stream - GDAC

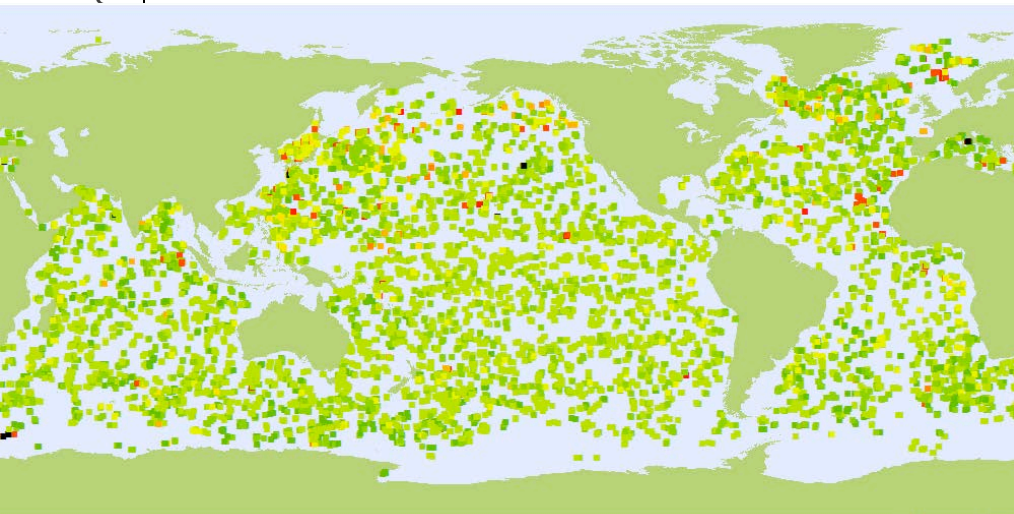
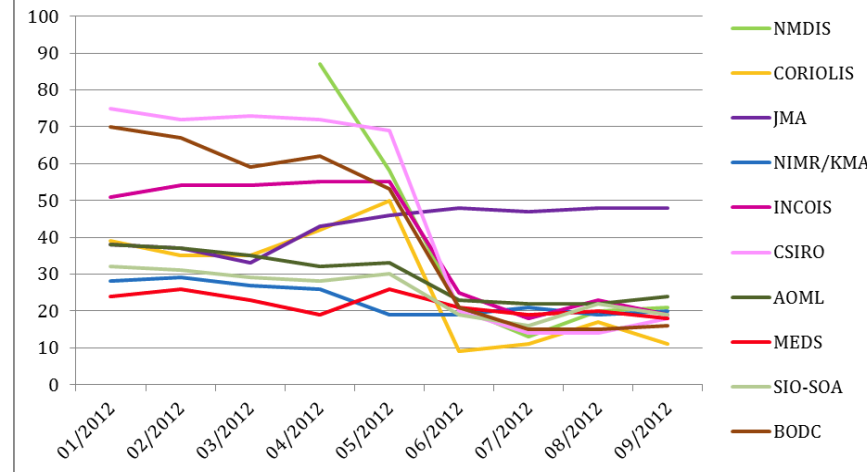
management 2012 Hyderabad India

- Monitoring data delivery to the GDACs:
 - Delay monitoring has been continued by AIC using French GDAC
 - Recent changes have improved the results: avg Feb 2013 delay is now 18 hours; compare to 39 hours in 2011
 - Large drop in delay in June – cause unknown but probably a Coriolis GDAC update

GDAC Average Delays, by DAC
October 2012



GDAC Average Delays, by DAC
2012



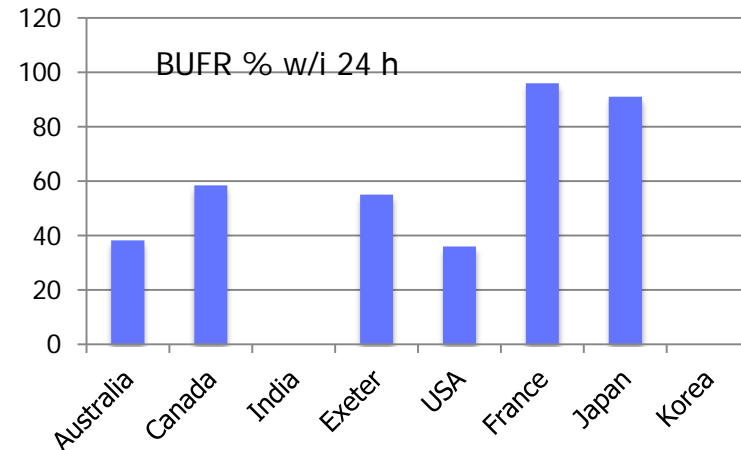
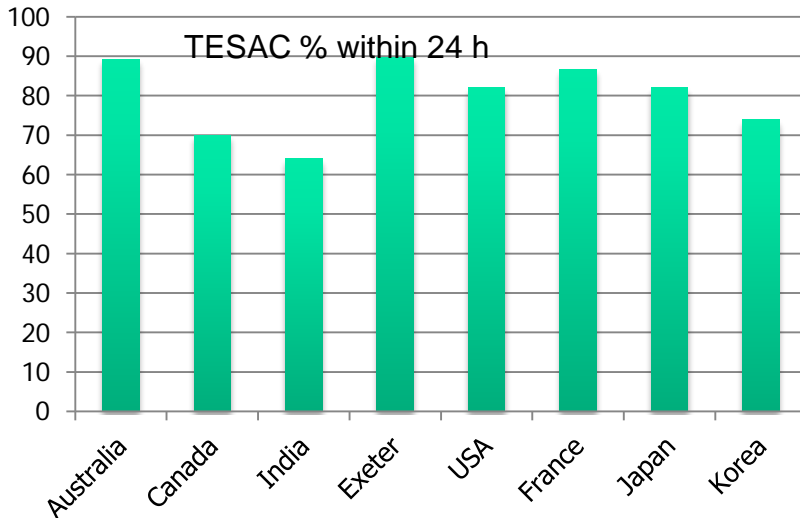
GDACs Delays
 $\mu=18$ hours $me=14$
 11365 Obs.

Errors	7 - 12h	25 - 48h	73 - 96h	121 - 240h
< 6h	13 - 24h	49 - 72h	97 - 120h	> 240h

February 2013
 jcomm ps

Real Time Data Stream - GTS

- 83% of data on GTS within 24 hours – this is down from 88% last year – improved to 89% in Feb
- Continuing transition to BUFR data transmission – 75% of the GTS data also arrives via BUFR (this could be higher – some headers might be missed)
 - Most DACs now transmit data in BUFR
 - Nov - 47% of the BUFR RT data arrived w/I 24 hours - significantly delayed compared to TESAC data – extra processing steps?
 - Large improvements – in Feb, 64% Iridium BUFRs and 59% Argos BUFRs arrived within 24 hours



- Continuing manual QC checks
 - Daily: Coriolis performing an OA to detect anomalies and send “standardized” feedback to DACs. Dacs then provide feedback to Coriolis (is the data truly good or bad?) to improve the system and decrease false alerts
 - Monthly: Coriolis sends a report to the DACs summarizing all OA anomalies detected
 - Results are good with <150 failures per month
 - DACs are correcting profiles and providing feedback
 - Some DACS not populating adjusted fields correctly
 - Only P Adjusted is filled
 - No adjusted fields when float requires correction for SPO



- Quarterly: Comparison to altimetry : A summary is produced and sent to all DACs. Individual messages are sent by AIC when the same anomaly is present in two consecutive runs without any action from DACs/Pis. Most of the DM operators are now responding to these reports and fixing their anomalies
 - Data quality better this year than last.
 - A new test to compare SLA and DHA differences helps detect floats that should be grey listed
 - also detects inconsistent corrections between D files and R files when D correction is not applied to later R files
- Floats should only be on the grey list until they are dead and have finished DMQC. Mathieu will monitor this and alert DACs when floats should be removed from the list



- Improvement in procedures
 - Test on density inversion that takes into account a threshold
 - Implemented after ADMT12 – progress very good but some large density inversions remain both in RT and DM data. Reports have been sent to the DACs so they can be investigated
 - TNDP floats – some still not being flagged correctly but all DACs are working towards this – problem decreasing as floats get older and disappear
 - Surface Pressure Corrections still a problem for some DACs in RT (CSIO, KMA), though most D files are compliant
 - Agreed all adjusted fields will be filled if SPO = 0 and whenever any one variable is adjusted



General Data Management

- Data formats will all be CF compliant – now version 3.0
- We can adopt ACDD conventions for data discovery which would make it easier for some systems to use Argo data – to be investigated. The necessary attributes could possibly be populated by the GDACs from the file contents requiring no action from the DACs
- Multi-profile files are ready to be delivered by many DACs
- Metadata files are still in flux
- Trajectory file progress – to be discussed in detail later
 - Trajectory format now to hold all information from the ANDRO DEP files so no need for the TRAJ 2 format and the conversion will be performed by Coriolis this summer
 - Trajectory files will now also be more consistent with the other Argo NetCDF formats
- DAC instruction/cookbook is well underway and will be an evolving document as new processes are required and defined



General Data Management

- Further data issues – post ADMT and not yet discussed in detail:
 - Data mode for secondary profiles – we need to dimension this variable (B King to propose solution to ADMT)
 - Instrument codes for new float types have been slow so data is flowing without a valid WMO instrument ID – we need a more efficient process for this



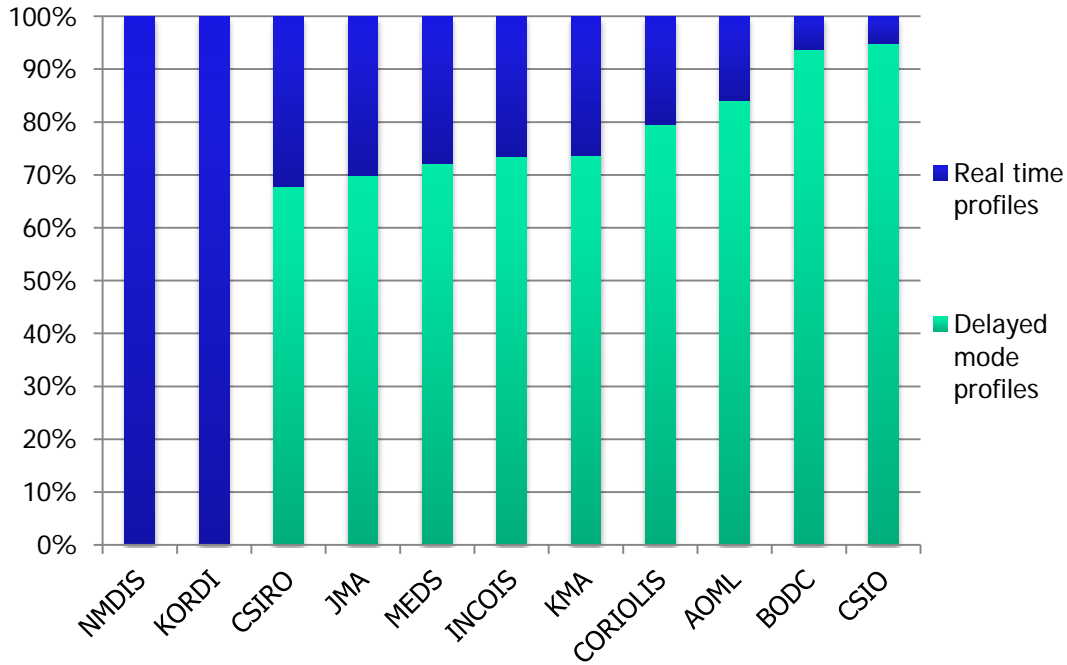
- Non Argo-Core issues
 - Near Surface Data needs to be identified and treated correctly
 - It may contain atmospheric measurements, unpumped data, pumped data for calibration...
 - Requires different QC
 - Unpumped and surface data will be carried in one of the additional profile dimensions
 - Users need to be educated about this data and its very specialized purpose.
 - Oxygen QC and data handling now agreed
 - 8 RT tests are recommended – detailed in the cookbook and QC manual
 - Variable names for the raw parameters are agreed
 - Raw variables will NOT have adjusted fields – only the derived parameters will require adjustment
 - Bio-Argo data handling and RTQC definition discussed at Bio-Argo workshop and actions identified



Delayed mode QC

13th Argo data management 2012 Hyderabad India

- Backlog as at ADMT – 21%
- 1.5% of the files have format errors
- Reducing the backlog has stalled for past 2 years



Histogram of DM profiles by DAC as of November 2012



GDAC status –

- Zipped DAC and GEO data files are now available and updated weekly at Coriolis
- The format checker developed by Mark at GODAE ran in test mode for several months with selected DACS. They provided feedback but we do not yet know when this will be operational.
- Detailed index files are updated every 6 hours at Coriolis (this will become hourly) to allow monitoring of delivery delays. Also to be implemented at USGDAC
- The GADR at NODC is serving RAW data in their multi-profile files, not adjusted – this needs to be fixed as soon as possible



- GDAC performance (Coriolis)
 - Data volumes grew by ~15% compared to 2011.
 - More than One million profiles are now available
 - In 2012, Coriolis server availability was 99,98% - representing 113 minutes of down-time in the entire year (longest 52 min)
 - File transfer time was significantly faster
 - Files are transferred from the DAC directories every 30 minutes.
- 1384 floats are on the master grey list – compared with 1181 a year ago



ARC activities

- **South Atlantic ARC:** Ran the consistency check on all floats passing through this region. Continue with product development including annual and semi-annual means of T, S and dynamic height, and maps of various parameters – all available to users through WWW.
- **MedArgo ARC:** Newly established for Mediterranean and Black Seas. Handles overall coordinating of float operations including deployments and production of Argo products. Also perform DMQC on the Argo data from this region. Have 151 floats with a maximum of 36 active at one time but new deployments will increase this soon.
- **North Atlantic ARC:** producing maps and gridded products as well as some time series plots. All available on the web. They also provide a data mining tool for the N Atl Argo data. While analyzing the OW software performance in the N Atl, they have shown that OW tends to over correct salinity in some specific regions – they are investigating the cause of this in order to improve the tools. To be discussed at DMQC5



ARC activities - con

- **Pacific ARC:** no report.
- **Southern Ocean ARC:** BODC is hosting the new data and information pages for this ARC, including products, maps and animated FOAM outputs. They aim to automated CTD data delivery to CCHDO for inclusion in the Argo reference database. There is also an ongoing effort to automate the POGO cruise planning database making it easier to identify potential sources of high quality CTD data.
- **Indian ARC:** working to archive IO data and making them available though a web-GIS, producing a DVD of Argo data for students and others with low bandwidth web connections, continuing DMQC on Indian floats, identifying and supplying CTD data from Indian research cruises for inclusion in the Argo reference database, held workshops to demonstrate the use of Argo data to researchers and generated value added products to be served on the web.



Remaining issue

- Argo Citations:
 - Last year discussed DOI (Digital Object Identifier)
 - Works well for our documentation but not for our data
 - Still preferred option but the data set is continually changing –
 - Do we freeze it with a new DOI at regular intervals?
 - Do we adopt an 'open time series' model so we can append data? Problem: we can't change data in this series
 - Both solutions require archiving multiple copies of the complete dataset, each with a unique DOI
 - Will also investigate 'minted DOI' we would then apply our own doi with an Argo identifier
 - doi:10.5285/**argo**9c42-4dfb-4da9-be97-c532ce13922
 - **Need clarification – what do we want from DOIs?**



Challenges:

- Data format volatility:
 - Only the technical file is stable
 - Metadata files are still under development
 - Multi-profile files are 'finished' but need to be slightly enhanced (mission number needs to be included)
 - Trajectory files are the biggest problem – very complex, require reprocessing of the hex, and understanding/implementation of the DM process – will not be quick
- MANPOWER – some DACs are falling behind in DMQC, programming the new formats, processing data from new float types – and new data such as oxygen and Bio-Argo. If the DACs are not adequately resourced, particularly for new data types, then we will need to prioritize.



- Action still pending since ADMT13 to AST chairs
 - AST to document the different issues that have affected the Argo data to inform users, e.g., pressure correction, micro-leak...

- DMQC 5 : will be prior to ADMT14 in Liverpool

- Next ADMT at BODC in Liverpool/UK – proposed dates
 - 14th-18th October or
 - 11th-15 November

