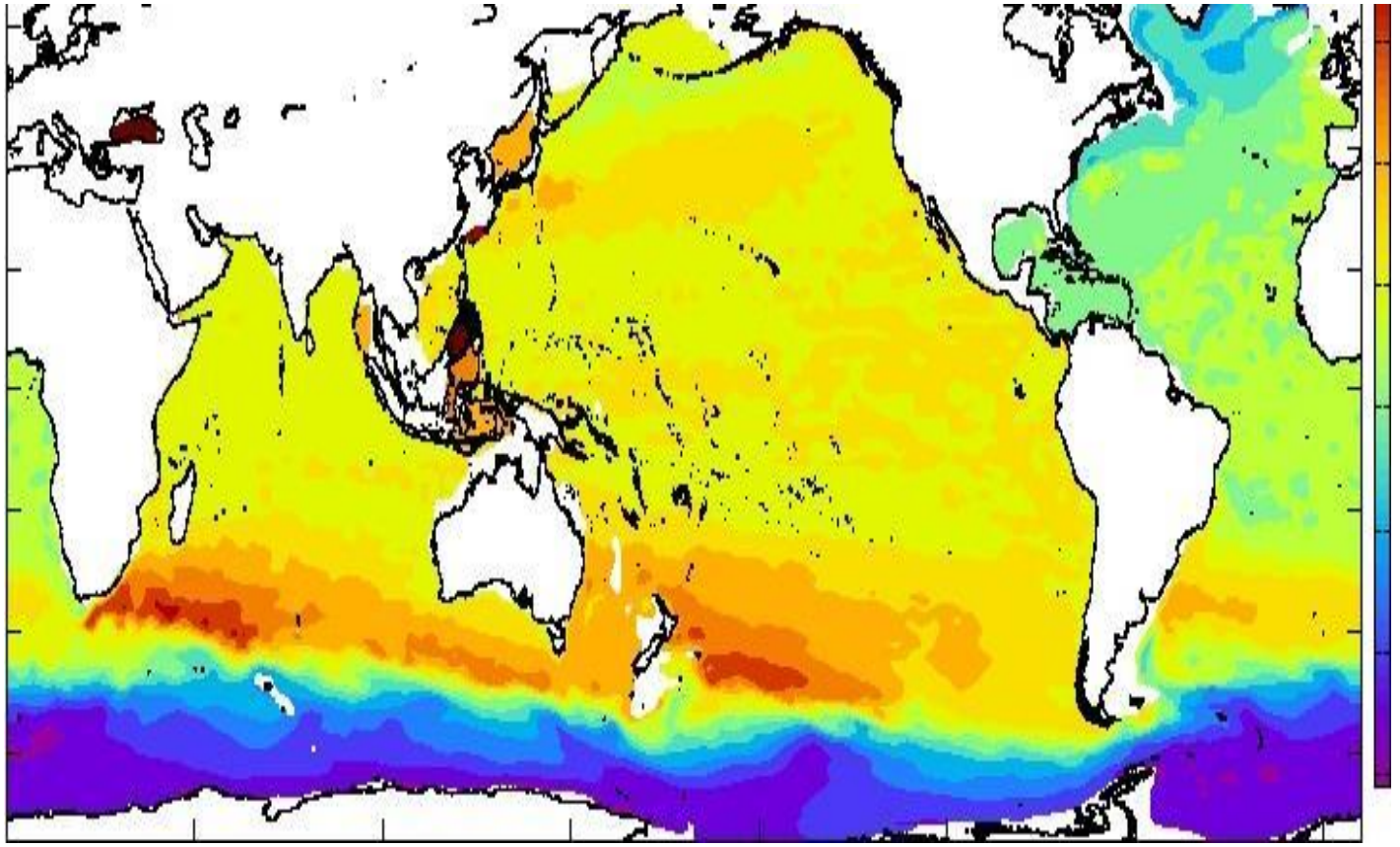
The CSIRO Ocean Data Archive (CODA) and CSIRO Atlas of Regional Seas v2 (CARSv2)

Chris Chapman, Thomas Moore
Rebecca Cowley, Ash Parker (CSIRO)

With contributions from Didier Monselesan, Bernadatte Sloyan, Chris Roach and many, many others!

The Original CARS - Product and Philosophy

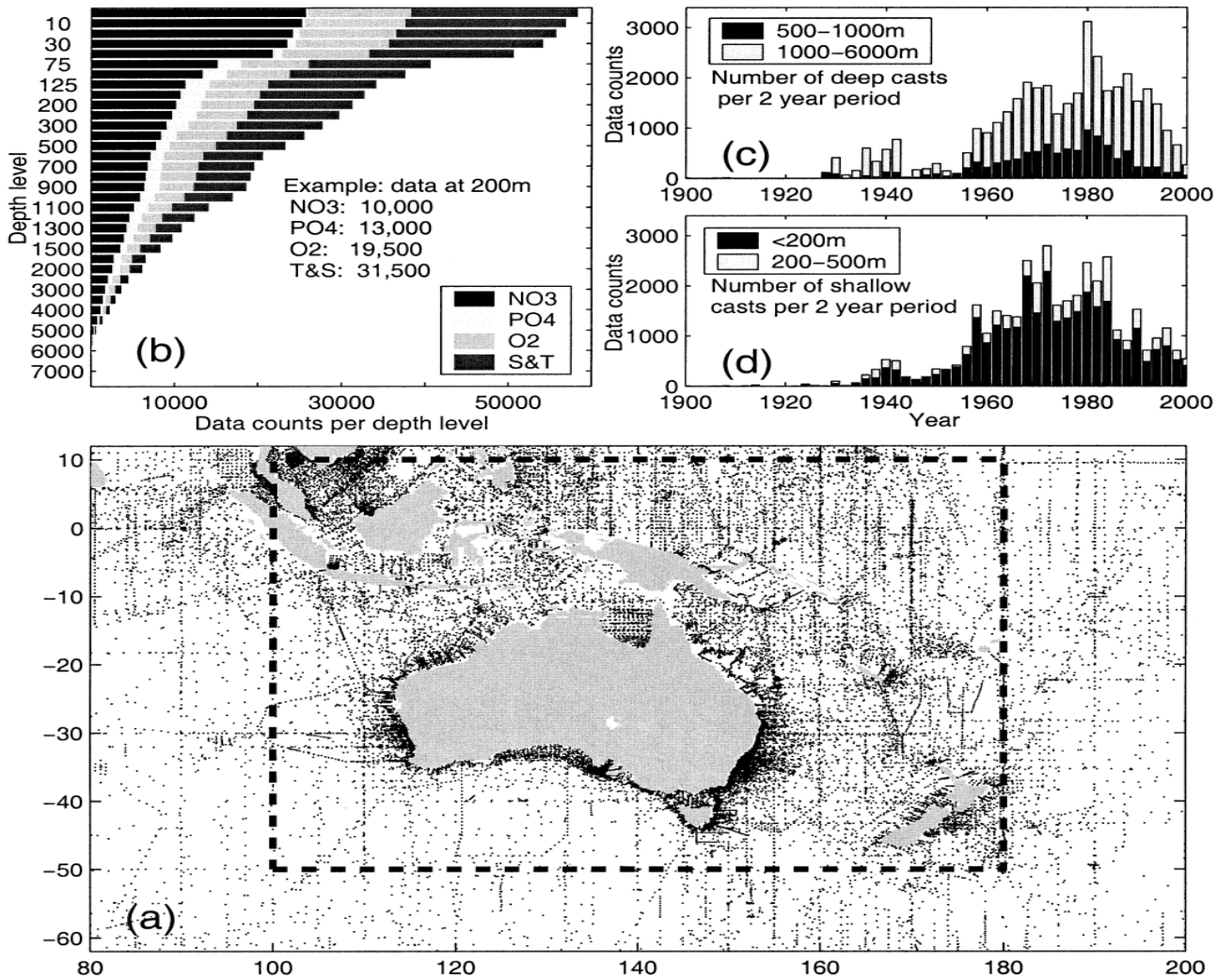
- An "atlas" of seasonal ocean properties;
- Gridded (½° x ½° and 1/8° x 1/8°) horizontal grid spacing;
  - Local least-squares fit;
  - Constraints on bathymetry and land barriers;
  - Attempts to improve *resolution* in areas of high data density;
- Derived from an "in-house" cultivated database of oceanographic observations (Bluelink Ocean Archive – BOA);
- **Extremely successful –** *Ridgeway et al. (2002)* cited 635 times (Google Scholar)



Source: CARS 2009, dynamic height

The Original CARS – where are they now?

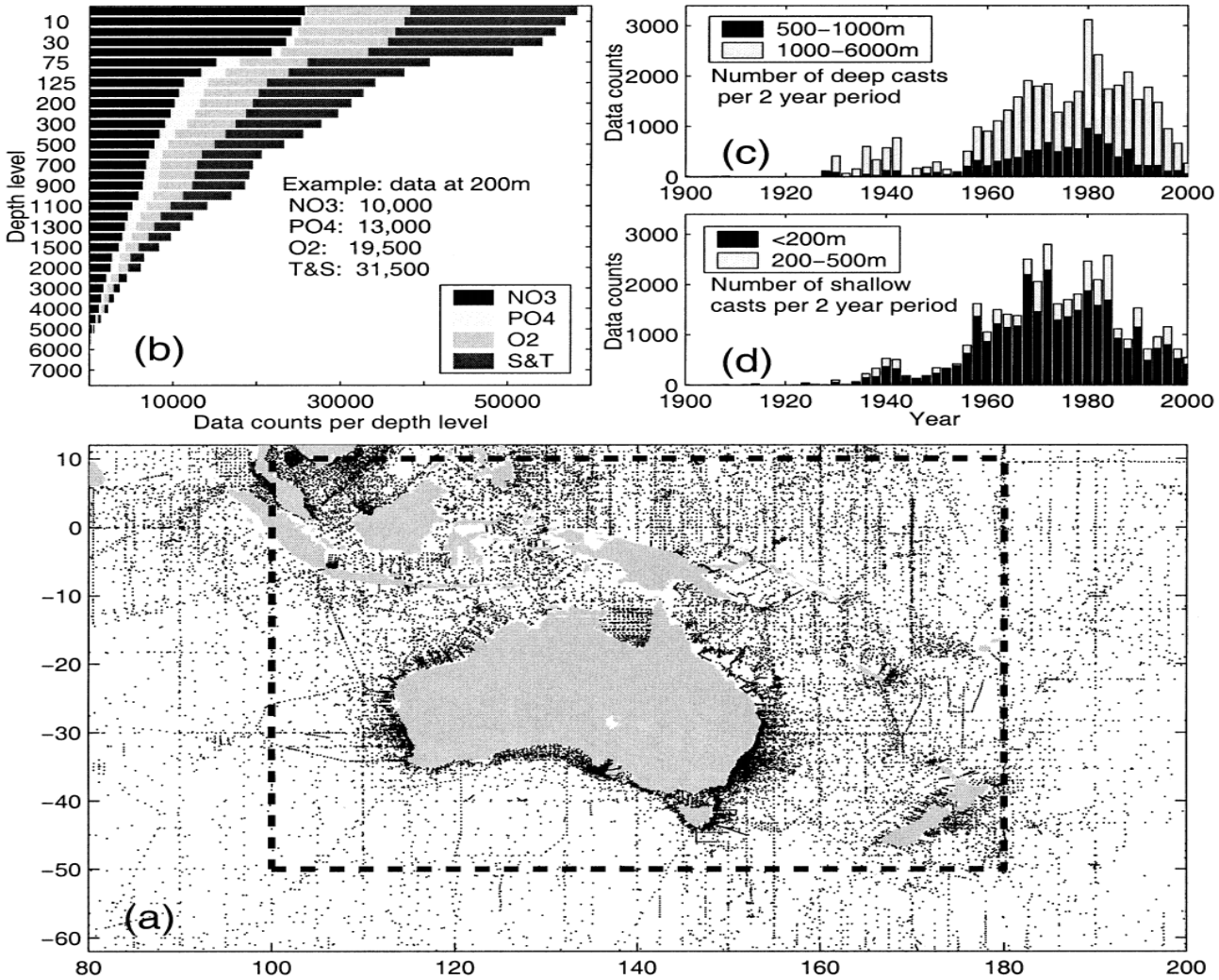
- Last comprehensive update - 2009;
- Loss of key personnel;
- Code base no longer maintainable;
- Lots of new obs and platforms since (ARGO, gliders, animal borne sensors, mooring programs, etc…)



Source: Ridgeway et al (2002) - data distribution of the original CARS product

The Original CARS – where are they now?

- Last comprehensive update - 2009;
- Loss of key personnel;
- Code base no longer maintainable;
- Lots of new obs and platforms since (ARGO, gliders, animal borne sensors, mooring programs, etc…)
- However – strong demand within the community for updated **an Australian Ocean Atlas** with accompanying **High Quality In-Situ Database**



Source: Ridgeway et al (2002) - data distribution of the original CARS product

What are we doing about it?
CARS2022 – design and philosophy

What are we doing about it?
CARSv2 – design and philosophy

- A near **global** product, with an **Australian focus**;

What are we doing about it?
CARSv2 – design and philosophy

- A near **global** product, with an **Australian focus**;
- **Best-in-class** data curation, quality control, data accessibility and gridding methodology

What are we doing about it?
CARSv2 – design and philosophy

- A near **global** product, with an **Australian focus**;
- **Best-in-class** data curation, quality control, data accessibility and gridding methodology
- Follow FAIR (Findable, Accessible, Interoperable, Reusable) principles and industry standard software engineering

What are we doing about it?
CARSv2 – design and philosophy

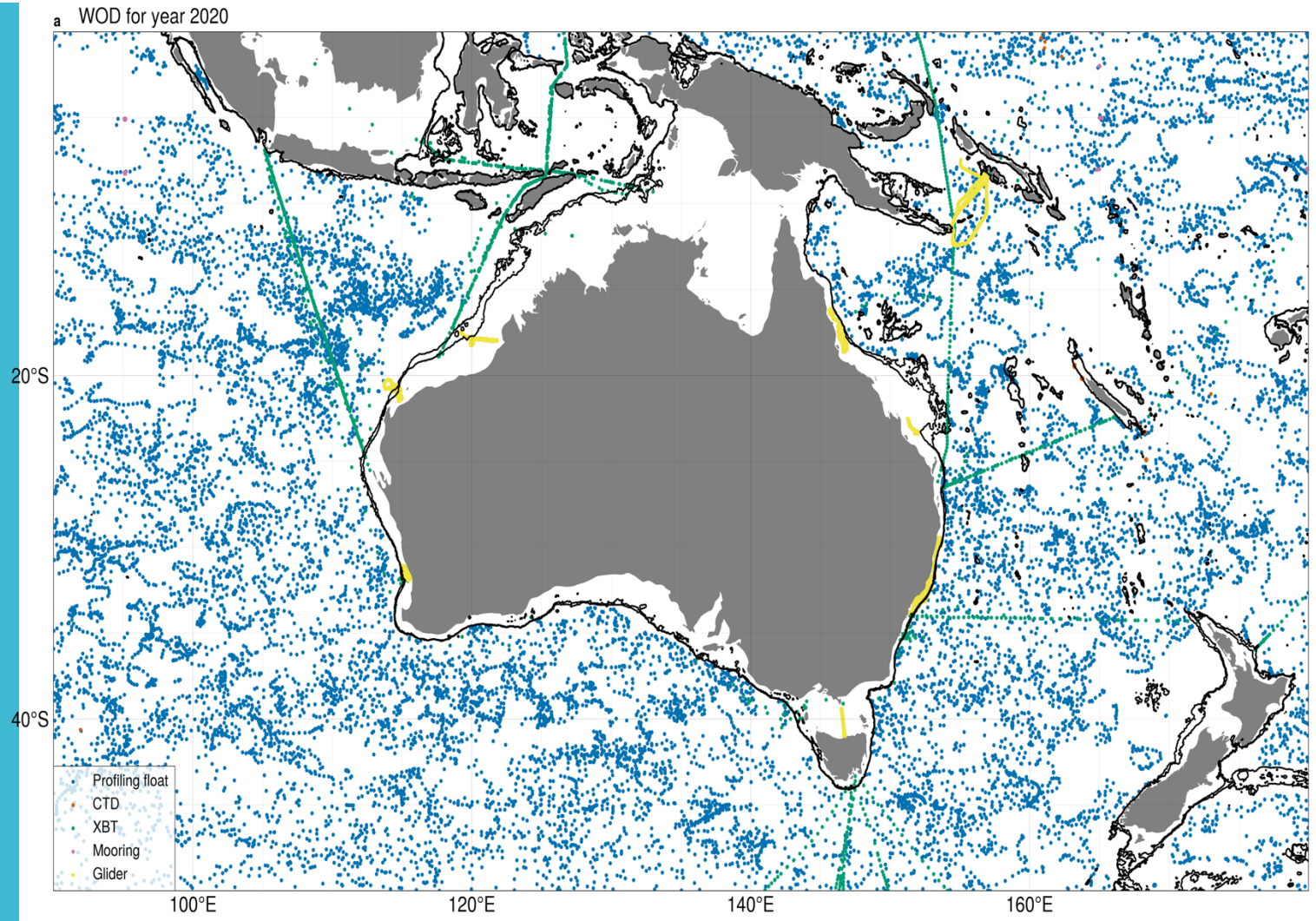
- A near **global** product, with an **Australian focus**;
- **Best-in-class** data curation, quality control, data accessibility and gridding methodology
- Follow FAIR (Findable, Accessible, Interoperable, Reusable) principles and industry standard software engineering
- Deliver two products:
  - The CSIRO Ocean Data Archive – complete database of in-situ oceanographic profiles with standard file format, metadata and quality control

What are we doing about it?
CARSv2 – design and philosophy

- A near **global** product, with an **Australian focus**;
- **Best-in-class** data curation, quality control, data accessibility and gridding methodology
- Follow FAIR (Findable, Accessible, Interoperable, Reusable) principles and industry standard software engineering
- Deliver two products:
  - The CSIRO Ocean Data Archive – complete database of in-situ oceanographic profiles with standard file format, metadata and quality control
  - The CSIRO Atlas of Regional Seas 2025 – a gridded climatology of ocean conditions around Australia

CSIRO Ocean Data Archive

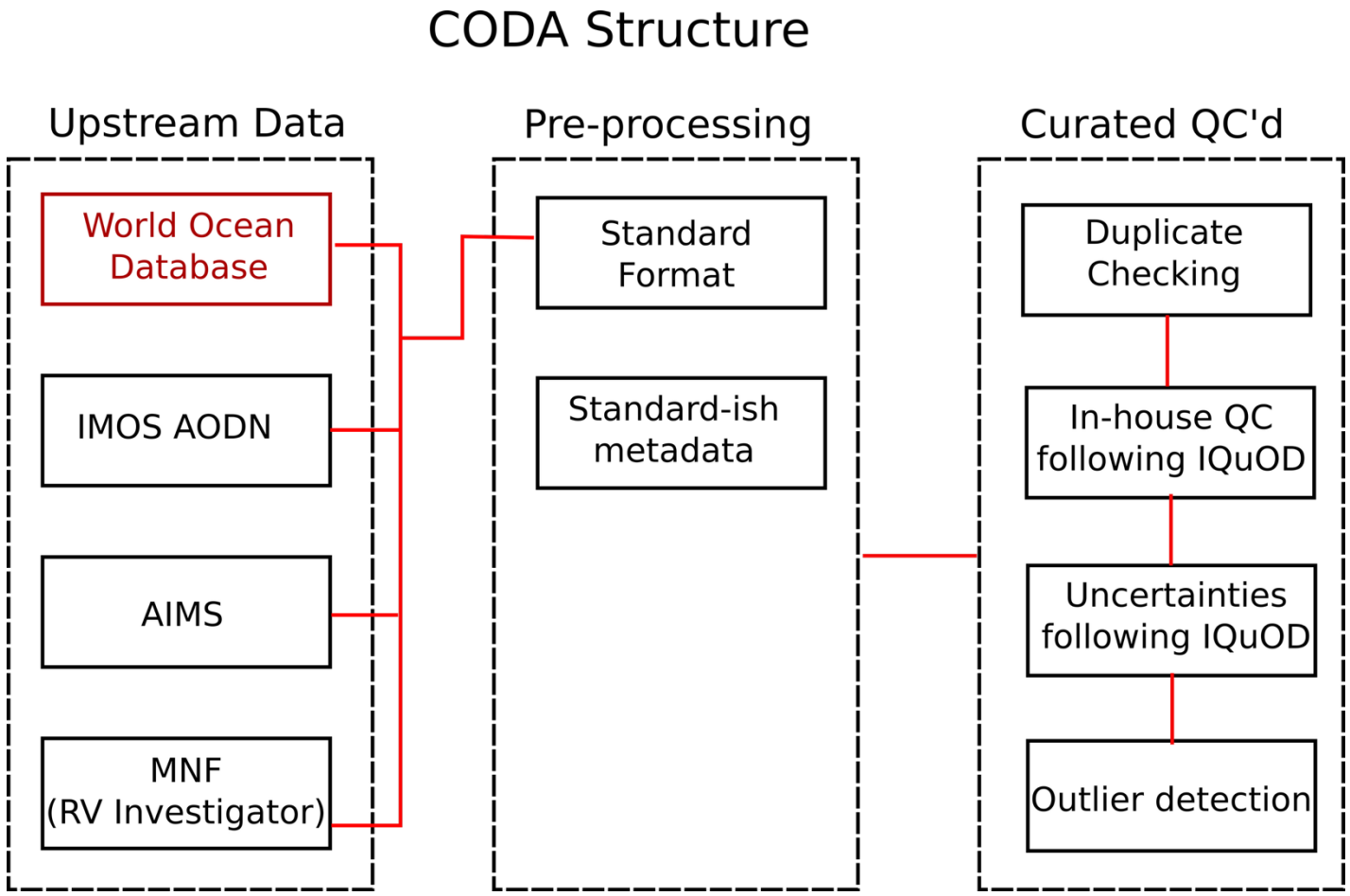
- Limited observations over shelf and shallow seas;



Source: Data from CODA (WOD2018 + AIMS + MNF + RAN/AODN)

The CSIRO Ocean Data Archive (CODA)

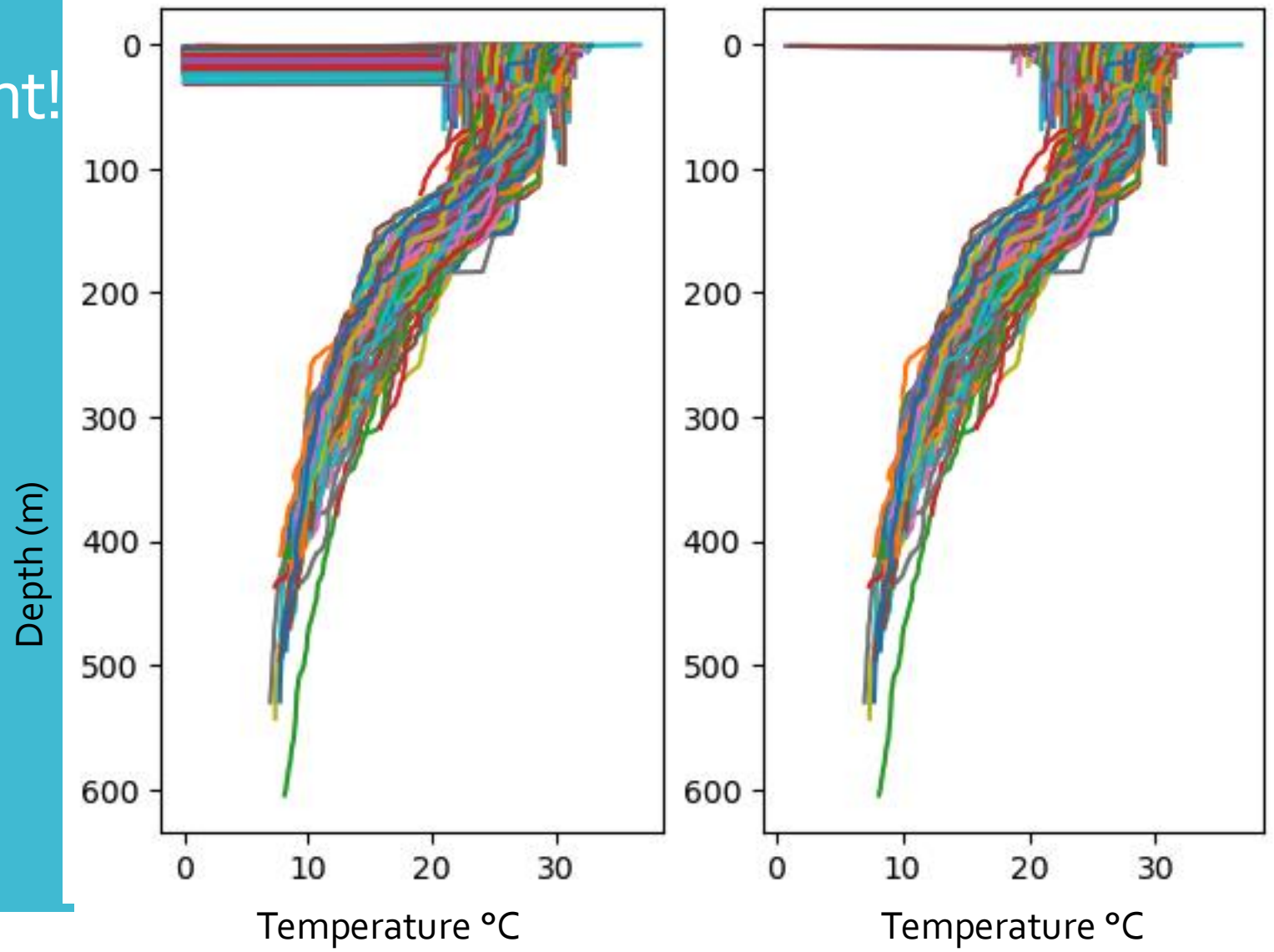
- Multi-platform database of ocean observations with an Australian focus
- Complement the global database (WOD) with "boutique" data;
- Data has a simple to read common file format(s):
  - Follows cf-conventions;
  - Standardised meta-data;
  - Netcdf and tabular (parquet) form (in progress);
- Automated, standardised, best-practice QC and uncertainties
- Potential use cases:
  - Research
  - Data-Assimilation



Data Quality is Paramount!

Simple example of CTD data with no QC flags included

- Remove the Temperature data with values of zero gives us the plot on the right



Data Quality

- Some of the data sourced is QC'd – some isn't;
- Even QC data may have very different standards/methods!
- Automated QC tools developed by IQuOD allow for ***scalable and unified apporach***
- Requires us to adapt the tools for file formats or transform the data recovered into a standard format for QC and data ingestion

www.iquod.org

IQuOD aims to maximize the quality, consistency and completeness of the long-term global subsurface ocean temperature database. IQuOD is a product served alongside the World Ocean Database (WOD).

IQuOD includes:

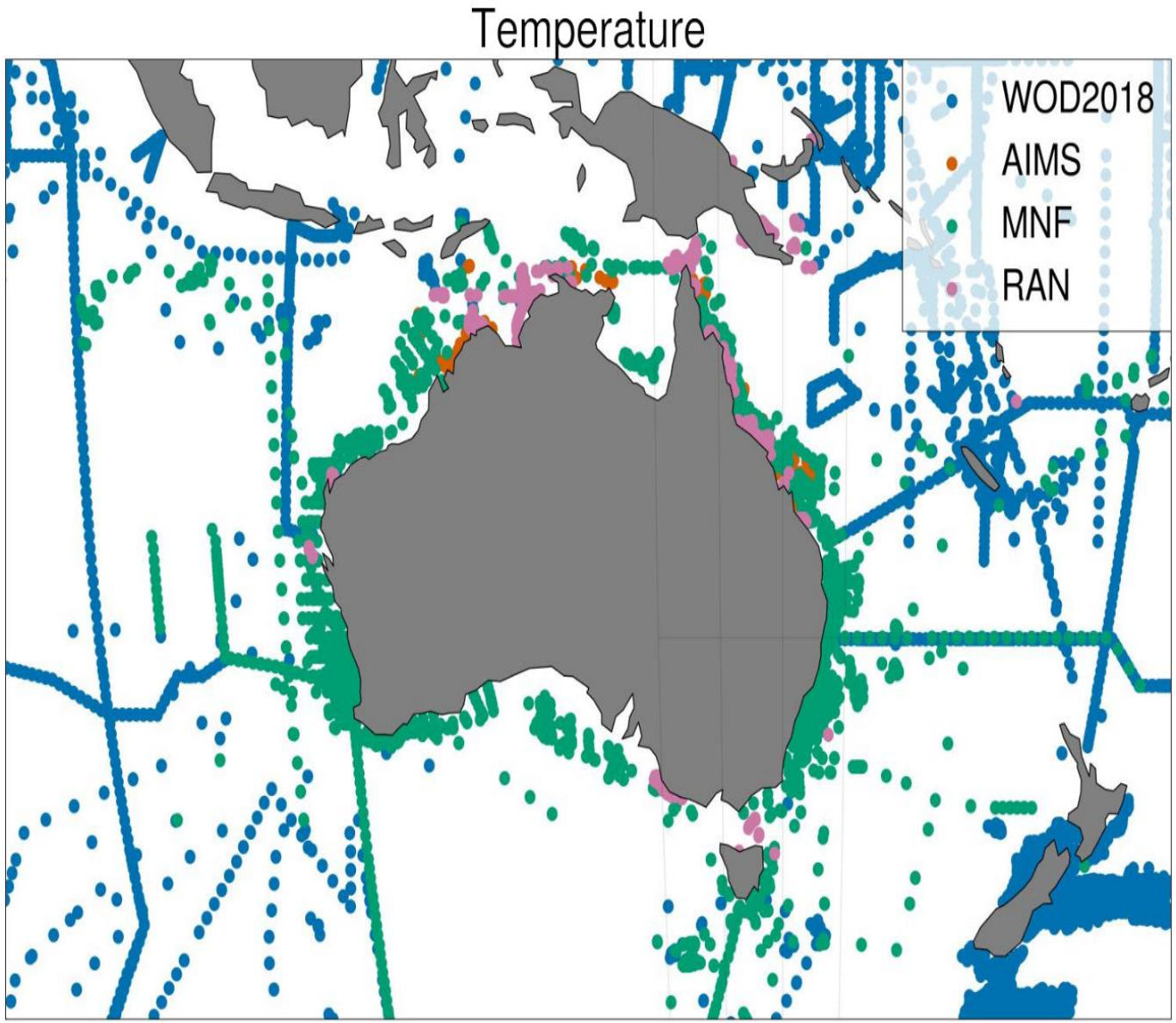
- Intelligent metadata for XBTs.
- Uncertainties assigned to each individual temperature observation.
- Some uncertainties assigned to depth and salinity.

IQuOD will soon include:

- Automated QC flags from the IQuOD community benchmarking tests.
- Duplicate checking tools.
- Machine learning tools implemented alongside expert/visual QC.

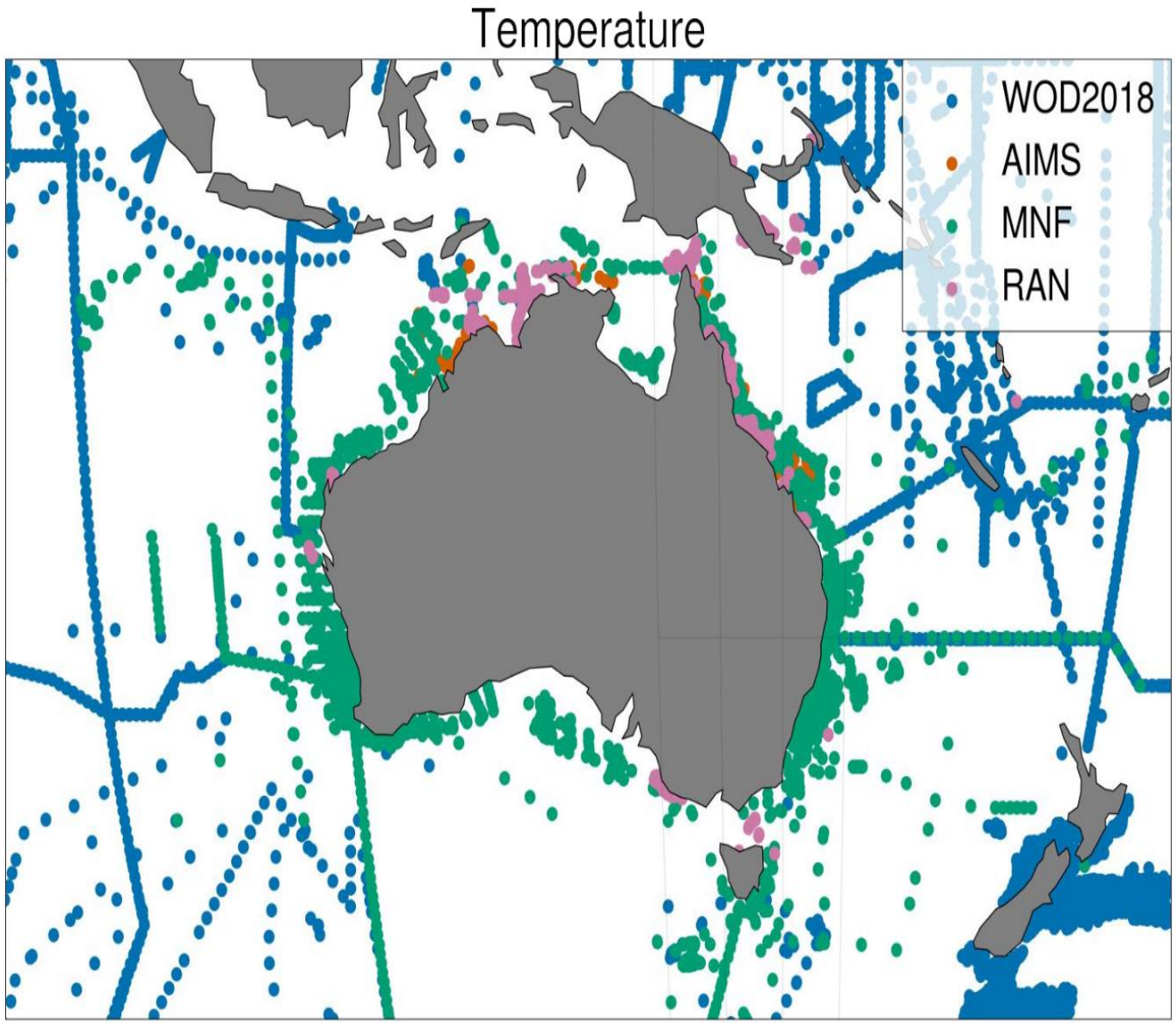
The CSIRO Ocean Data Archive (CODA)

- An **Australian product** for **Australian conditions;**



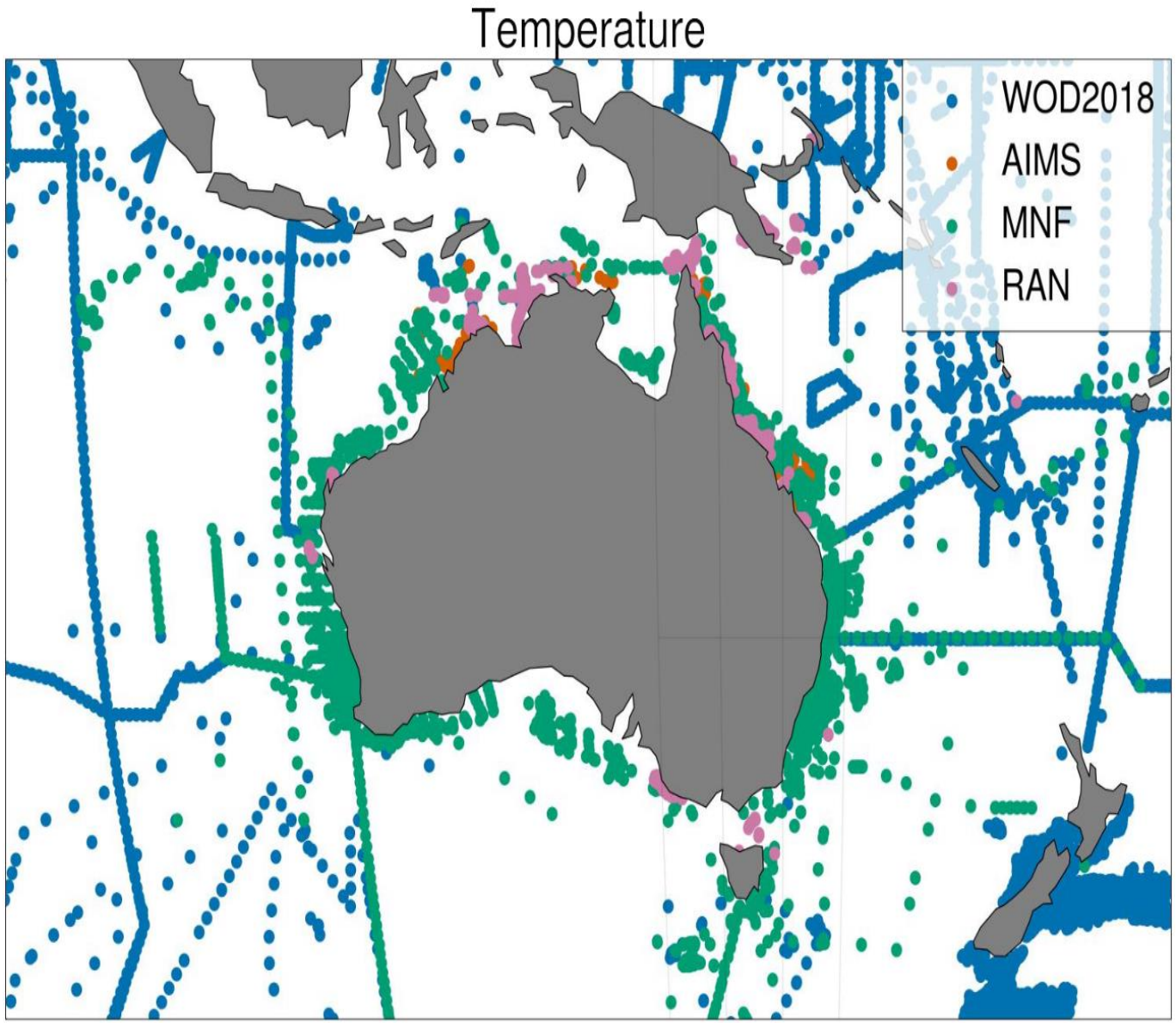
The CSIRO Ocean Data Archive (CODA)

- An **Australian product** for **Australian conditions;**
- Improved observational density over shelf and coastal region**;**



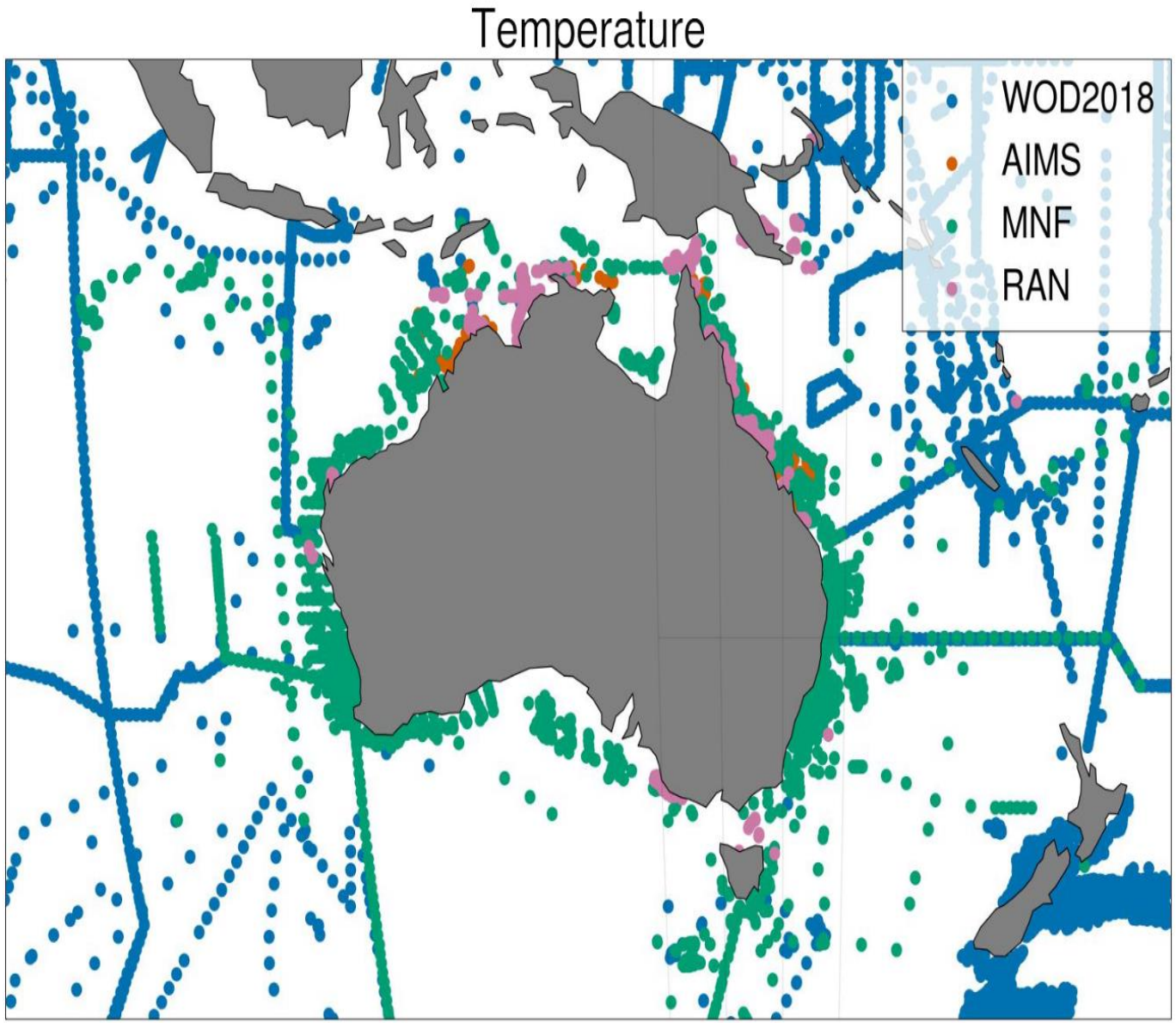
The CSIRO Ocean Data Archive (CODA)

- An **Australian product** for **Australian conditions;**
- Improved observational density over shelf and coastal region;
- Advanced duplicate checking;



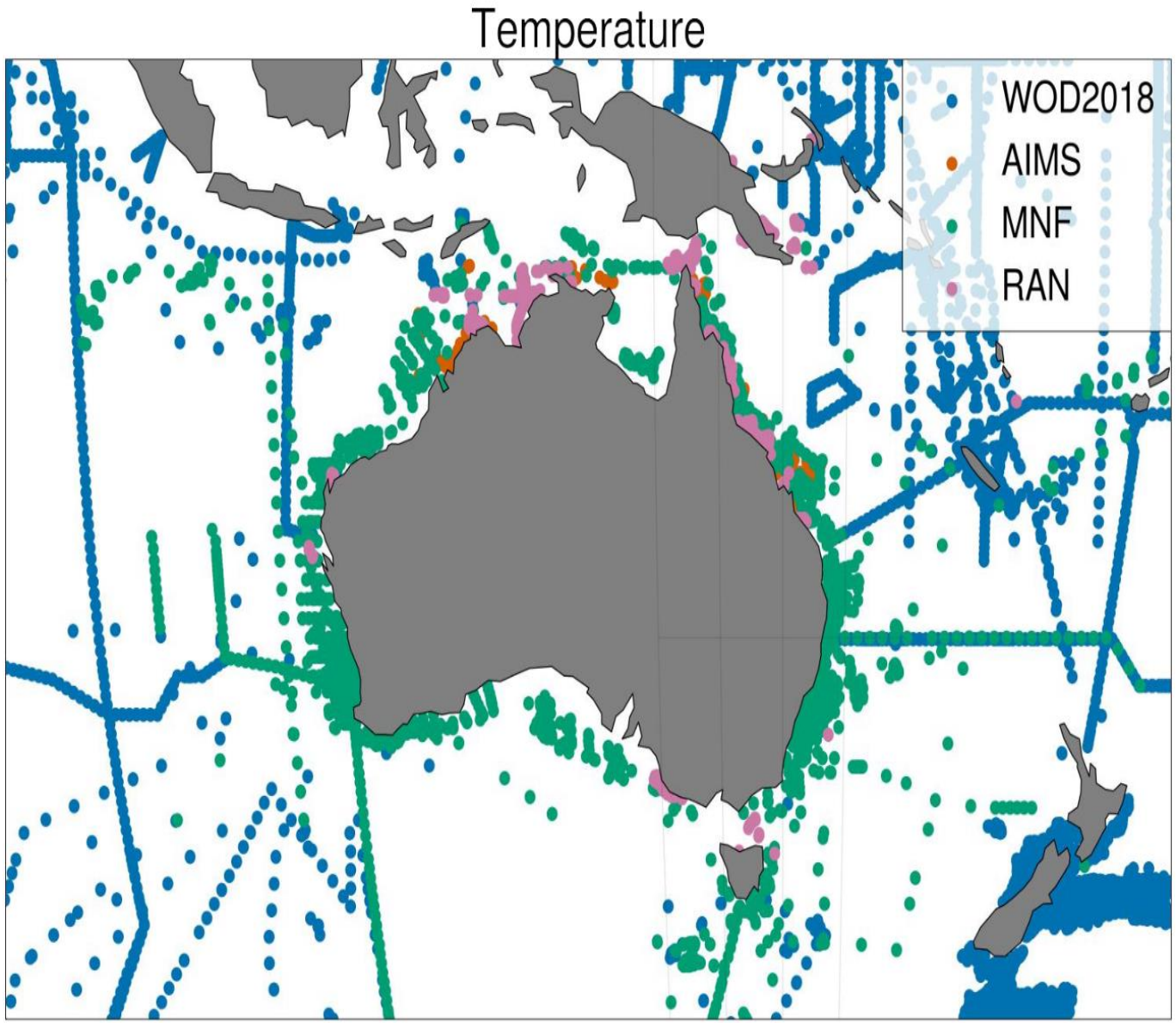
The CSIRO Ocean Data Archive (CODA)

- An **Australian product** for **Australian conditions;**
- Improved observational density over shelf and coastal region;
- Duplicate checking (in progress);
- Standard file format, standard-ish meta data, standard QC flags (with original flags carried through);



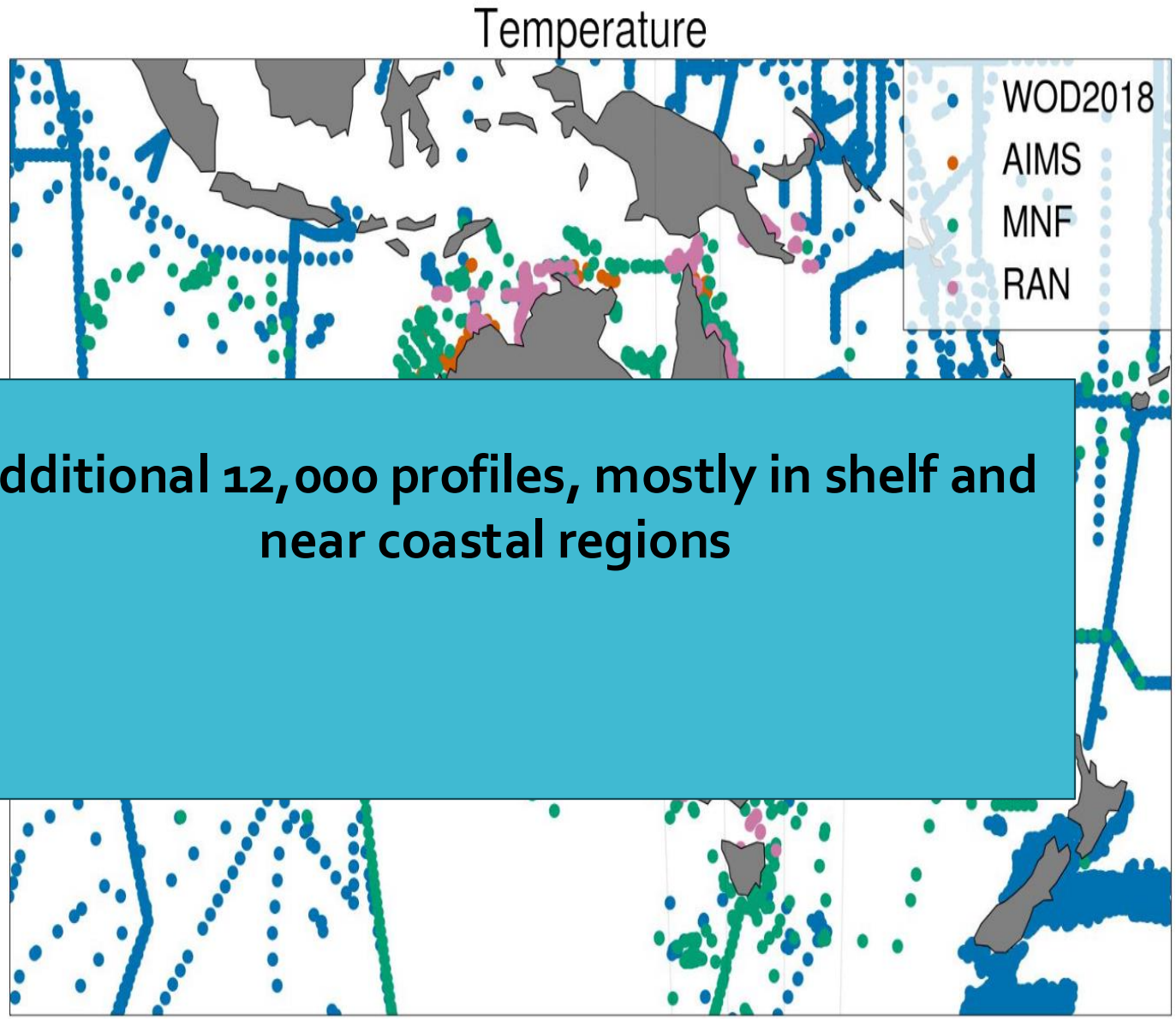
The CSIRO Ocean Data Archive (CODA)

- An **Australian product** for **Australian conditions**;
- Improved observational density over shelf and coastal region;
- Duplicate checking (in progress);
- Standard file format, standard-ish meta data, standard QC flags (with original flags carried through);
- Adding Observational Uncertainties (work in progress)



The CSIRO Ocean Data Archive (CODA)

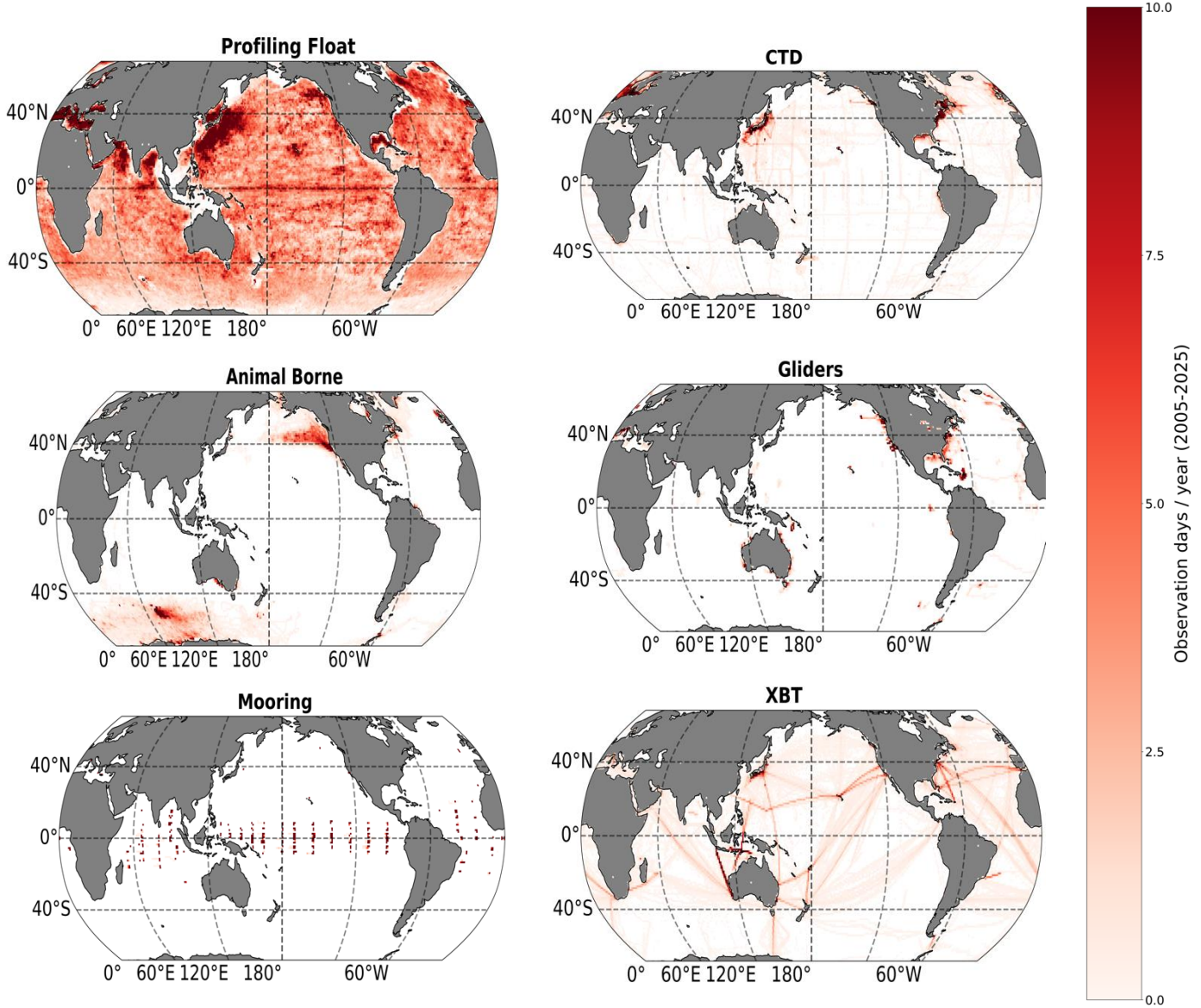
- An **Australian product** for **Australian conditions**;
- Improved observational density over shelf and coastal region;
- Duplicate checking (in progress);
- Standard file format, standard-ish meta data, standard QC flags (with original flags carried through);
- Adding Observational Uncertainties (work in progress)



Additional 12,000 profiles, mostly in shelf and near coastal regions

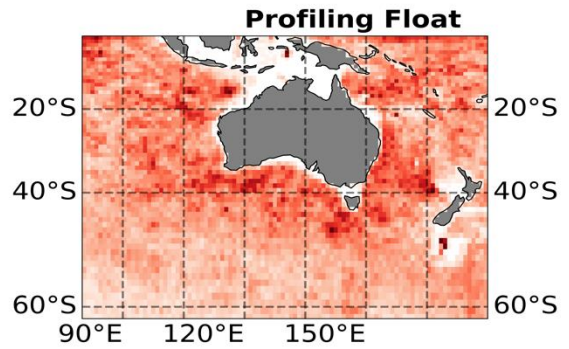
The CSIRO Ocean Data Archive (CODA)

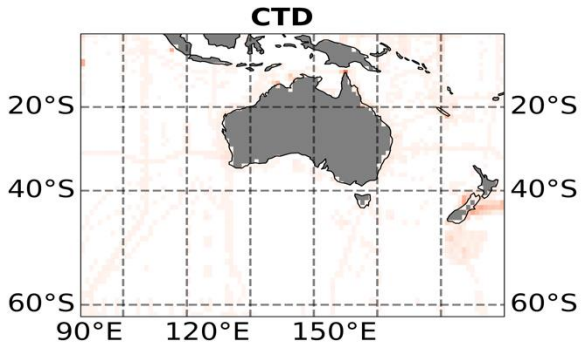
- No one observational platform delivers data everywhere;
- Poses challenges for climatologies and DA:
  - Different precisions and errors;
  - Inhomogenous sampling in space and time;
  - Biases from instrumentation and deployment strategies

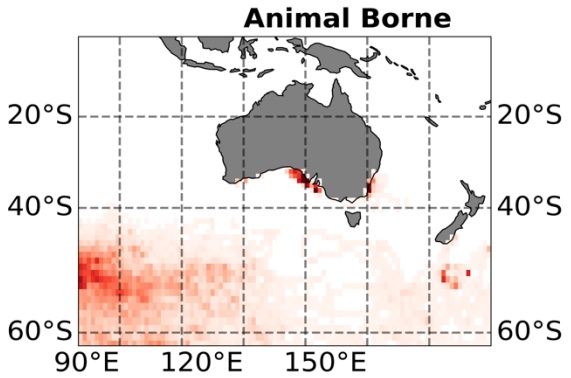


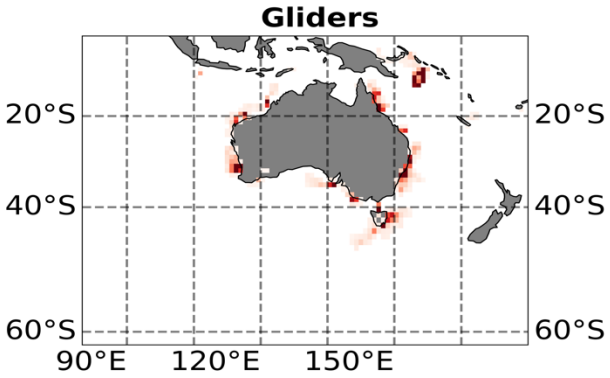
The CSIRO Ocean Data Archive (CODA)

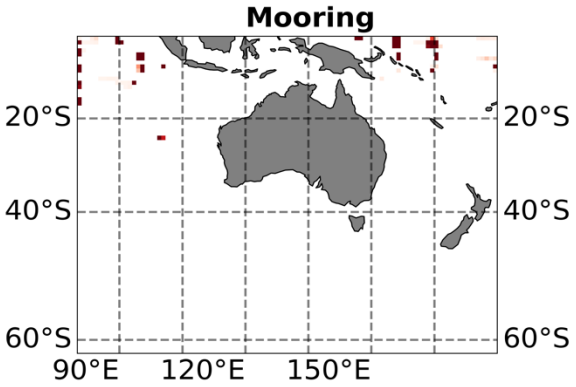
- No one observational platform delivers data everywhere;
- Poses challenges for climatologies and DA:
  - Different precisions and errors;
  - Inhomogenous sampling in space and time;
  - Biases from instrumentation and deployment strategies

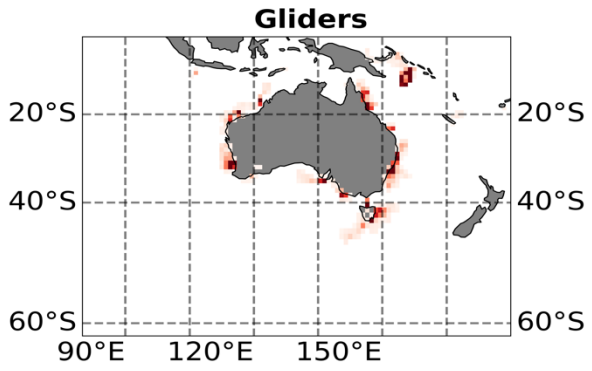












The CSIRO Ocean Data Archive (CODA)

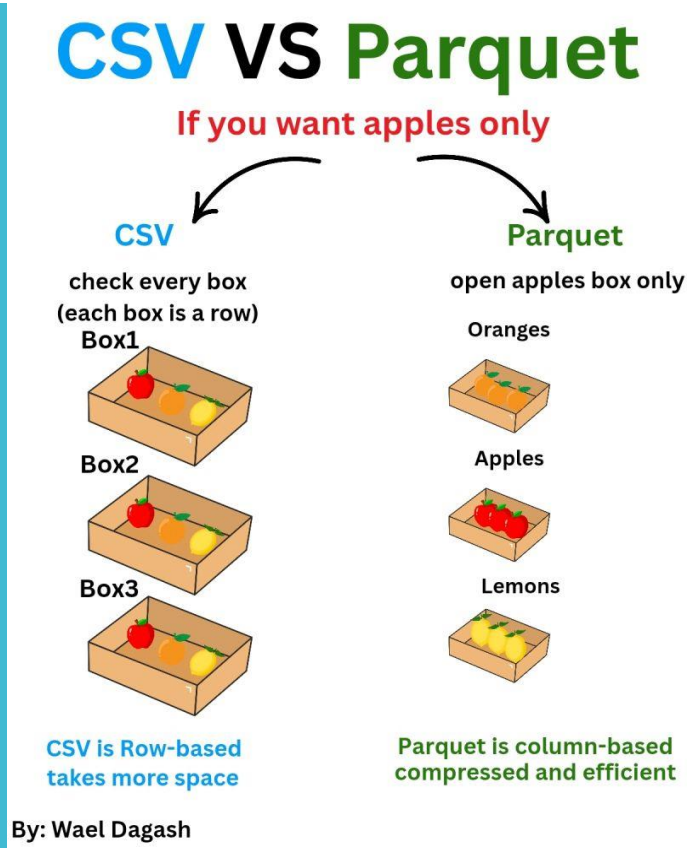
- Wrangling data from various archives can be a nightmare

The CSIRO Ocean Data Archive (CODA)

- Wrangling data from various archives can be a nightmare
- We've developed a set of tools to make that easier

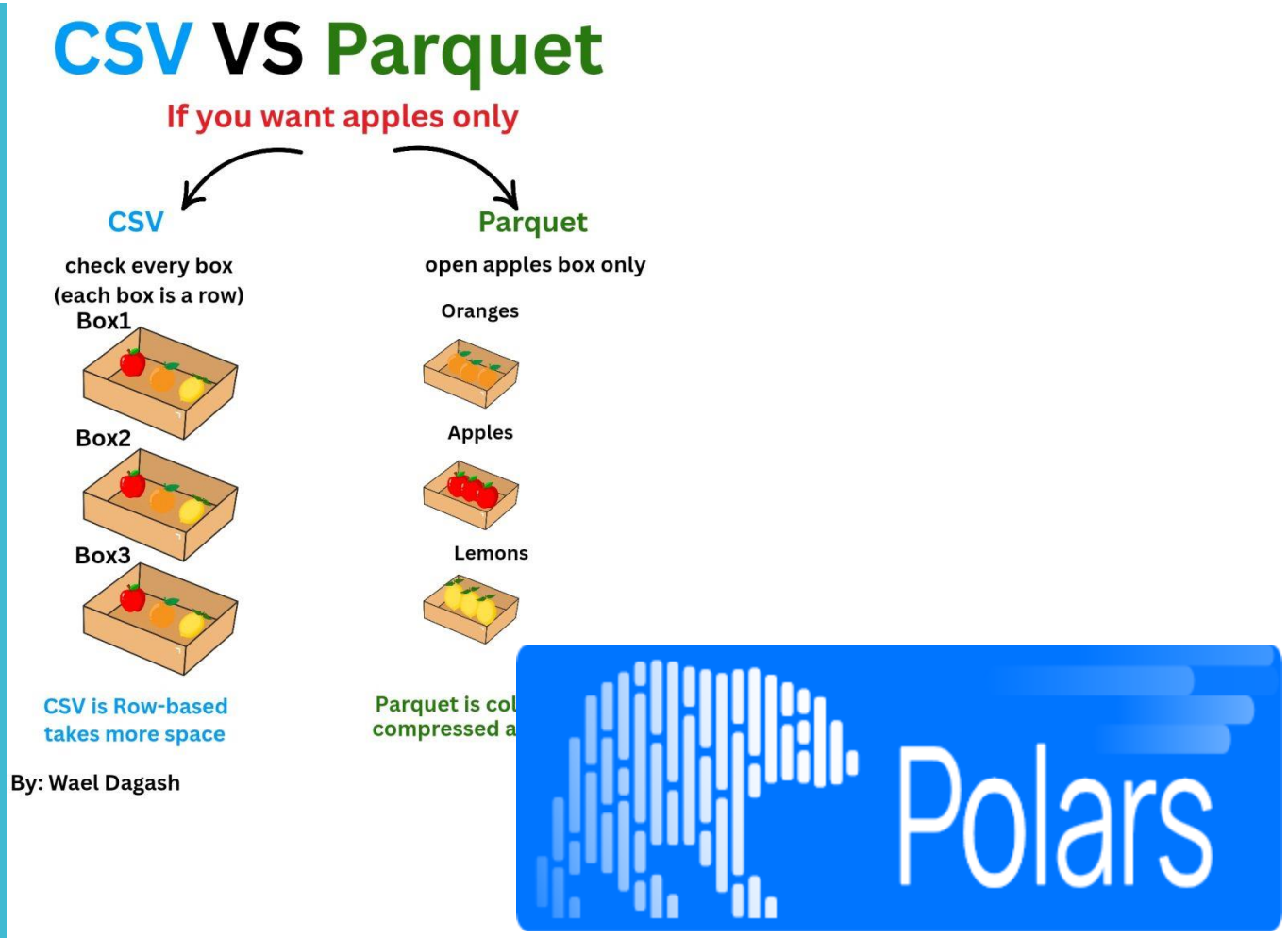
The CSIRO Ocean Data Archive (CODA)

- Wrangling data from various archives can be a nightmare
- We've developed a set of tools to make that easier
  - Based on *parquet* tabular format



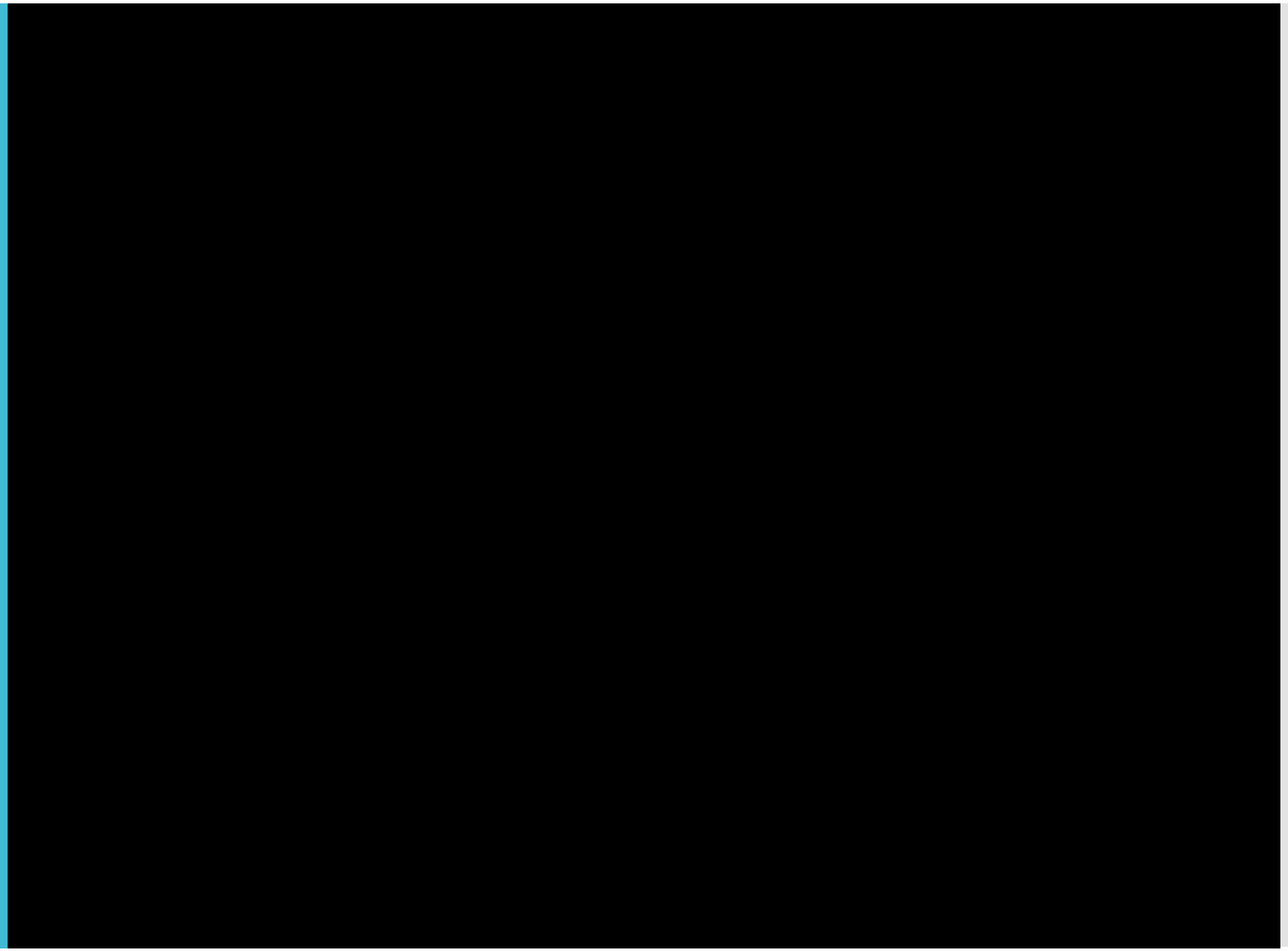
The CSIRO Ocean Data Archive (CODA)

- Wrangling data from various archives can be a nightmare
- We've developed a set of tools to make that easier
  - Based on *parquet* tabular format
  - Uses the open source *Polars* software

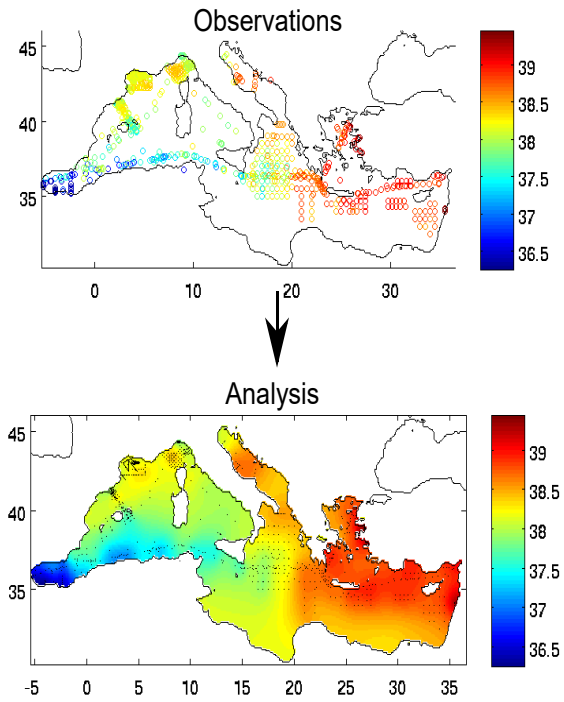


The CSIRO Ocean Data Archive (CODA)

- Wrangling data from various archives can be a nightmare
- We've developed a set of tools to make that easier
  - Based on *parquet* tabular format
  - Uses the open source *Polars* software

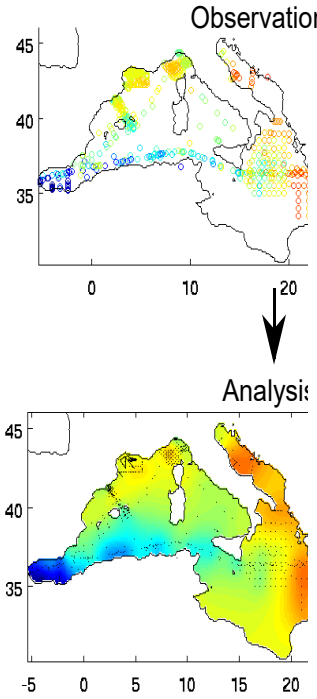


CSIRO Atlas of Regional Seas (CARS)



Source: DIVAnd website: https://gher-uliege.github.io/DIVAnd.jl/

CSIRO Atlas of Regional Seas (CARS)



Source: DIVAnd website: https://

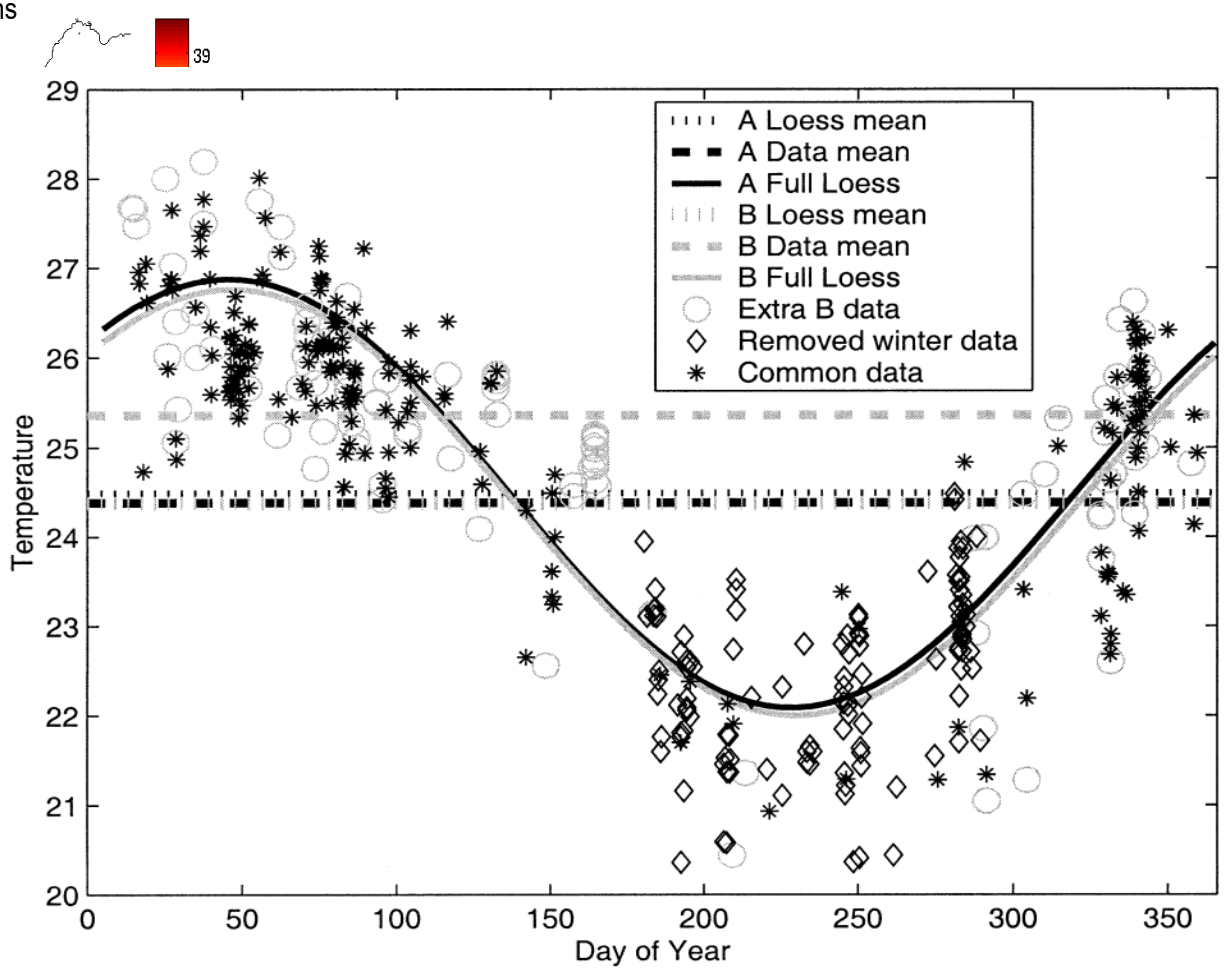
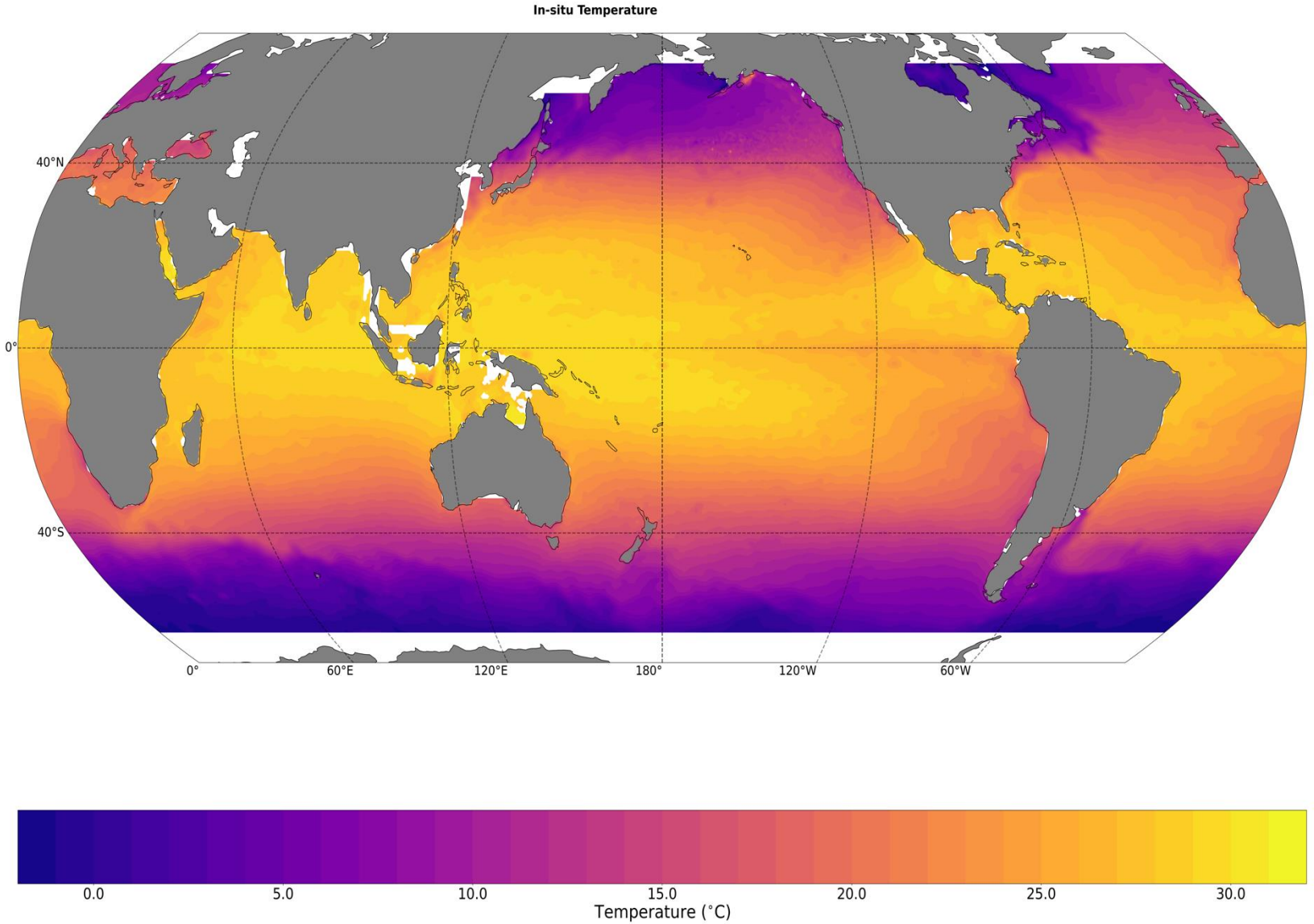


FIG. 9. Surface temperature plotted against day of year for the region (28.2°–21.85°S, 153.5°–160.5°E). The figure legend shows the origin of each symbol and curve plotted in the figure.

Source: Ridgeway et al. (2002)

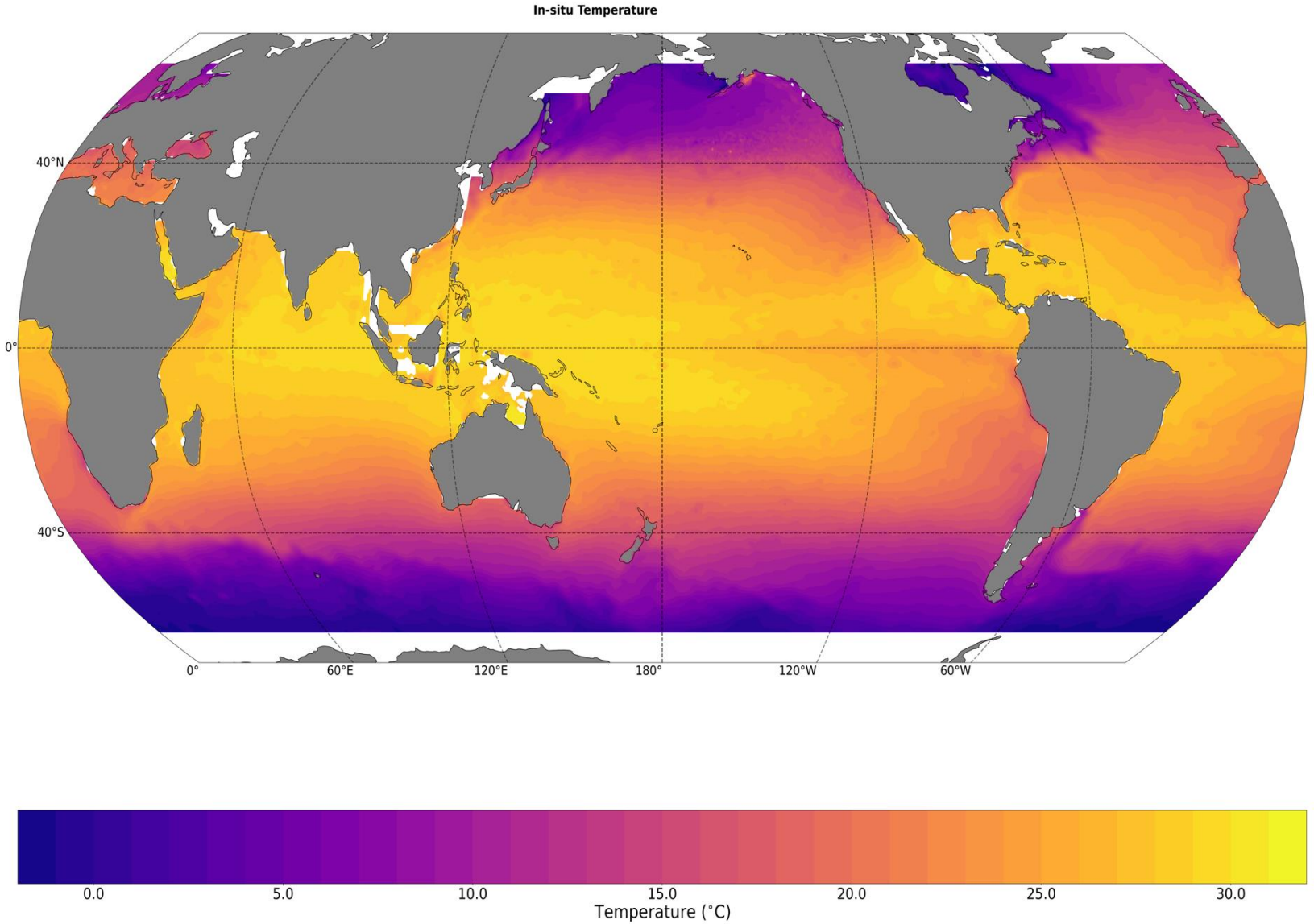
CSIRO Atlas of Regional Seas (CARS)

- First pass of CARSv2



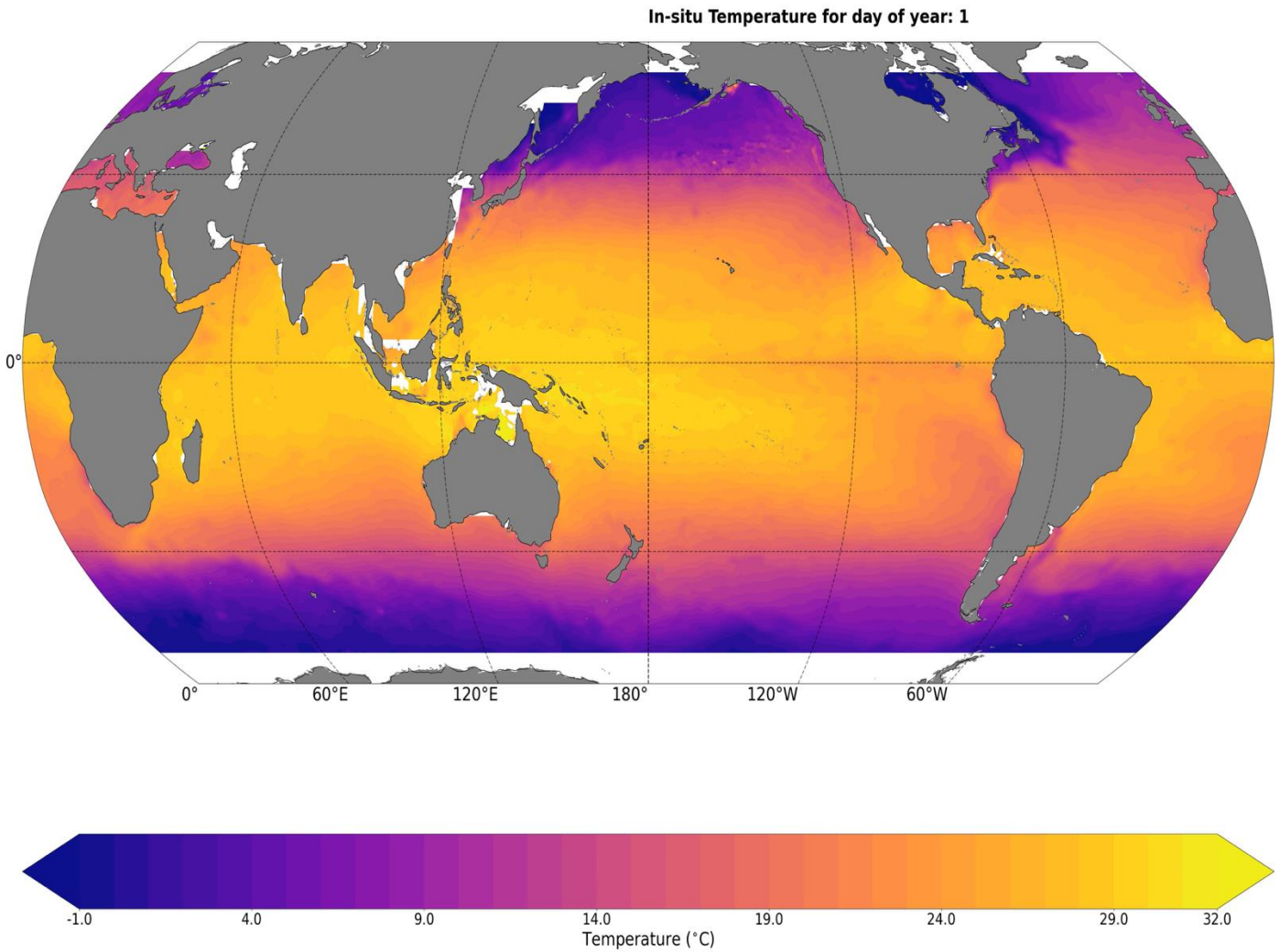
CSIRO Atlas of Regional Seas (CARS)

- First pass of CARSv2
  - ✓ Only using Argo, CTD and Marine Mammals for now



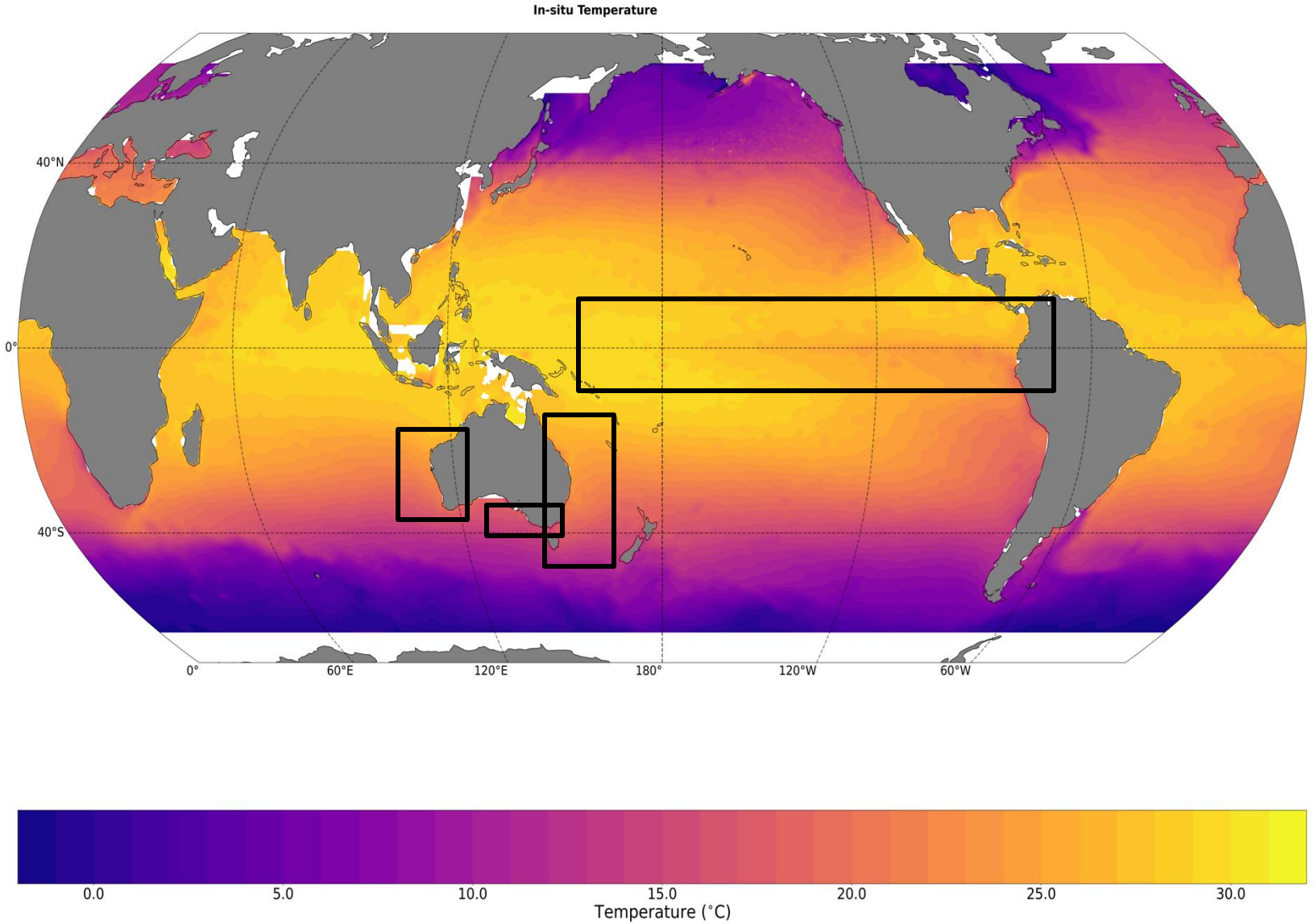
CSIRO Atlas of Regional Seas (CARS)

- First pass of CARSv2
  - ✓ Only using Argo, CTD and Marine Mammals for now
  - ✓ Fancy bathymetry weighting and imposed barriers applied
  - ✓ Updated bathymetry, post map residual screening;

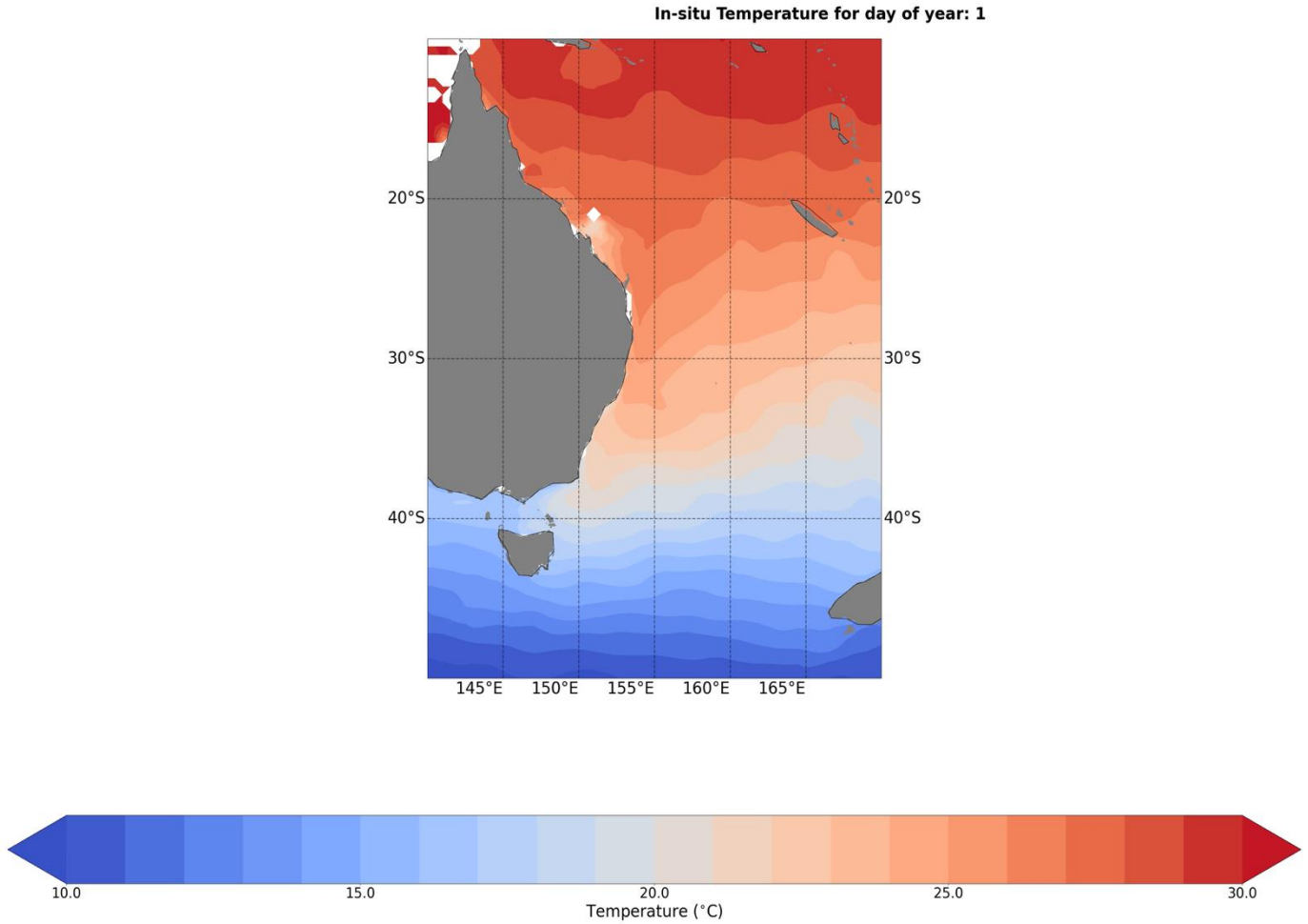


CSIRO Atlas of Regional Seas (CARS)

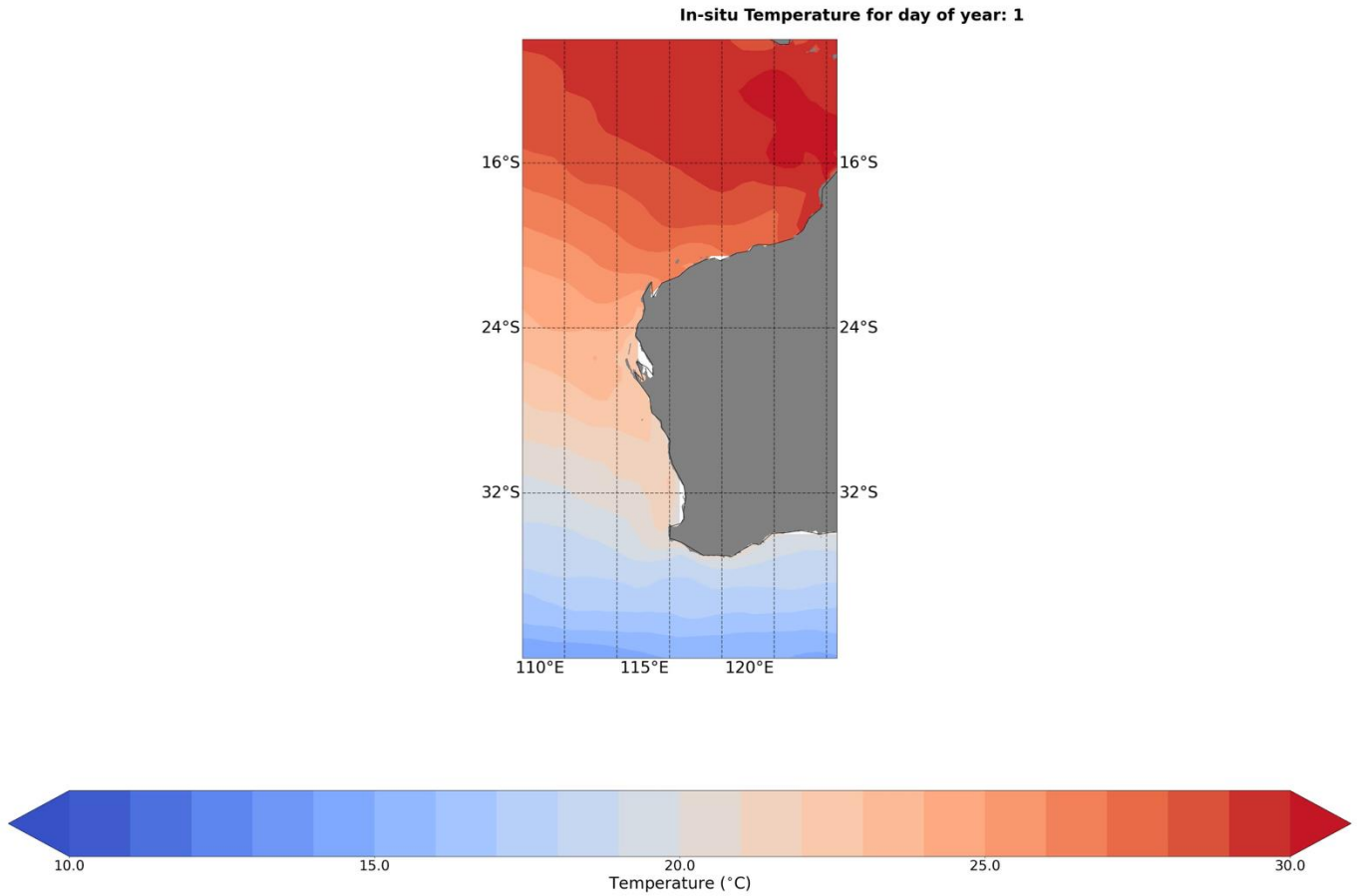
- First pass of CARSv2
  - ✓ Only using Argo, CTD and Marine Mammals for now
  - ✓ Fancy bathymetry weighting and imposed barriers applied



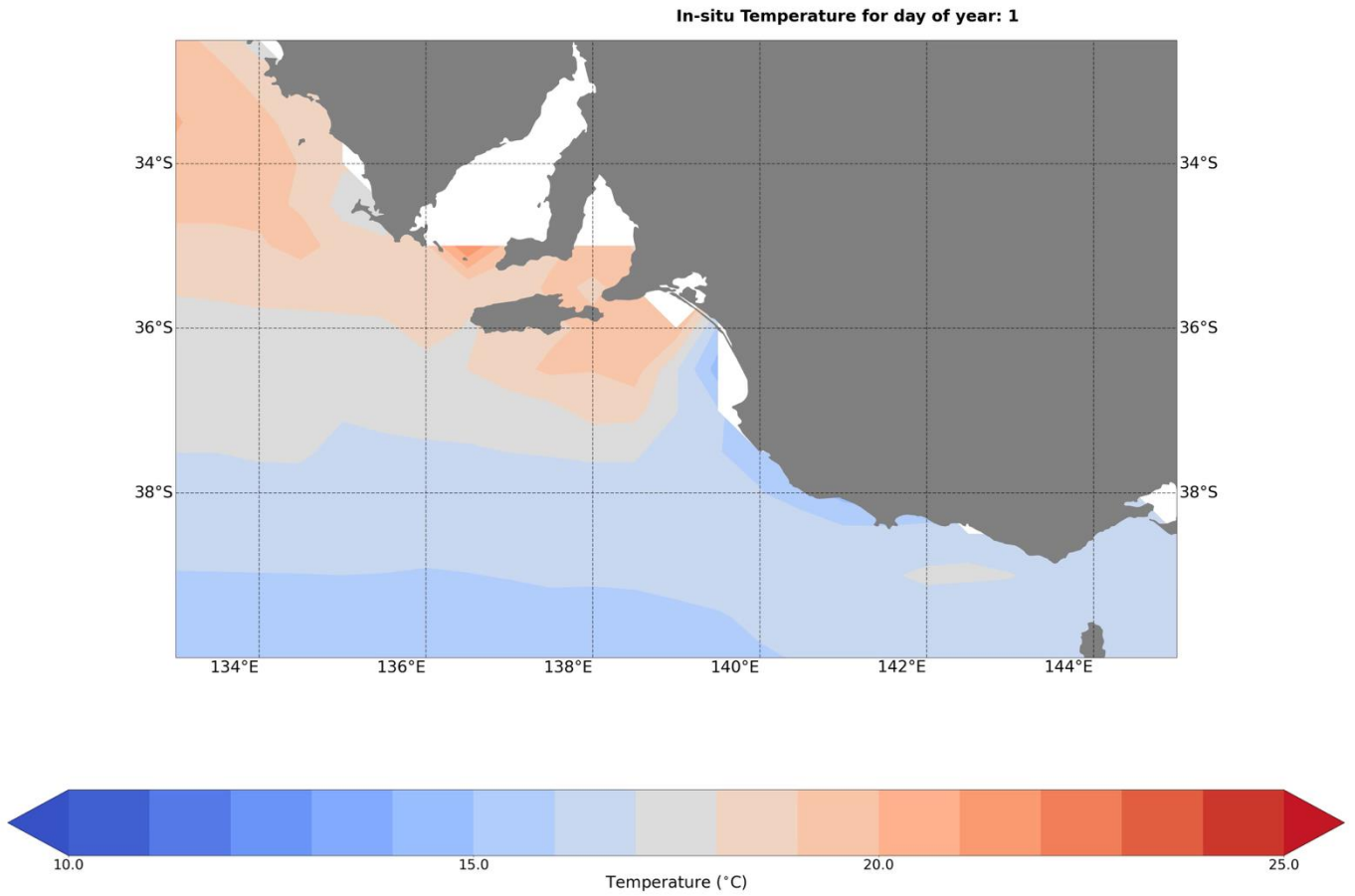
EAC region


In-situ Temperature for day of year: 1
20°S
30°S
40°S
145°E
150°E
155°E
160°E
165°E
10.0
15.0
20.0
25.0
30.0
Temperature (°C)

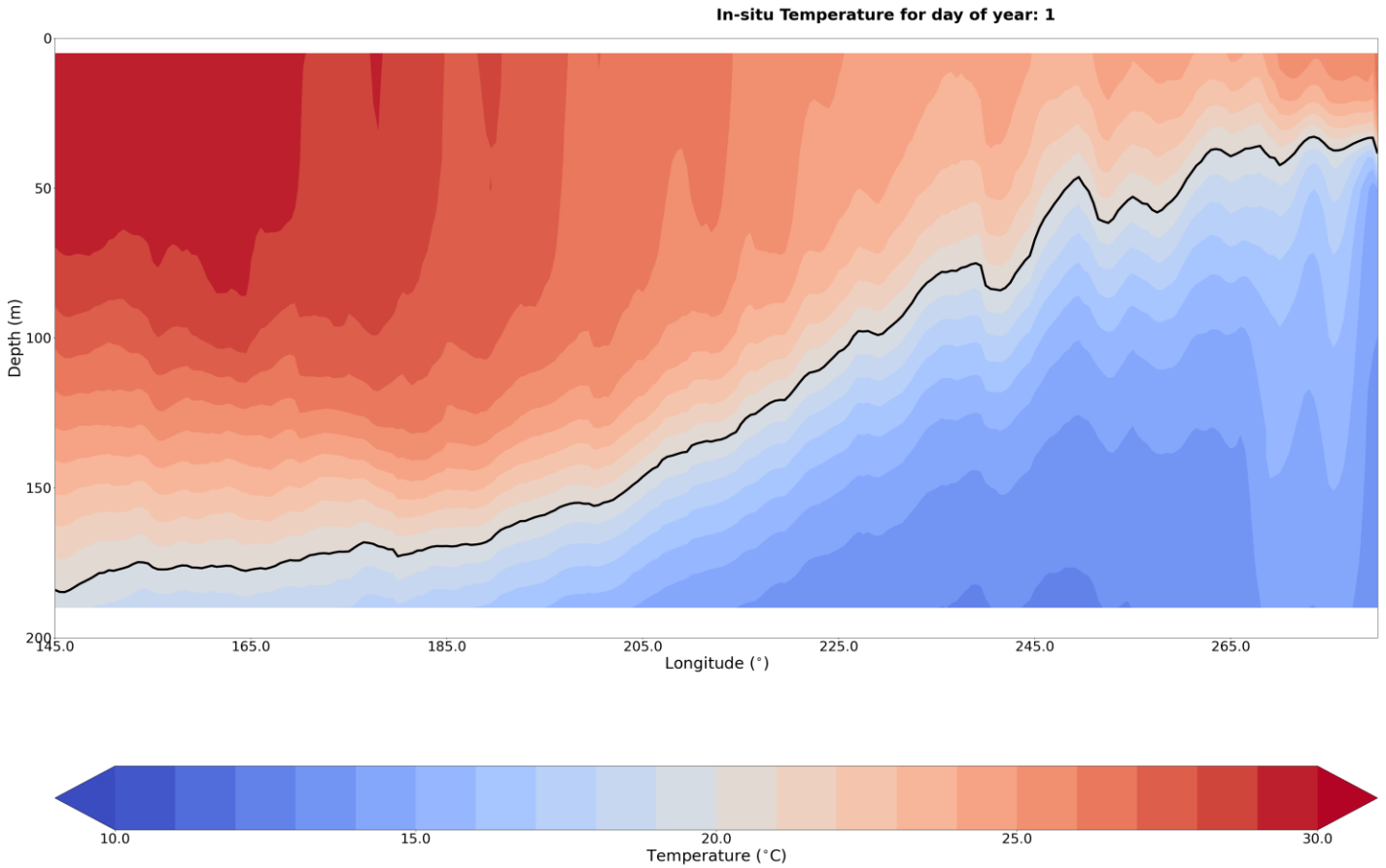
Western Australia/Leeuwin



Bonney upwelling

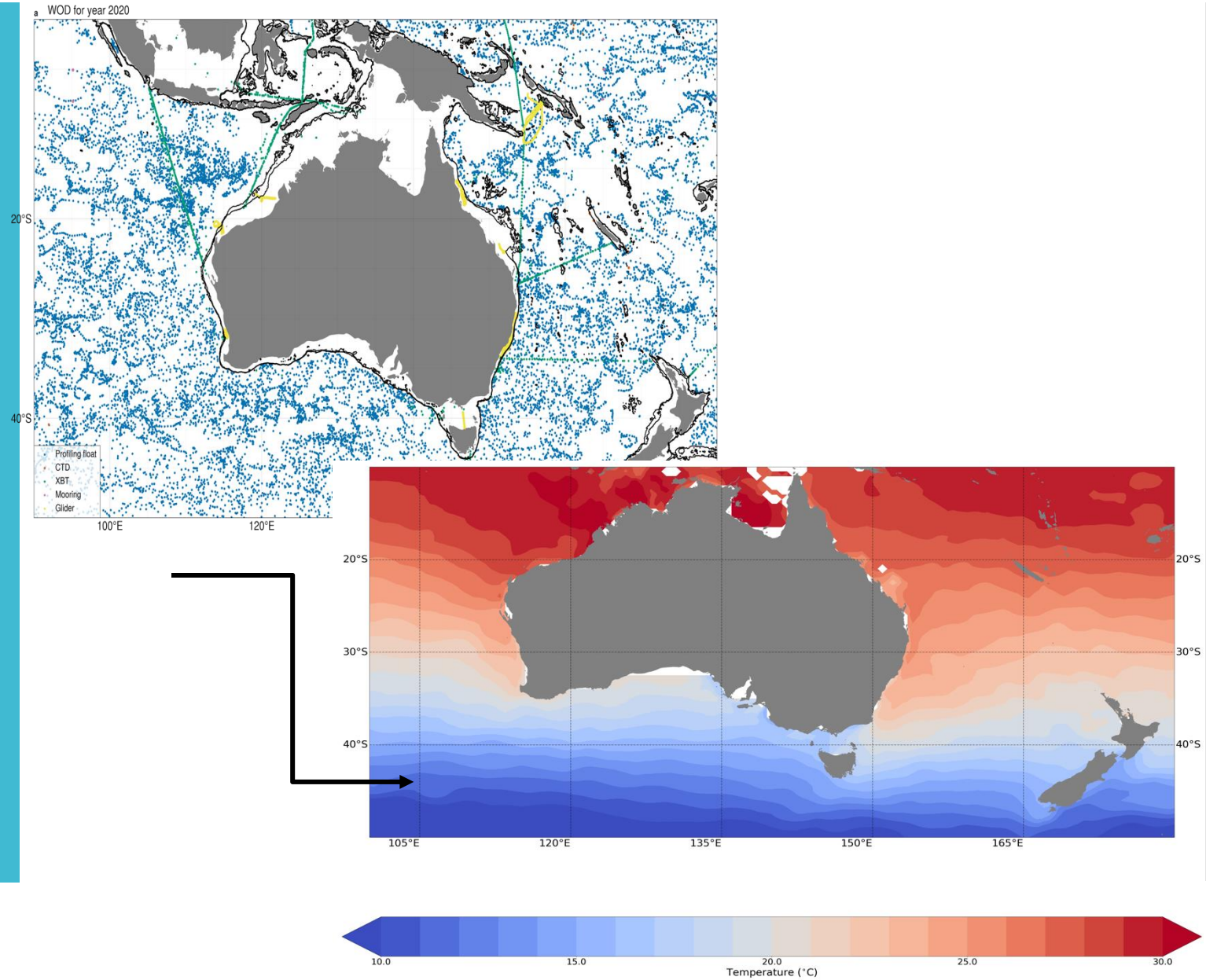


Equatorial thermocline



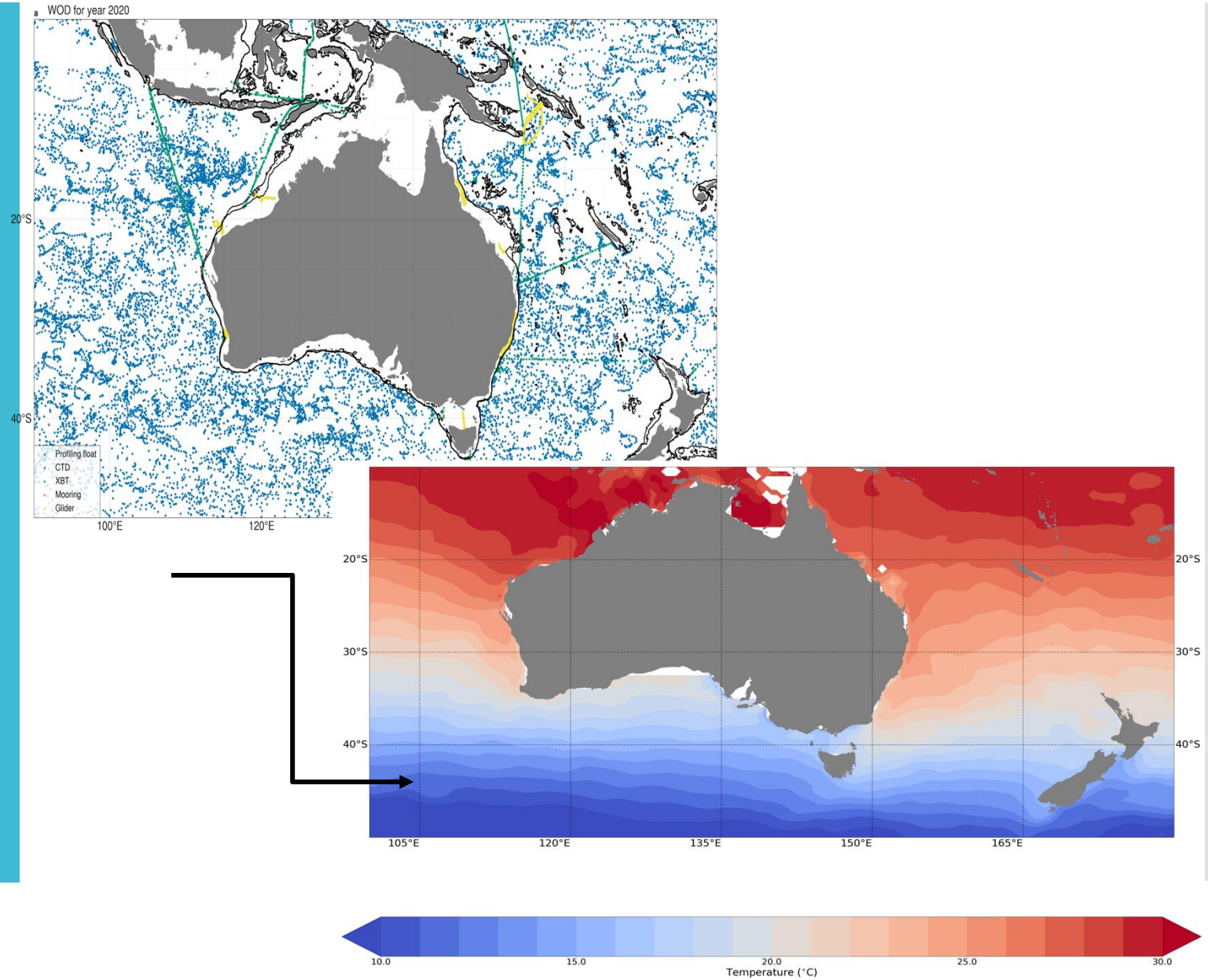
Roadmap forward

- Incorporating "boutique" datasets into CODA one-by-one (tedious process but we are getting there)



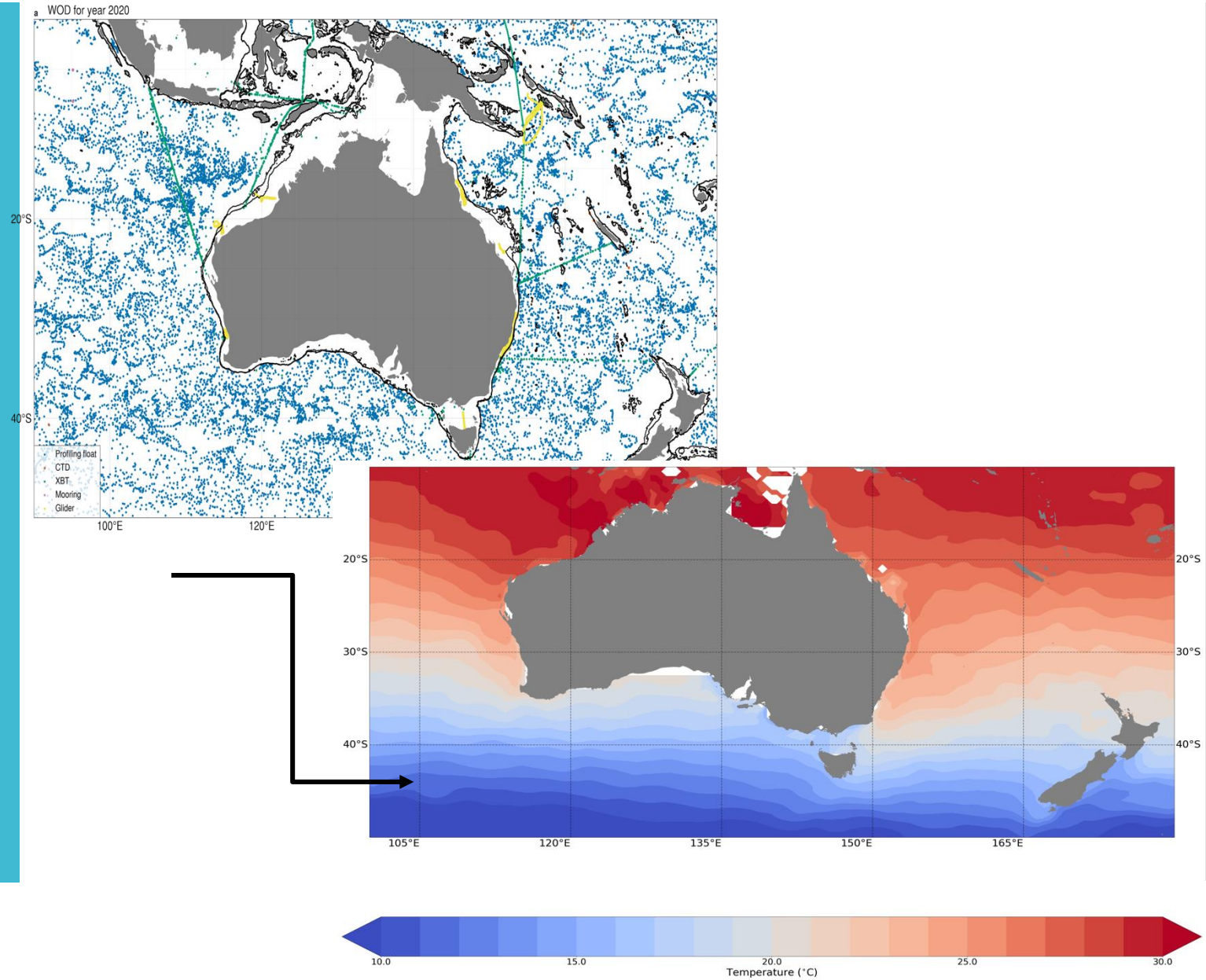
Roadmap forward

- Incorporating "boutique" datasets into CODA one-by-one (tedious process but we are getting there)
- Putting the CODA toolbox into a simple package;



Roadmap forward

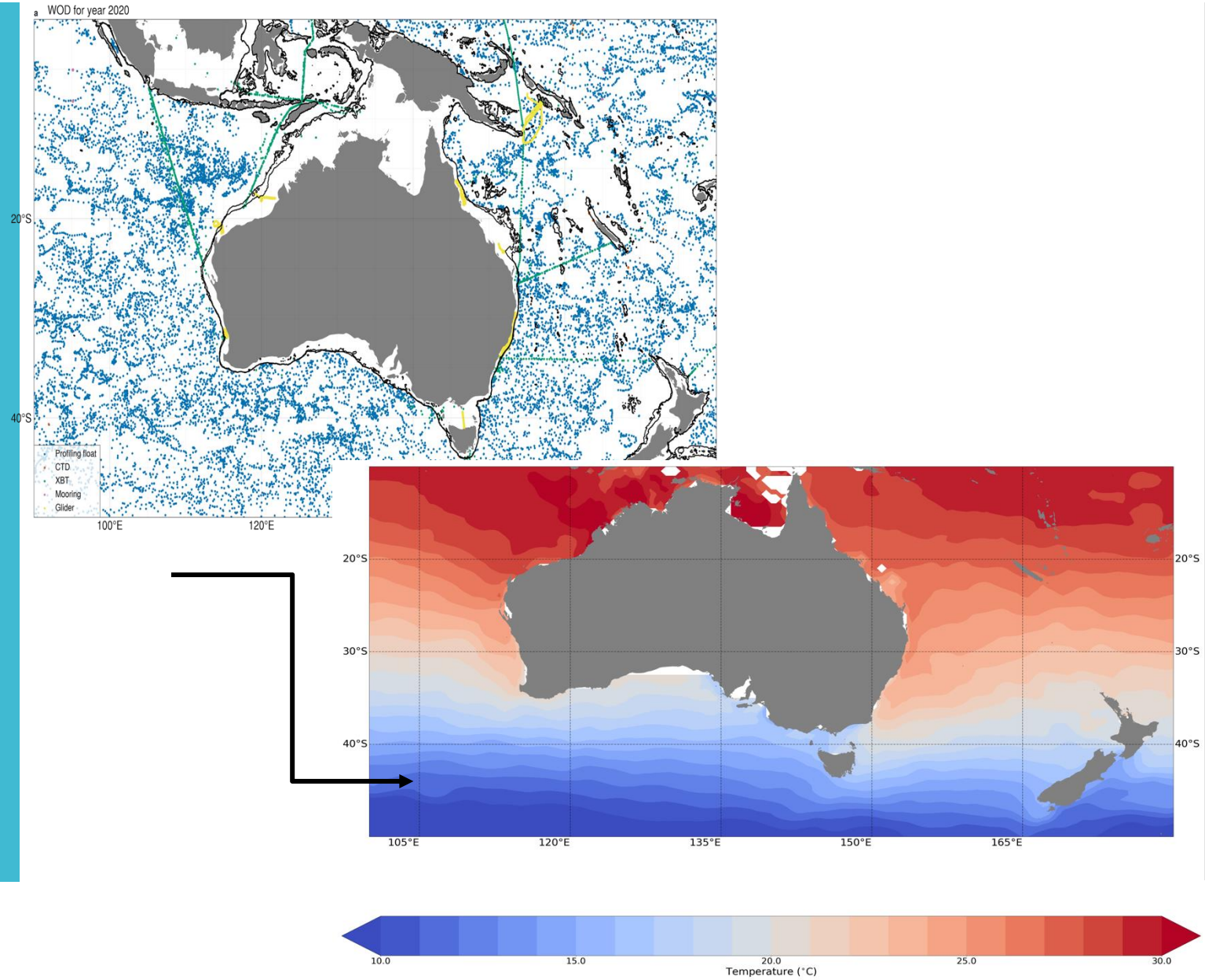
- Incorporating "boutique" datasets into CODA one-by-one (tedious process but we are getting there)
- Putting the CODA toolbox into a simple package;
- Aiming for a CARS release to beta testers around AMOS (February) 2025;



Roadmap forward

- Incorporating "boutique" datasets into CODA one-by-one (tedious process but we are getting there)
- Putting the CODA toolbox into a simple package;
- Aiming for a CARS release to beta testers around AMOS (February) 2025;
- Taking requests! For a limited time only! Do you need:
  - A bespoke set of profiles?
  - A bespoke climatology?

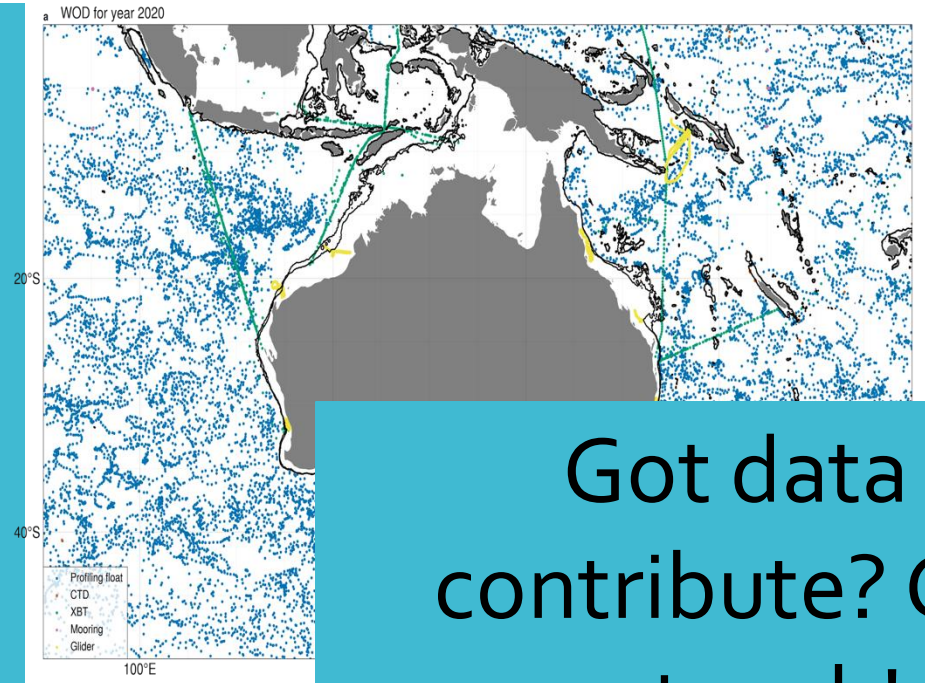
Get in touch!



Roadmap forward

- Incorporating "boutique" datasets into CODA one-by-one (tedious process but we are getting there)
- Putting the CODA toolbox into a simple package;
- Aiming for a CARS release to beta testers around AMOS (February) 2025;
- Taking requests! For a limited time only! Do you need:
  - A bespoke set of profiles?
  - A bespoke climatology?

Get in touch!



Got data to contribute? Get in touch!