**Strengthening the ocean data expert’s community for timely and high-quality data provision**

**Joint Meeting of IQuOD, GTSPP, SOOP, and XBT science groups**

Simona Simoncelli1, Rebecca Cowley2,Lijing Cheng3, Shenfu Dong4, Justine Parks5, Andrea Storto6, Chris Paver7, Chunxue Yang6, Gui Castelao, Claudia Fratianni1

1 *Istituto Nazionale di Geofisica e Vulcanologia, Sezione di Bologna, Italy* [*http://orcid.org/0000-0003-1283-2798*](http://orcid.org/0000-0003-1283-2798)

2 *Environment, Commonwealth Scientific and Industrial Research Organisation (CSIRO), Hobart, TAS, Australia* [*http://orcid.org/0000-0001-8541-3573*](http://orcid.org/0000-0001-8541-3573)

*3 State Key Laboratory of Earth System Numerical Modeling and Application, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China.* [*http://orcid.org/0000-0002-9854-0392*](http://orcid.org/0000-0002-9854-0392)

4 *Atlantic Oceanographic and Meteorological Laboratory/NOAA, Miami, FL, U.S.*

5 *Scripps Institution of Oceanography, University of California, San Diego, U.S.*

6 *Institute of Marine Science, National Research Council of Italy, Rome, Italy*

7 *National Centers for Environmental Information/NOAA, Silver Spring, MD, U.S.*

*Corresponding author*: Simona Simoncelli, simona.simoncelli@ingv.it

**What:** More than 50 international experts of ocean observations, data quality control, and data management came together, under the umbrella of the International Oceanographic Data and Information Exchange (IODE) of the Intergovernmental Oceanographic Commission (IOC) of UNESCO,to explore future collaborations and synergies for efficient ocean in situ data and products provision.

**When:** 11-15 November 2024.

**Where:** Bologna (Italy) and online.

ABSTRACT

Ocean science relies heavily on the collection, management, and analysis of vast amounts of data. The joint meeting of IQuOD, GTSPP, SOOPIP, and XBT Science Team groups brought together experts to discuss critical issues, advancements, and future plans in ocean data management. This report aims to synthesize the key themes and discussions from this meeting, providing a comprehensive overview of the current state and future directions in the field.

**1. Introduction**

The proposal of a joint meeting of the International Quality-controlled Ocean Database (IQuOD), Global Temperature and Salinity Profile Program (GTSPP), Ship Of Opportunity Programme Implementation Panel (SOOPIP), and eXpendable Bathythermograph (XBT) Science Team (XST) was launched early in 2024 by their chairs recognizing the need to discuss the status of the groups’ activities, their overlaps and possible synergies for improving the Global Ocean Observing System (GOOS, https://goosocean.org/) and boosting its data value chain.

The meeting was supported by the International Oceanographic Data and Information Exchange (IODE) program (<https://www.iode.org>) of the Intergovernmental Oceanographic Commission (IOC) of UNESCO, whose mission is enhancing marine research by facilitating the exchange of oceanographic data and information by meeting the needs of users (<https://iode.org/about/>). The *International Quality-controlled Ocean Database (IQuOD)* and the *Global Temperature and Salinity Profile Programme (GTSPP*) are both IODE program activities (<https://iode.org/programme-activities/>), i.e. long-term action receiving minimal UNESCO/IOC funding.

The IQuOD (<https://www.iquod.org/index.html>) group includes experts in data quality and management, ocean and climate modelers and the broader climate-related community. Its operational structure comprises a dynamic workforce of 30-50 international members organized into specialized task teams. It aims to produce and freely distribute the highest quality and complete single ocean profile repository for use in ocean climate research applications (Simoncelli et al., 2024).

The GTSPP (<https://www.ncei.noaa.gov/products/global-temperature-and-salinity-profile-programme>), established by IODE in 1990, is an international cooperative developed by a group of marine and oceanic science organizations to provide researchers and marine operations managers with accurate, up-to-date temperature and salinity profile data. The World Meteorological Organization (WMO) and the IOC jointly manage the network of data capture, archive, and dissemination systems to ensure sustained data quality control, storage, and access. Presently the National Centers for Environmental Information (NCEI, <https://www.ncei.noaa.gov/>) of NOAA maintains the program’s long-term archive by providing storage and quality control services to ensure that best copy versions of GTSPP data are properly preserved and available to the public.

The *Ship of Opportunity Programme (SOOP)* is an observing network of the GOOS, a global platform to deploy and operate oceanographic instrumentation from commercial ships and research vessels. The SOOP panel, as part of the broader Ship Observations Team (SOT), is mainly involved in the data collection, maintenance, enhancement, and data management of upper ocean temperature profile measurements from XBT instruments.

The *XBT science team (XST)* is a group of science users of XBT data. The XST was formed in 2011 with the aim to communicate XBT-related scientific results and to make recommendations on XBT transect prioritization and data management to SOOP and linking with other operational panels.

Common critical issues about the sustainability and efficiency of ocean monitoring and data management activities emerged from the initial chairs’ reports, setting the scene for further discussion. On one side the recognized reduction of workforce and funding resources represents the initial incentive to join the forces and better define the priorities of the ocean observing community to feed the users with data and products fitting their needs. On the other side, the emerging societal challenges require a community effort to share and manage all data acquired for rapid reuse through streamlined data flows, innovative technologies and the adoption of community standards.

The multi-day meeting was designed to review the activities of each participating group, introduce new initiatives, and set the stage for future collaborations. The IODE vision on the global ocean data landscape and the long-term strategy to reduce its present fragmentation represents the overarching context to achieve and it is presented in section 2. The other main meeting topics are organized as follows:

* section 3: observing network optimization and sustainability;
* section 4: data management and sharing;
* section 5: quality control (QC) and bias correction;
* section 6: scientific studies;
* section 7: user requirements.

The challenges and future directions are summarized in section 8.

**2. The Global Ocean Data Landscape**

The advent of the United Nations Decade of Ocean Science for Sustainable Development 2021-2030 (the ‘Ocean Decade’) affirmed the value of data-driven science and the importance of data management to advance ocean knowledge and to tackle emerging societal challenges through an informed decision-making process. The continuous observation of the ocean is at the foundation of this process because it allows systematic monitoring of the ocean state and understanding its functioning and evolution in a rapidly changing environment. Beside the need to sustain and continuously tailor an adequate GOOS, sharing and efficiently managing large amounts of derived data are equally important to transform all existing data into value for society (Simoncelli et al. 2022).

The Ocean Decade Data & Information strategy (<https://unesdoc.unesco.org/ark:/48223/pf0000385542>) has been proposed after a preliminary review and consultation process, to interconnect the existing ocean data systems (nodes) to a sustainable and inclusive digital ecosystem for all ocean stakeholders through the implementation of interoperable layers (UNESCO-IOC, 2023). The Ocean Decade Vision 2030 White Paper – Challenge 8: Create a Digital Representation of the Ocean (Calewaert et al., 2024, <https://unesdoc.unesco.org/ark:/48223/pf0000390123>) details how to create a digital representation of the ocean through recommended practices, data standards, shared architectures and tools to foster ocean data sharing leveraging innovations (artificial intelligence and digital twinning). The key drivers of this blue revolution are the open science and FAIR (Findable, Accessible, Interoperable, Reusable) principles that promote the efficient sharing of data, products and tools, maximizing their re-use and speeding up the value chain for transparent scientific and decision-making processes.

The Ocean Data and Information System **(**ODIS) (<https://odis.org>) is the new program component of IODE which aims to create a sustainable, interoperable, and inclusive digital ecosystem for ocean stakeholders. ODIS can help to shape the future data landscape without fragmentation, preserving diversity and independence of its nodes through the development and implementation of interoperability layers. Thus, ODIS contributes to the implementation of the Ocean Decade strategy, speaking with all the interested actors to regulate and organize their plug-in. An important component of IODE, functional to this process, is the Ocean Teacher Global Academy (<https://classroom.oceanteacher.org/>) which can equalize the digital capacities around the world.

ODIS is thus a program component but also the hub of ODIS federation of regional and thematic nodes that share (meta)data, with IODE acting as a monitoring framework. Ocean InfoHub (OIH) (<https://oceaninfohub.org>) offers a portal allowing discovery of all ODIS content from the network of ODIS partners. An OIH dashboard (<http://dashboard.odis.org>) has been built to monitor ODIS nodes (data exchange) and the metadata describing their assets (FAIRness). ODIS federation is FAIR in practice and in a federated way.

In this context GOOS represents an hypernode component of ODIS for different Essential Ocean Variables (EOVs). The GOOS Observations Coordination Group (OCG, <https://goosocean.org/who-we-are/observations-coordination-group/>) recently released a cross-networks Data Implementation Strategy (2024, GOOS Report No.296) that fits the digital ecosystem ambition of the Ocean Decade and outlines several requirements which are designed to improve the data and metadata flows from the global oceanographic and marine meteorological observing networks to the GOOS hypernode through a federated approach (Fig. 1). At the core of the GOOS federated data system is ERDDAP (Environmental Research Division Data Access Program), an open-source environmental data server software developed by NOAA, which can easily connect distributed data nodes providing data brokering services. The GOOS Data Implementation Strategy also supports the WMO Integrated Global Observing System (WIGOS) and Information System (WIS), based on the Global Telecommunication System (GTS). A phasing out of the GTS and transition to WIS 2.0 is planned in early 2030s but the conceived GOOS ERDDAP federation will maintain its central role interconnecting with both the GTS and the WIS2.0, facilitating a smooth transition.

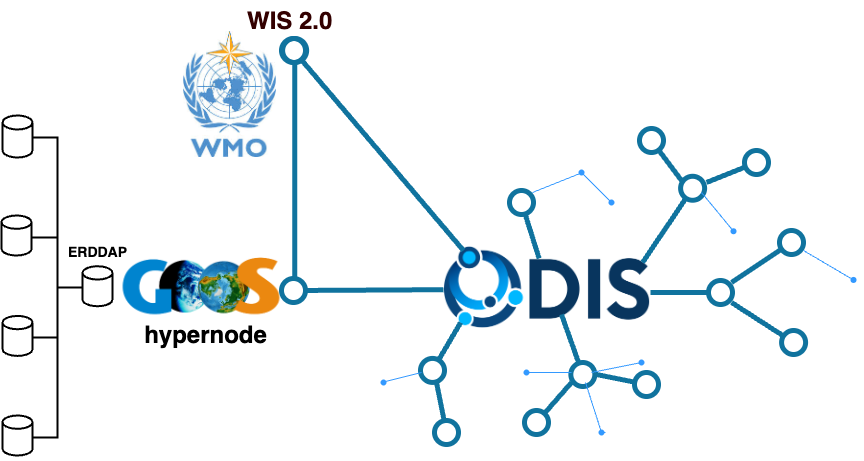


Fig. 1. Positioning of the GOOS OGC federated data system in the ODIS digital ecosystem.

**3. SOOP network optimization and sustainability**

Several observational platforms have been coordinated in SOOP over the years to gather upper ocean data, including XBTs, ThermoSalinoGraphs (TSGs), and Partial Pressure of Carbon Dioxide (pCO2). The SOOP network and implementation panel were formalized in the 1980s and originally dedicated to XBT data gathering and delivery. In recent years, it expanded to include other emerging platforms that lacked organization at the GOOS network level. Each monitoring platform has its own peculiarities (network coverage, sampling frequency, operational equipment, quality assurance practices and data type) and communities of practice, which require specific coordinated actions, capacity development and technology transfer to successfully meet the GOOS targets. Attempting to coordinate the activities of so many diverse groups has strained the leadership of SOOP and has not proven practical. Therefore, the network plans to refocus on XBT measurements. The additional SOOP observational groups are currently coalescing within their related programs and networks under the GOOS umbrella.

The WMO-IOC technical coordination focal point for the GOOS networks is OceanOps (formerly Joint Centre for Oceanography and Marine Meteorology in situ Observations Programmes Support, JCOMMOPS). OceanOPS monitors the status and growth of the GOOS networks through a dedicated dashboard (<https://www.ocean-ops.org/board>) which tracks timely exchange of data and metadata. The reorganization of OceanOPS due to reduced resources required the creation of service level agreements (SLAs) for each group, where services beyond the baseline level are tiered according to additional financial resources allocated to OceanOps by each network. In addition, there is the need to optimize the SOOP metadata flow to OceanOPS in order to constantly monitor the observing system effort and maximize its impact.

Deployment of XBTs began in the 1960s to collect upper ocean temperature profiles. The number of XBT deployments reached maximum in the ‘80s and ‘90s then decreased significantly since the advent of Argo profiling floats (Fig. 2), but the Argo observing system lacks some coverage where XBTs can fill the gap. The Argo ocean observing system is global and can resolve the large-scale circulation but is not adequately sampling the mesoscale features, boundary currents along repeat fixed lines, and coastal regions as the XBT network does. Additionally, the network sustains long time series data and is an extremely cost-effective method for performing repeat transects. Thus, the XBT network is an important complement to the Argo observing system.

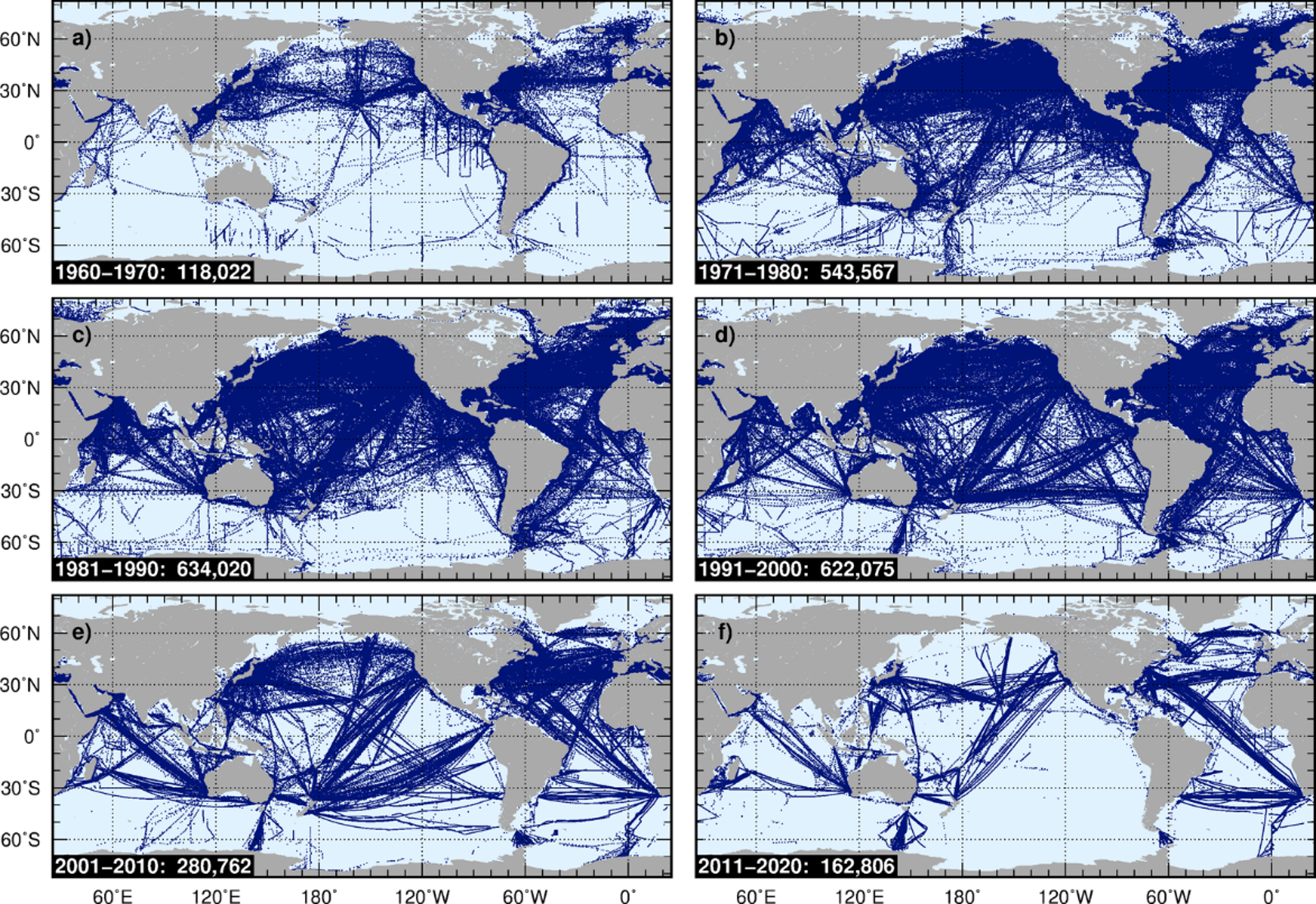


Fig. 2. Evolution of the global XBT network during the 60 years of XBT deployments.

Presently the XBT network consists of 32 transects focusing on monitoring the boundary currents and meridional overturning circulation and its associated mass, heat and freshwater transports (Fig.4). The XBT lines are mostly in high resolution mode (5-30km profile spacing) providing long time series of data for climate studies. Some lines are equipped with additional sensors sampling sea surface atmospheric parameters providing interdisciplinary observations for estimating air-sea heat exchange. The number of deployments in the last decade saw a substantial reduction from 83,450 (yearly 17,300) in the time period 2015-2019 to approximately 50,000 (yearly 10,000) in 2020-2024. The impact of COVID caused loss of some transects due to complication of operations and ship decommissioning (Fig. 3) (Boyer et al 2023). Both funding and resources issues are also reported due to increasing probe costs and lack of technical personnel for operational activities, data managers and scientific experts. The US XBT program is anticipating a significant reduction in support starting mid 2025 and other networks are also experiencing funding challenges, posing the need for international cooperation to find funding mechanisms for future observing activities and minimize the impact on data products and applications for the users, and on the present ocean prediction capabilities.

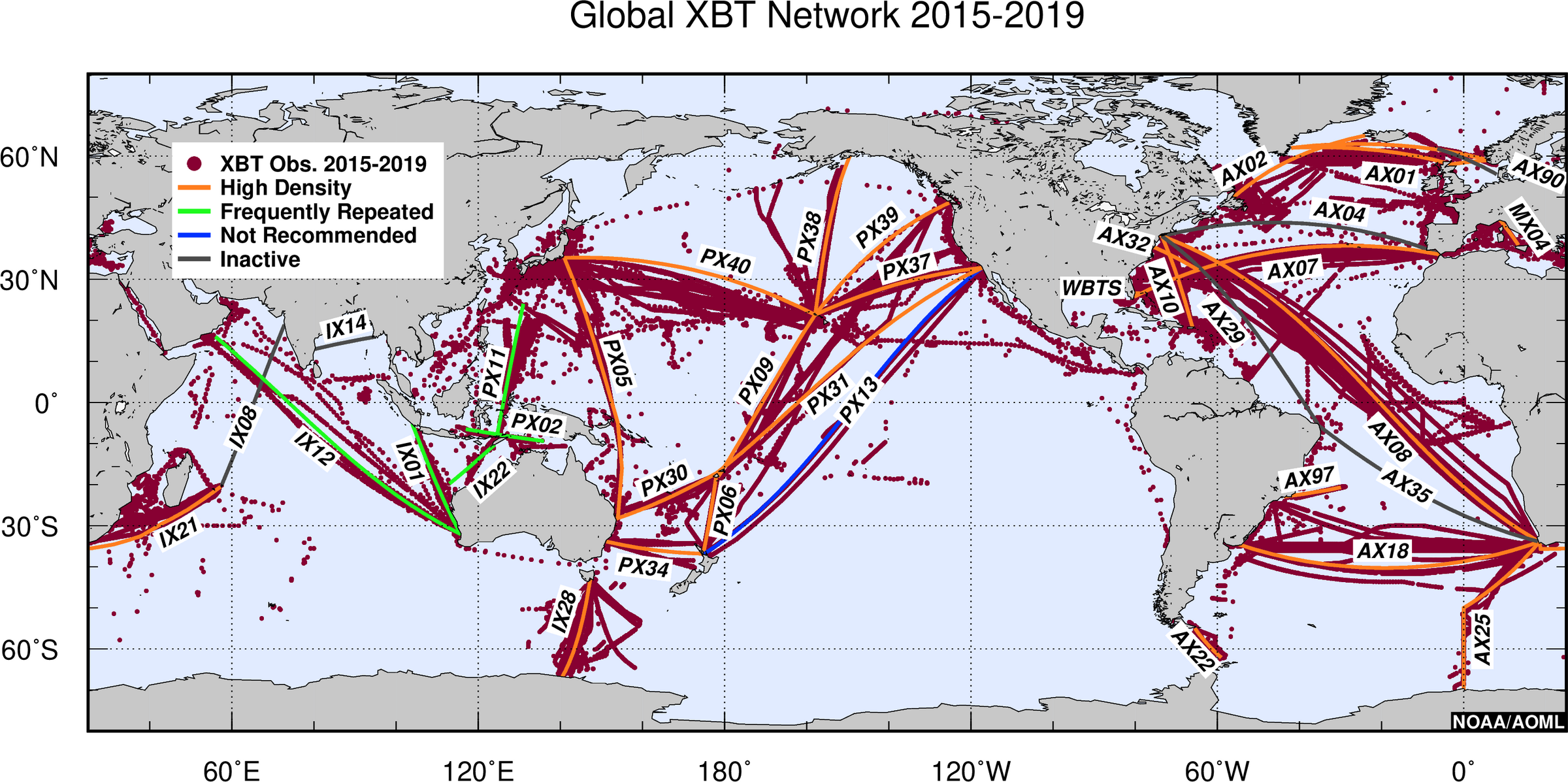
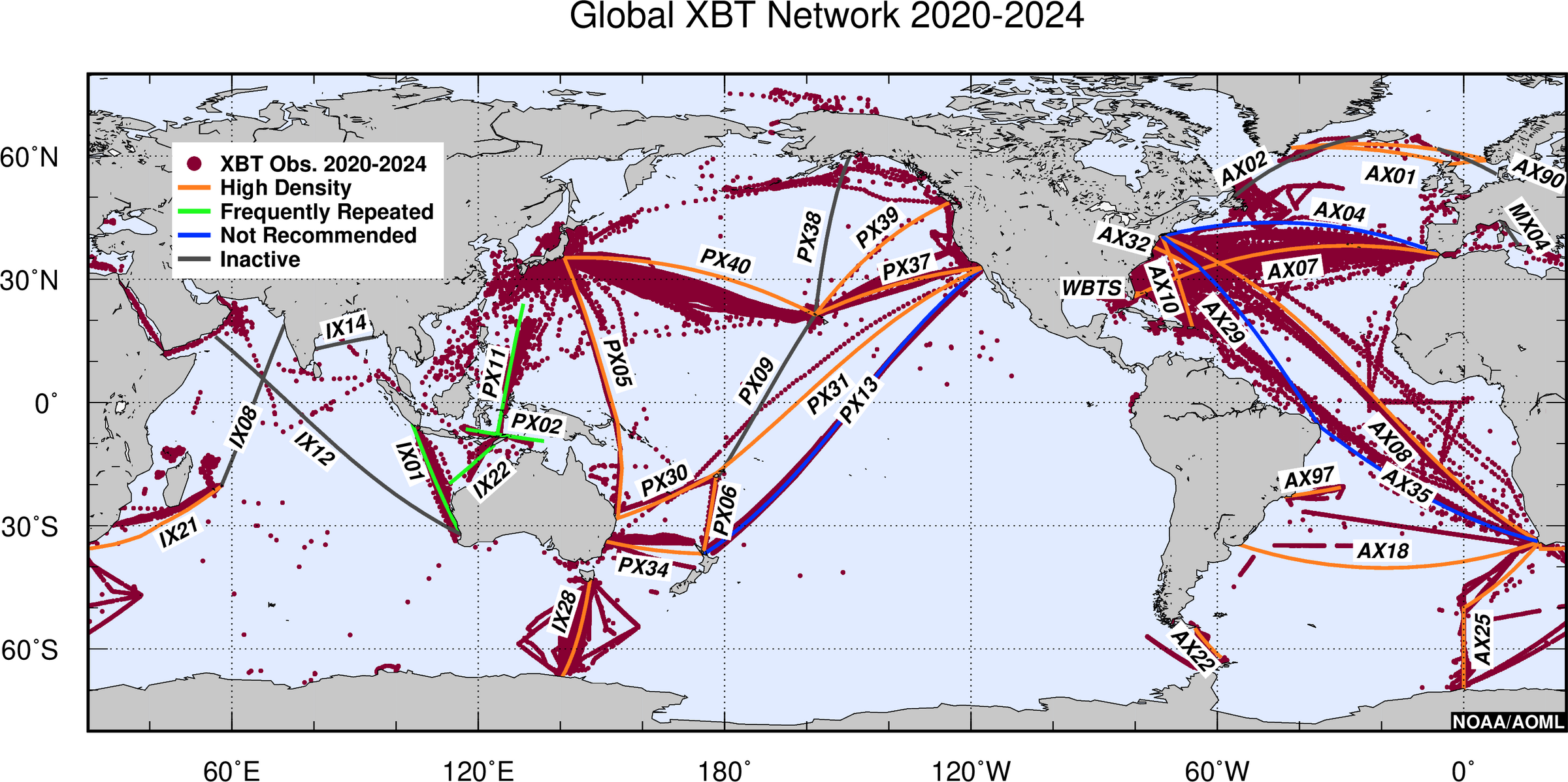
(a)(b)

Fig. 3. The global XBT network: (a) 2015-2019; (b) 2020-2024.

**4. Data Management and Sharing**

The SOOP, GTSPP, IQuOD and XST groups have different positioning along the ocean data value chain (Fig. 4): SOOP is an operational group out in the field collecting upper ocean data that need to be shared rapidly; GTSPP acts as data assembly center that archives and shares the near real-time (NRT) data from the GTS. They also collect delayed-mode QC’d data, sharing them with the World Ocean Database (WOD, also an IODE programme activity); IQuOD value adds to the WOD data by applying bench-marked quality control testing, uncertainty estimates and metadata corrections; XST develops science products from SOOP XBT data. Their mutual collaboration in optimizing the data flow, quality and FAIRness would make the data management framework within and across the observing networks more efficient and transparent. The adoption of standard vocabularies and harmonized metadata information would improve interoperability for further efficient data integration and product generation, and by prompt sharing this information with OceanOPS enables a systematic GOOS effort for ocean monitoring and data flow optimization.

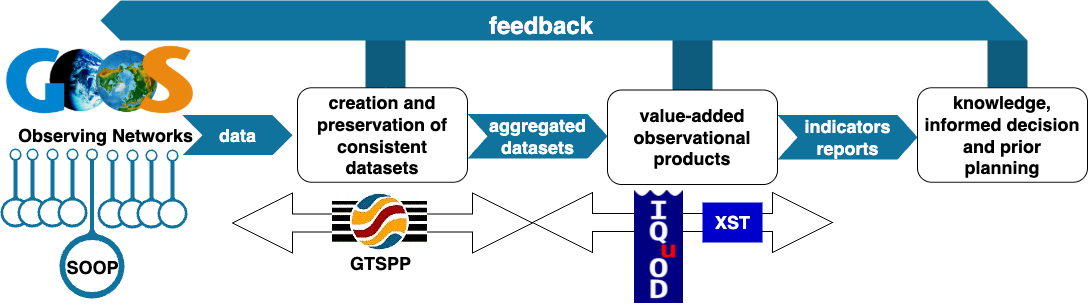


Fig. 4. Positioning of SOOP, GTSPP, IQuOD and XST along the ocean data value chain.

One of the global observing network attributes states that data should be free, open, and available in a timely manner through a dedicated data management infrastructure and FAIR-compliant data services for RT and DM data provision. This process requires that each data provider defines a clear data lifecycle and management plan to share data and derived products, while a common strategy to optimize the network data flow would boost the GOOS data value chain (Fig.5), facilitating the overall data integration process. The definition of retrieval and formatting procedures through common templates including all essential metadata information for NRT data submission and long-term preservation would harmonize the raw data (level 0 in Fig. 5) through the network and assure their completeness for the successive QC analysis and any eventual future re-processing.

The OCG Data Implementation Strategy (2024) outlines several requirements which are designed to improve flows of data and metadata in GOOS through the development of a federated network based on ERDDAP technology, which not only provides data brokering services but can also easily connect distributed data nodes. The adoption of ERDDAP to document, serve and archive data is thus recommended to SOOP data providers as endpoint data service because it can abstract both producers and users from data formats and it allows machine-to-machine interactions. One example of the usage of ERDDAP to share XBT data is in Simoncelli et al. (2024).

In GTSPP the NRT data are gathered weekly from the GTS, then archived, automatically quality controlled and disseminated. The development and adoption of common automatic QC procedures would provide harmonized and consistent NRT data through the GTS/WIS2.0 and the GTSPP. NRT data can then be subject to secondary QC procedures for DM data provision. The generation of shared software code for data QC, processing and reuse would: make the workflow more efficient and transparent; make possible the implementation of a feedback loop on data errors; favor the generation of consistent products (e.g. gridded datasets); increase proper data reuse. This collaborative effort for shared procedures through software repositories (e.g. <https://github.com/>) would favor the realization of web portals or catalogs for better communicating to the users about the SOOP(GOOS) network(s) and its data products.

The impending move to cloud-based data services for enhanced storage (e.g. Argo GDAC daily synchronization on Amazon S3, https://registry.opendata.aws/argo-gdac-marinedata/) and computing capabilities requires taking into consideration new data formats (e.g. Zarr and Parquet), innovative data access services and data lake solutions to be able to consume big data volumes on Virtual Research Environments (Simoncelli et al. 2022). The World Ocean Database cloud (WODc) and Blue Cloud 2026 (<https://www.blue-cloud.org/>) projects are examples of this transformation providing collaborative environments for both data producers and users to manage, quality control and analyze big amounts of data thanks to shared software tools. The webODV (<https://webodv.awi.de/>) is one of these tools already deployed on Blue Cloud VRE to support the generation of integrated EOVs data collections (<https://blue-cloud.org/workbenches-essential-ocean-variables-eovs>). The webODV service also shares community datasets providing online data access, thus it could support our community on the generation and promotion of new common data products but also the training of young scientists.

Key aspects of cloud infrastructures are scalability, reliability and robustness which can also assure system resilience through redundant archives and community data mirroring. The development of open, FAIR and collaborative digital assets is a necessary step to become a node of the expected digital ecosystem.

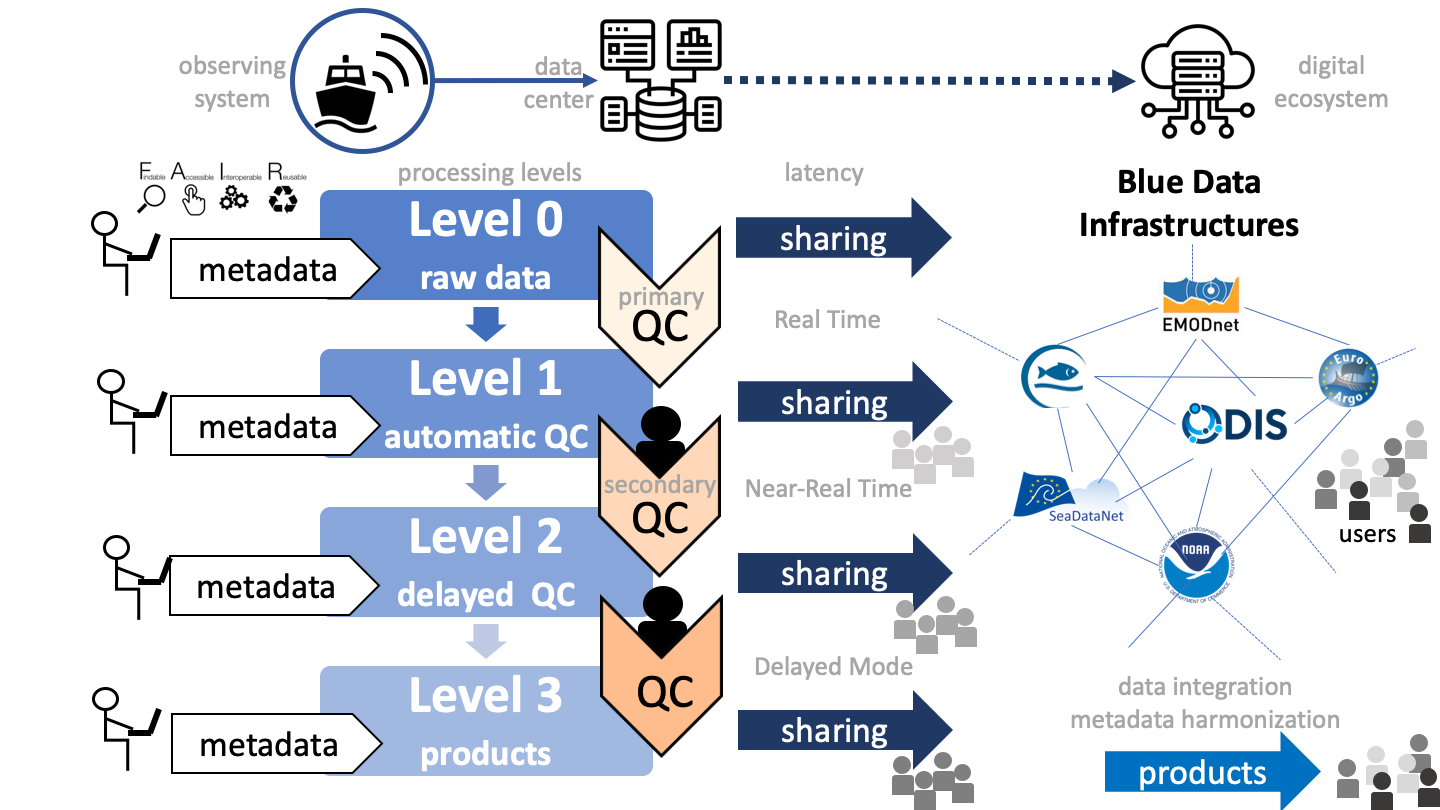


Fig. 5. Ocean data processing levels and data flow from the data provider towards different blue data infrastructures, which are interconnected in a digital ecosystem landscape. Intermediary and end users can access data at different processing or integration levels with different latencies according to their requirements.

**5. Quality Control (QC) and Bias Correction**

QC procedures for temperature and salinity data were a significant topic, including discussions on automated QC (AutoQC) tools and the need for benchmark datasets for salinity QC in the context of IQuOD and Argo. Salinity QC was identified as an area requiring further attention, with a proposal to create a task team for comparing different QC schemes and identifying best practices. The challenge of salinity systematic bias, particularly in Argo data, was also discussed.

The first version of the IQuOD dataset (IQuODv0.1) is built upon the WOD, with the addition of measurement uncertainty for each observation of temperature and in some cases for depth/pressure and salinity (Cowley et al. 2021). Intelligent metadata is included in IQuODv0.1 to provide Expendable Bathythermograph probe type and drop rate (Haddad et al. 2022). IQuODv0.1 is currently available via NCEI. IQuOD also maintains several code repositories on github (https://github.com/IQuOD), including one for AutoQC. Starting from IQuODv0.1, the IQuOD team presented updates on their AutoQC efforts after Good et al. (2023) and the upcoming release of the updated product that includes the IQuOD AutoQC flags (currently, IQuODv0.1 contains WOD flags in the IQuOD flag fields).. The IQuOD community has worked to remove duplicates from the World Ocean Database and correct and add metadata where possible (Song et al. 2023). The uncertainty assignment for each measurement will also be updated to include the representativeness error (e.g., Forget and Wunsch, 2007), which will better support the objective analysis and reanalysis community.

Data useability was an important topic. The Argo team is building new ARCO (Analysis Ready Cloud Optimized) products in parquet format with only selected good quality data (EasyOneArgo) and selected depth levels (EasyOneArgoLight) to fit the needs of ocean and climate modelers. Similar activities are underway from the Chinese Academy of Science Ocean Data Center (CODC dataset) (Zhang et al. 2024). The provision of simplified datasets could be a big step forward in making observational data more accessible to users, no longer requiring them to read and understand the flagging schemes or different data mode variables (e.g. real-time, adjusted real-time, and delayed mode) for different datasets. IQuOD, WOD and the XST recognised the need for such products. A task team was created to work towards new data products for XBT data (both in situ and gridded XBT datasets). These products should be cloud-accessible as well as accessible via more traditional data servers (eg, OpenDAP, ftp etc).

Biases in various ocean instruments (XBT, MBT, Argo, bottle data) have been identified in the past decades and the correction schemes were reviewed, highlighting the progress made in addressing systematic errors in historical data (e.g.,, Gouretski et al. 2020, 2022). The need for caution when using reference data for bias estimation was emphasized.

**6. Scientific Studies and Key Applications**

One of the key applications of high-quality data is to support scientific findings. The meeting discussed the recent scientific achievements and new understandings based on data/observations, but also highlighted potential scientific questions that can be addressed with advanced datasets. The topics discussed include ocean heat content, salinity and freshwater flow, sea level budget, heat/freshwater transport, ocean circulations, mesoscale and sub-mesoscale eddies, and uncertainty estimates.

The improved data quality has led to the recent establishment of global ocean warming acceleration (Cheng et al. 2022) and better closure of sea level budget (Fredriske et al. 2020). An investigation of two different QC schemes can lead to a 8% difference in the upper 2000m OHC estimate, suggesting the importance of QC in climate monitoring (Tan et al. 2023). However, gaps remain, for example, there is still a mismatch of the inter-annual variation between OHC and net radiation imbalance observed at the top of the atmosphere (Hakuba et al. 2024); the budget of sea level acceleration is still not closed since ~1960. Further improvements of data quality have potential to support addressing such scientific questions.

Besides global-scale research, high-quality data also supports the data-driven (and machine learning) reconstructions of ocean dynamics (e.g., Smith et al. 2023), meridional ocean heat and freshwater transports (e.g., Pita et al., 2024), and understanding regional processes (e.g., Yashayaev 2024). An example is high spatial resolution XBT sections. While multidecadal XBT sections along fixed lines are repeated about 4 times per year crossing boundary currents, this schedule is not frequent enough to fully resolve shorter period marine heat waves (MHWs). Various methods have been developed to produce higher temporal resolution temperature sections by combining XBT data with daily remotely-sensed data, which then supports MHWs studies. XBT data were recently used to characterize seabed physical properties (reference?). A novel study by Hornbach et al. (2024) demonstrated that in addition to recording water temperature as they fall through the water column, XBT probes continue to record temperature after they impact the seabed, producing unique temperature responses that depend directly on seabed physical properties ().

**7. User requirements**

To increase fitness for use of GOOS observations, it is important to engage more with the users’ community and gather their requirements. Specifically, a meeting session was dedicated to ocean reanalyses and the requirements that the reanalysis developers pose to the observing community to advance their research and development targets. The main challenges for the next generation of ocean reanalyses (Yang at al., 2025) were summarized, bringing some examples of global and regional systems.

In order to produce longer time series of physical and biochemical reanalyses, there is a need to extend EOVs products prior to the 1950s for data assimilation. Storto and Yang (2024) report that observation-induced uncertainty is one of the major factors impacting the total ocean reanalysis uncertainty, implying that high-quality and cross-network consistent observations with associated uncertainty quantification are in high demand. Developers would like to access long time series of global in situ datasets for assimilation, but also best quality benchmark datasets and products (gridded fields, climatologies) for reanalyses evaluation to be used in model intercomparison exercises within the Ocean Decade MER-EP project (<https://oceandecade.org/actions/marine-environment-reanalyses-evaluation-project/>). Benchmark datasets should be homogenous in space and time, of the highest quality, bias corrected, and organized per instrument type. These products would also be very useful for training data-driven AI models for oceanographic and climate-related studies.

The users’ demand is rapidly evolving with the advent of machine learning technologies, acting as big data consumers for ocean predictions, early warning systems, and “what if scenarios” applications (digital twins), which presume their FAIRness compliance. The adoption of consistent data formats and the inclusion of essential metadata, such as the instrument type, were emphasized to facilitate their use and the assignment to measurements of differential instrumental errors in data assimilation schemes.

Some of the spatial (horizontal and vertical) and temporal gaps could be filled promoting the openness and sharing of existing data and through more efficient assembly efforts but there is a need of data rescue initiatives, such as the one launched by the Copernicus Climate Service and the WMO (https://datarescue.climate.copernicus.eu/), that would unlock historical records and make them available to the scientific community for historical reanalyses generation before their complete loss. However, constantly assessing the fitness-for-purpose of GOOS effort through a systematic review of EOVs, Essential Climate Variables (ECVs), and Essential Biological Variables (EBVs), monitoring adequacy towards users’ requirements, and sustaining it through continuous investments, remains a top priority.

The final discussion focused on how to strengthen collaboration and communication through mutual participation in relevant community workshops for a continuous tuning of needs and priorities, to increase awareness of the available products and to implement a feedback loop for reporting on in situ data issues.

**8. Summary and Future Directions**

Several challenges and future directions emerged from the meeting:

* **Contribution to the global ocean digital ecosystem:** discussions revolved around pushing data and metadata to ODIS nodes, ensuring data discoverability and accessibility.
* **Standardization and Interoperability:** continued efforts are needed to standardize data formats, metadata, and QC procedures to enhance interoperability. Adoption of common vocabularies and best practices is critical.
* **International coordination in developing QC and bias correction best practices**: bringing specific expertise from different countries and groups is essential to form an integrated high-quality global ocean database.
* **Software Development and Sharing:** developing and sharing well-documented, FAIR software tools is essential for data reuse and harmonization. Platforms like GitHub and Docker were encouraged.
* **Cloud Computing:** transitioning to cloud infrastructures for data storage, processing, and analysis was identified as a key direction and it requires the creation of high quality and ready-to-use data products in cloud optimized formats. Addressing distributed management and international cooperation issues is necessary.
* **Community Engagement:** fostering more frequent interaction between observing communities, modelers, and reanalysis groups is vital. Establishing feedback mechanisms and creating user-friendly data products will enhance data utilization.
* **Data Rescue:** the need for organized data rescue efforts, particularly for historical data, was emphasized. Creating an inventory of available data and securing funding for digitization are crucial.
* **Funding:** funding challenges, particularly for long-term data management and maintenance of observing networks like the XBT network, were a significant concern. International collaboration is needed to secure sustained funding.
* **Training and Outreach:** the group recognises the need to share the knowledge and best practices of data QC to wider community and less-developed countries. Thus, the Ocean Teacher Global Academy (OTGA) is an ideal platform for teaching tools for data access and understanding of data QC procedures. A task team was established to investigate course development with OTGA.

*Acknowledgments.*

The meeting was locally hosted by the Istituto Nazionale di Geofisica e Vulcanologia (INGV, Sezione di Bologna, Italy) and most participants travel costs were covered by the International Oceanographic Data and Information Exchange (IODE) of the Intergovernmental Oceanographic Commission (IOC) of UNESCO. SD was supported by the Global Ocean Monitoring and Observing (GOMO) Program and NOAA’s Atlantic Oceanographic and Meteorological Laboratory (AOML). NOAA coauthors acknowledge that the scientific results and conclusions, as well as any views or opinions expressed herein, are those of the coauthors and do not necessarily reflect the views of NOAA or the Department of Commerce. L.C. acknowledges the International Partnership Program of the Chinese Academy of Sciences (Grant No. 060GJHZ2024064MI), the Chinese Academy of Sciences and the National Research Council of Italy Scientific Cooperative Programme. We finally thank the SAVHOTEL for the pleasant stay.

APPENDIX A

**Agenda**

**Joint Meeting of IQuOD/GTSPP/SOOP/XBT**

Bologna (Italy), 11-15 November, 2024

**Day-1** **Setting the Scene**

| **Time** | **Topic/Title** | **Presenter** | **Moderator** |
| --- | --- | --- | --- |
| 14:00-14:10 | Opening | Simona Simoncelli  Organizer/ Sponsor | Simona Simoncelli  (INGV) |
| 14:10-14:40 | ODIS introductory talk | Pier Luigi Buttigieg |
| 14:40-15:10 | IQuOD Review | Guilherme Castelao and  Lijing Cheng |
| 15:10-15:50 | GTSPP Review | Chris Paver |
| 15:50-16:20 | Coffee break | |
| 16:20-17:00 | SOOP Review | Justine Parks |
| 17:00-17:40 | XBT Review | Shenfu Dong |

**Day-2**

**Morning session**: Quality Control, flagging practices/standards.

|  |  |  |  |
| --- | --- | --- | --- |
| **Time** | **Topic/Title** | **Presenter** | **Moderator** |
| 9:00-9:15 | AutoQC implementation and future with WOD cloud | Tim Boyer  (NOAA/NCEI) | Guilherme Castelão  Rebecca Cowley |
| 9:15-9:30 | Salinity QC, report on pilot and prospect for the future | Zhetao Tan,  Yingfan,  Lijing Cheng  (IAP/CAS) |
| 9:30-9:45 | IODE QC flag scheme | Toru Suzuki  (MIRC, Japan) |
| 9:45-10:00 | Argo QC, latest changes and wish list for the future | Thierry Carval  (Ifremer) |
| 10:00-10:15 | OTGA and IQuOD | Uday Bhaskar  (INCOIS) |
| 10:15-10:40 | Discussion | |
| 10:40-11:00 | Coffee break | | |
| 11:00-13:00 | Panel discussion on QC | | |

**Afternoon session**: Pathways to public sharing of (meta)data and code.

| **Time** | **Topic/Title** | **Presenter** | **Moderator** |
| --- | --- | --- | --- |
| 14:00-14:15 | Developing a GOOS Data and Metadata Hypernode through ERDDAP Federation | Kevin O'Brian  (NOAA) | Chris Paver  Simona Simoncelli |
| 14:15-14:30 | Sharing NRT and DM XBT data through ERDDAP | Claudia Fratianni  (INGV) |
| 14:30-15:45 | Data integration and metadata harmonization in cloud-based workflows: the Blue-Cloud workbenches | Simona Simoncelli  (INGV) |
| 14:45-15:00 | GTSPP data product access and looking forward | Chris Paver  (NOAA/NCEI) |
| 15:00-15:15 | Roles and Data Flow of GTSPP, GOSUD, and IQuOD: How to Improve Collaboration and Data Management Practices | Ludovic Drouineau  (Ifremer) |
| 15:15:15:30 | Discussion | |
| 15:30-16:00 | Coffee break | | |
| 16:00-17:30 | Panel discussion on data/code/metadata sharing and data management | | |

**Day-3**

**Morning session:** XBT science and data bias.

|  |  |  |  |
| --- | --- | --- | --- |
| **Time** | **Topic/Title** | **Presenter** | **Moderator** |
| 9:00-9:15 | Marine Heatwaves in the Kuroshio | Janet Sprintall (SCRIPPS) | Shenfu Dong  Justine Parks |
| 9:15-9:30 | XBTs Provide First-Order Characterization of Seabed Physical Properties | Taylor Lee (US Naval Research Lab.) |
| 9:30-9:45 | Update on bias corrections (XBT/MBT/bottle/APB and CH14 update) | Viktor Gouretski (IAP/CAS) |
| 9:45-10:00 | IAP OHC update and a new OHC uncertainty estimate | Lijing Cheng (IAP/CAS) |
| 10:00-10:15 | Reliability of XBT system measurements in oceanographic datasets | Franco Reseghetti (INGV) |  |
| 10:15-10:30 | Discussion | | |
| 10:30-11:00 | Coffee break | | |
| 11:00-13:00 | Panel discussion or group discussion | | |

**Afternoon session**: Perspectives and requirements from the data user side.

|  |  |  |  |
| --- | --- | --- | --- |
| **Time** | **Topic/Title** | **Presenter** | **Moderator** |
| 14:00-14:15 | GLORAN and OceanDecade | Chunxue Yang (CNR) | A. Storto,  C. Yang,  S. Simoncelli |
| 14:15-14:30 | CNR reanalysis | Andrea Storto (CNR) |
| 14:30-14:45 | In-situ observations in CMCC reanalysis systems | A. Aydogdu (CMCC) |
| 14:45-15:00 | ECCO reanalysis | Gael Forget (MIT) |
| 15:00-15:30 | Discussion | | |
| 15:30-16:00 | Coffee break | | |
| 16:00-17:30 | Panel discussion on how to better support user communities such as reanalysis/climate model requirements, metrics and also observing system integration/gaps. | | |

**Day-4**

M**orning session:** Scientific talks, new scientific understandings based on data/observations.

|  |  |  |  |
| --- | --- | --- | --- |
| **Time** | **Topic/Title** | **Presenter** | **Moderator** |
| 9:00-9:15 | OHC: GEWEX group assessment on OHC/EEI | Tim Boyer  (NOAA) | Lijing Cheng  Marlos Goes |
| 9:15-9:30 | Improving data-driven reconstructions of ocean dynamics: from operational to scientific applications | B. Buongiorno Nardelli (CNR) |
| 9:30-9:45 | ICOADS: A Globally Merged Surface Marine and Meteorological Dataset | Chunyin Liu (ICOADS) |
| 9:45-10:00 | Observations methods for salinity and freshwater transport in the Atlantic Ocean | Marlos Goes (NOAA/AOML) |
| 10:00-10:15 | Sea level and its acceleration budget in light of recent observational advances since 1960 | Huayi Zheng  (IAP/CAS) |
| 10:15-10:30 | IAP OHC update and a new OHC uncertainty estimate | Lijing Cheng (IAP/CAS) |
| 10:30-10:45 | ODV/webODV - Low-barrier Access to Environmental Data | Reiner Schlitzer (AWI) |
| 10:45-11:00 | Multiplatform multisensor approach to studying processes and variability in the subpolar North Atlantic. XBT as a key exploratory tool | Igor Yashayev (Bedford Institute of Oceanography) |
| 11:00-11:30 | Coffee break | | |
| 11:30-13:00 | Panel discussion on gaps/research needs in answering some key scientific questions. | | |

**Afternoon session**: Public access platforms (major databases), instruments, Ferrybox, CPR, pCO2, TSG SOOP

|  |  |  |  |
| --- | --- | --- | --- |
| Time | **Topic/Title** | **Presenter** | **Moderator** |
| 14:00-14:15 | The Eurogoos Ferrybox Task Team | Anna Wranne (Voice of the Ocean Foundation) | Justine Parks |
| 14:15-14:30 | Introduction to OceanOps | Martin Kramp (OceanOps) |
| 14:30-14:45 | WIS 2.0 Introduction | David Berry  (WMO) |
| 14:45-15:00 | Ocean data management and dataset developments in IAP/CAS | Lijing Cheng (IAP/CAS) |
| 15:15-15:25 | Introduction to the organization of IOC/GOOS and SOOP’s place within it | Emma Heslop (UNESCO/IOC) |
| 15:25-15:45 | Discussion | | |
| 15:45-16:15 | Coffee break | | |
| 16:15-17:30 | Panel discussion on platforms and instruments: how to better coordinate across the major groups | | |

**Day-5**

**Summary and wrap-up**

|  |  |  |  |
| --- | --- | --- | --- |
| **Time** | **Topic/Title** | **Presenter** | **Moderator** |
| 10:15-10:30 | Share EuroGOOS strategy, side meeting outcomes | Thierry Carval (Ifremer) | Simona Simoncelli |
| 10:30-11:00 | IQuOD Summary Report and Final discussion (including action items, and drafting a provisional budget) | Lijing Cheng  (IAP/CAS) |
| 11:00-11:30 | GTSPP Summary Report and Final discussion (including action items, and drafting a provisional budget) | Chris Paver (NOAA) |
| 11:30-12:00 | Coffee break | |
| 12:00-12:30 | SOOP Summary Report and Final discussion (including action items) | Justine Parks  (SCRIPPS) |
| 13:00-13:00 | XBT Summary Report and Final discussion (including action items) | Shenfu Dong  (NOAA) |
| 13:00-13:30 | Wrap-up | Simona Simoncelli  (INGV) |

APPENDIX B

**List of Participants**

|  |  |  |  |
| --- | --- | --- | --- |
| **Onsite Participants** | | **Remote Participants** | |
| *Rebecca Cowley* | CSIRO | *Kevin O'Brien* | NOAA |
| *Gui Castelao* | NREL | *Rachel Killick* | UK Met Office |
| *Simona Simoncelli* | INGV, Italy | *Catia Domingues* | NOC |
| *Franco Reseghetti* | INGV, Italy | *Tammy Morris* | SAEON |
| *Claudia Fratianni* | INGV, Italy | *Huai Min Zhang* | SOT Chair |
| *Marjahan Finlayson* | UniBo and INGV, Italy | *Kathy Tedesco* | NOAA |
| *Mauro Cirano* | Federal University of Rio de Janeiro (UFRJ) | *Suneel Vasimalla* | Goa, INDIA |
| *Lijing Cheng* | IAP/CAS, China | *Jenny Chiu* | Canada DFO |
| *Zhetao Tan* | IAP/CAS, China | *Simon Good* | Met Office |
| *Uday Bhaskar* | INCOIS | *Lisa Krummel* | BoM, Australia |
| *Gael Forget* | MIT | *Christine Coatanoan* | IFREMER |
| *Alexey V. Mishonov* | NCEI | *Luca Repetti* | Istituto Idrografico della Marina |
| *Christoph Waldmann* | University of Bremen | *Rebecca Zitoun* | IMOS |
| *Alison Macdonald* | WHOI | *Igor Yashayaev* | DFO-MPO |
| *Justine Parks* | SIO | *Cooper Van Vranken* | FVON/ODN |
| *Shenfu Dong* | NOAA/AOML | *Julia Oelker* | ICBM, Universität Oldenburg |
| *Janet Sprintall* | SIO | *Juan Leonardo Moreno* | DIMAR |
| *Chunying Liu* | NOAA/NECI | *Emma Heslop* | IOC/GOOS |
| *Yuri Cotroneo* | University Parthenope, Italy | *Julie Jakoboski* | FVON/Aotearoa Moana Observing System |
| *Antonino Ian Ferola* | University Parthenope, Italy | *Faiza Al-Yamani* | KISR |
| *Chris Paver* | NOAA/NCEI | *Aidan McMahon* | BoM, Australia |
| *Francis Bringas* | NOAA/AOML | *Veronique Lago* | UNSW, Australia |
| *Reiner Schlitzer* | AWI | *Giuseppe Aulicino* | University Parthenope, Italy |
| *Thierry Carval* | IFREMER | *Ivana Melillo* | Grandi Navi Veloci S.p.A. |
| *Andrea Storto* | CNR-ISMAR, Italy | *Yulong LIU* | NMDIS, China |
| *Chunxue Yang* | CNR-ISMAR, Italy | *Leonardo Lima* | CMCC |
| *Marlos Goes* | NOAA/AOML | *Andrea Cipollone* | CMCC |
| *Viktor Gouretski* | IAP/CAS, China | *Alessandro Grandi* | CMCC |
| *Huifeng Yuan* | Computer Network Information Center, CAS, China | *Jenny Pistoia* | CMCC |
| *Toru Suzuki* | MIRC, Japan | *David Berry* | WMO |
| *Xinyi Song* | IAP/CAS, China | *Tim Boyer* | NCEI |
| *Yujing Zhu* | IAP/CAS, China |  |  |
| *Yuying Pan* | IAP/CAS, China |  |  |
| *Joaquin Trinanes* | NOAA/AOML |  |  |
| *Anna Willstrand Wranne* | Voice of the Ocean Foundation |  |  |
| *Huayi Zheng* | IAP/CAS |  |  |
| *Carles Castro Muniain* | FVON/ODN |  |  |
| *Antonio Novellino* | ETT, Italy |  |  |
| *Taylor R. Lee* | US Naval Research Lab. |  |  |
| *Bruno Bongiorno Nardelli* | CNR, Italy |  |  |
| *Martin Kramp* | OceanOPS |  |  |
| *Ludovic Drouineu* | IFREMER |  |  |
| *Ali Aydogdu* | CMCC, Italy |  |  |
| *Patrick Gorringe* | FVON/SMHI |  |  |
| *Pier Luigi Buttigieg* | MPI for Marine Microbiology |  |  |
| *Jerome Gourrion* | PokaPok |  |  |
| *Pietro Miraglio* | CMCC, Italy |  |  |

REFERENCES

Boyer, T., and Coauthors, 2023: Effects of the Pandemic on Observing the Global Ocean. Bull. Amer. Meteor. Soc., 104, E389–E410, <https://doi.org/10.1175/BAMS-D-21-0210.1>.

Cheng, L. et al. Past and future ocean warming. Nature Reviews Earth & Environment. 3, 776-794 (2022).

Cowley R, Killick RE, Boyer T, Gouretski V, Reseghetti F, Kizu S, Palmer MD, Cheng L, Storto A, Le Menn M, Simoncelli S, Macdonald AM and Domingues CM (2021) International Quality-Controlled Ocean Database (IQuOD) v0.1: The Temperature Uncertainty Specification. Front. Mar. Sci. 8:689695. doi: 10.3389/fmars.2021.689695

Forget, G., and C. Wunsch, 2007: Estimated Global Hydrographic Variability. J. Phys. Oceanogr., 37, 1997–2008, <https://doi.org/10.1175/JPO3072.1>.

Frederikse, T. et al., The causes of sea-level rise since 1900. Nature 584, 393-397 (2020).

Good S, Mills B, Boyer T, Bringas F, Castelão G, Cowley R, Goni G, Gouretski V and Domingues CM (2023) Benchmarking of automatic quality control checks for ocean temperature profiles and recommendations for optimal sets. Front. Mar. Sci. 9:1075510. doi: 10.3389/fmars.2022.1075510.

Gouretski, V. & Cheng, L. Correction for Systematic Errors in the Global Dataset of Temperature Profiles from Mechanical Bathythermographs. *J. Atmos. Ocean. Technol.* **37**, 841-855 (2020).

Gouretski, V., Cheng, L. & Boyer, T. On the Consistency of the Bottle and CTD Profile Data. *Journal of Atmospheric and Oceanic Technology.* **39**, 1869-1887 (2022).

Haddad, S., R. E. Killick, M. D. Palmer, M. J. Webb, R. Prudden, F. Capponi, and S. V. Adams, 2022: Improved Infilling of Missing Metadata from Expendable Bathythermographs (XBTs) Using Multiple Machine Learning Methods. J. Atmos. Oceanic Technol., 39, 1367–1385, <https://doi.org/10.1175/JTECH-D-21-0117.1>.

Hakuba, M. *et al.* Trends and Variability in Earth’s Energy Imbalance and Ocean Heat Uptake since 2005. *Surveys in Geophysics.* **45**, 1721–1756 (2024).

Hornbach, M. J., Wood, W. T., Lee, T. R.,Phrampus, B. J., Abelev, A., Herdic, P. C.,et al. (2024). XBTs provide first‐order characterization of seabed physical properties. Earth and Space Science, 11,e2023EA003441. https://doi.org/10.1029/2023EA003441

Observations Coordination Group (OCG) Data Implementation Strategy (2024), GOOS Report No.296. <https://goosocean.org/document/33970>.

Pita I, Goes M, Volkov DL, Dong S and Schmid C (2024) South Atlantic meridional overturning circulation and its respective heat and freshwater transports from sustained observations near 34.5°S. Front. Mar. Sci. 11:1474133. doi: 10.3389/fmars.2024.1474133

Simoncelli, S., Manzella, G. M. R., Storto, A., Pisano, A., Lipizer, M., Barth, A., Myroshnychenko, V., Boyer, T., Troupin, C., Coatanoan, C., Pititto, A., Schlitzer, R., Schaap, D. M. A., & Diggs, S. (2022). A collaborative framework among data producers, managers, and users. In Ocean Science Data (pp. 197–280). Elsevier. <https://doi.org/10.1016/b978-0-12-823427-3.00001-3>

Simoncelli, S., Cowley, R., Tan, Z., Killick, R., Castelão, G., Cheng, L., Good, S., Boyer, T., Mills, B., Bhaskar, U., & Locarnini, R. (2024). The International Quality-controlled Ocean Database (IQuOD). Miscellanea INGV, 80, 139–140. <https://doi.org/10.13127/MISC/80/50>.

Simoncelli, S., Reseghetti, F., Fratianni, C., Cheng, L., and Raiteri, G.: Reprocessing of eXpendable BathyThermograph (XBT) profiles from the Ligurian and Tyrrhenian seas over the time period 1999–2019 with a full metadata upgrade, Earth Syst. Sci. Data, 16, 5531–5561,<https://doi.org/10.5194/essd-16-5531-2024>, 2024.

Smith PAH, Sørensen KA, Buongiorno Nardelli B, Chauhan A, Christensen A, St. John M, Rodrigues F and Mariani P (2023) Reconstruction of subsurface ocean state variables using Convolutional Neural Networks with combined satellite and *in situ* data. *Front. Mar. Sci.* 10:1218514. doi: 10.3389/fmars.2023.1218514

Song X, Tan Z, Locarnini R, Simoncelli S, Cowley R, Kizu S, Boyer T, Reseghetti F, Castelao G, Gouretski V and Cheng L (2024) DC\_OCEAN: an open-source algorithm for identification of duplicates in ocean databases. Front. Mar. Sci. 11:1403175. doi: 10.3389/fmars.2024.1403175.

Tan, Z., Cheng, L., Gouretski, V., Zhang, B., Wang, Y., Li, F., Liu, Z., & Zhu, J.: A new automatic quality control system for ocean profile observations and impact on ocean warming estimate. Deep Sea Research Part I: Oceanographic Research Papers, 194, 103961, <https://doi.org/10.1016/j.dsr.2022.103961>, 2023.

UNESCO-IOC (2023). Ocean Decade Data & Information Strategy. Paris, UNESCO. (The Ocean Decade Series, 45).

Yang, C., and Coauthors, 2025: Gathering Users and Developers to Shape Together the Next-Generation Ocean Reanalyses. Bull. Amer. Meteor. Soc., 106, E419–E429, <https://doi.org/10.1175/BAMS-D-24-0034.1>.

Yashayaev, I. Intensification and shutdown of deep convection in the Labrador Sea were caused by changes in atmospheric and freshwater dynamics. Commun Earth Environ 5, 156 (2024). <https://doi.org/10.1038/s43247-024-01296-9>.

Zhang, B., Cheng, L., Tan, Z. et al. CODC-v1: a quality-controlled and bias-corrected ocean temperature profile database from 1940–2023. Sci Data 11, 666 (2024). https://doi.org/10.1038/s41597-024-03494-8