# CEOS Ocean Variable Enabling Research & Applications for GEO (COVERAGE): A Platform to Simplify and Expand the Accessibility and Usage of Inter-agency Satellite and in-situ Oceanographic Data


Vardis M. Tsontos
*Jet Propulsion Laboratory, California Institute of Technology*
Pasadena, CA, USA
vtsontos@jpl.nasa.gov

Jorge Vazquez
*Jet Propulsion Laboratory, California Institute of Technology*
Pasadena, CA, USA
jvazquez@jpl.nasa.gov

Thomas Huang
*Jet Propulsion Laboratory, California Institute of Technology*
Pasadena, CA, USA
thomas.huang@jpl.nasa.gov

Toshio M. Chin
*Jet Propulsion Laboratory, California Institute of Technology*
Pasadena, CA, USA
toshio.m.chin@jpl.nasa.gov

Joe T. Roberts
*Jet Propulsion Laboratory, California Institute of Technology*
Pasadena, CA, USA
joe.t.roberts@jpl.nasa.gov

Joseph Jacob
*Jet Propulsion Laboratory, California Institute of Technology*
Pasadena, CA, USA
joseph.jacob@jpl.nasa.gov

Flynn Platt
*Jet Propulsion Laboratory, California Institute of Technology*
Pasadena, CA, USA
flynn.platt@jpl.nasa.gov

Vaishnavi Yathirajam
*Jet Propulsion Laboratory, California Institute of Technology*
Pasadena, CA, USA
vaishnavi.yathirajam@jpl.nasa.gov


OCEANS 2022, Hampton Roads | 978-1-6654-6809-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/OCEANS47191.2022.9977301


***Abstract*— Given the growing variety and volume of oceanographic observations, there is a need for data systems and services supplementing existing agency repository infrastructures to enable simpler access and synergistic use of multidisciplinary ocean data. The availability of a set of core software capabilities and data that can be reused, augmented, and scaled efficiently to support a wide range of regional use cases is also a related need. The CEOS (Committee on Earth Observation Satellites) Ocean Variables Enabling Research and Applications for GEO (COVERAGE) initiative seeks to address these imperatives, providing improved access to multi-agency, multidisciplinary remote sensing and in-situ oceanographic data in support of open science and applications for societal benefit. Here we provide an overview of COVERAGE, an initiative and 3-year pilot project within CEOS involving interagency participation. Emphasis is placed on describing technology and architectural aspects, available data, web-based tools and services in support of associated ecosystem thematic demonstration applications. International collaborative aspects of the project are discussed, including COVERAGE involvement in the UN Decade of Ocean Science for Sustainable Development.**



***Keywords—NASA, CEOS, informatics, cloud computing, information systems, web services, analytics visualization, remote sensing, in-situ, ocean data, integration, ecosystem applications, UN Decade of the Oceans.***


## I. INTRODUCTION

There is an imperative to better marshal and integrate the growing variety and volume of oceanographic data in support of open science and inter-disciplinary marine applications for societal benefit. However, this ocean of data remains overly fragmented and is characterized by heterogeneous, often complex modes of access across agency and domain-specific repositories. There is, therefore, a need for data systems and services that will supplement these existing data infrastructures to enable the necessary integration and synthesis of multidisciplinary satellite and in-situ ocean observations so as to more fully realize their potential. Where effective, these capabilities will facilitate broader uptake and synergistic use of ocean data to support both established and emerging applications involving a wider range of end-user communities. There is also a related need for the implementation of technology platforms delivering value-added data services; a set of advanced core capabilities and data that can be reused and augmented as necessary, and that are cloud-enabled and can be spun up and scaled agilely to support a suite of regional applications. This is the critical gap that the CEOS (Committee on Earth Observation Satellites) Ocean Variables Enabling Research and Applications for GEO (COVERAGE) initiative seeks to address.

Here we provide a general overview of COVERAGE as a community driven effort and formal NASA-led initiative within CEOS. We discuss the COVERAGE web-portal with its associated cloud-enabled data services and web-based tools, including our interactive Data Viewer, back-end satellite imagery services, analytics services, data access services, and overall system architecture. Constituent Earth Observation data and COVERAGE's data product development work are also presented together with a description of our thematic ecosystem demonstration applications and engagement with the UN Decade of the Oceans.

## II. OVERVIEW

COVERAGE is a cross-cutting, NASA-led initiative within CEOS, the Committee on Earth Observation Satellites, involving several agencies and GEO programs internationally. It is motivated by the need to address constraints to access and synergistic use of Earth Observations in interdisciplinary applications, particularly amongst emerging user communities with a need for such environmental information. These include applications related to marine biodiversity and ecosystem resource management. COVERAGE involves development of a technology platform and reusable, open source toolkit providing access to complementary satellite and in-situ datasets from distributed sources via a core set of value-added data

services. It facilitates access to a curated collection of fit-for-purpose, interagency satellite data products from the 4 Ocean Virtual constellations (sea surface temperature, ocean color radiometry, ocean surface topography, and ocean vector winds) at a common resolution chosen as a baseline dataset, with a focus on supporting additional essential parameters (e.g. remotely sensed sea surface salinity) and higher resolution products next. Utility of the approach is demonstrated in the context of a pilot thematic ecosystem application, currently with a marine biodiversity and high seas fisheries focus, but with other use cases also being explored. COVERAGE has adopted a phased, stakeholder driven development approach. Technical requirements gathering, a data inventory exercise, and user needs assessment were undertaken during an initial 9-month Phase A activity. Phase B involved implementation of the current prototype capability over the course of 12-months, with an 18-month Phase C project currently underway that will involve further technical enhancements and regional spin-off applications.

## III. Constituent Data

COVERAGE includes a coherent set of gap-free, multi-mission (Level 4) satellite remote sensing products at a common 0.25-degree gridding resolution for a range of essential ocean variables. Summarized in figure 1, this collection of high-value, fit-for-purpose, near real-time and delayed mode interagency datasets span the ocean virtual constellations plus key, new satellite sea surface salinity products from ESA's Climate Change Initiative (CCI) and NASA's Salinity Continuity Program. Selection of this curated set of data products occurred during initial project scoping (Phase A). This involved expert consultations and was based on a range of standard selection criteria relating to product quality, resolution, spatial and temporal coverage, latency, format interoperability and access mechanisms.

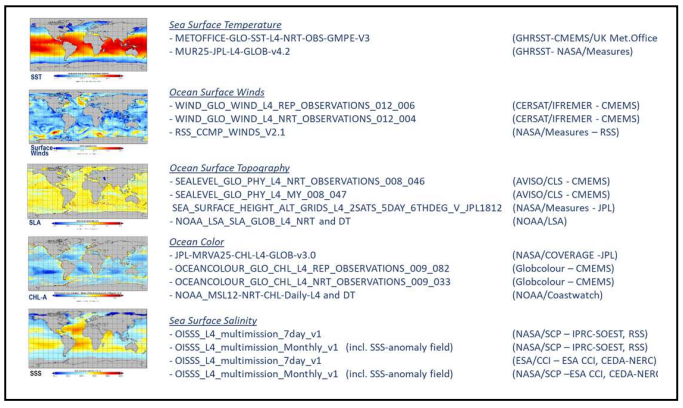



Fig. 1. Summary of high value, interagency satellite Level 4 (L4) near real-time and delayed time/reanalysis data products included in COVERAGE as a coherent baseline dataset spanning key ocean variables.

In support of our regional biological thematic applications for the Sargasso Sea and Eastern Tropical Pacific, COVERAGE additionally integrates some select value-added satellite datasets. These currently include the GEO-MBON Seascapes pelagic habitat classification product served via NOAA CoastWatch [1] and the Sargassum Floating Algal Index product for the Bermuda region produced by the satellite-based Sargassum Watch System (SaWS) at the University of South Florida [2].

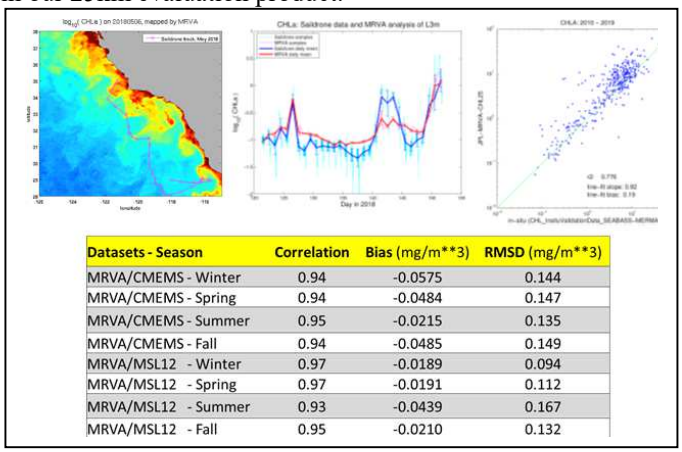
In-situ biological datasets integrated in support of the current ecosystem thematic pilot include: 1) animal telemetry datasets for various pelagic fish and marine mammal species from several agencies (e.g. NOAA, CSIRO, IMOS); 2) publicly available long-term (7 decades) fisheries catch and effort series for multiple species from the 5 tuna Regional Fisheries Management Organizations (RFMOs); and 3) extensive Automatic Identification System (AIS) fishing vessel positional data since 2012 available from Global Fishing Watch. Additionally, to demonstrate COVERAGE's, generalized oceanographic in-situ data support capability, we have recently successfully integrated the full suite of NASA "Salinity Processes in the Upper Ocean Regional Study (SPURS)" field campaign data. These are complex and heterogeneous data from more than thirty instrument platforms deployed during the SPURS1 and 2 campaigns in regions with contrasting salinity regimes (subtropical North Atlantic and Eastern Tropical Pacific respectively) [3, 4]. Integration of these data within COVERAGE's advanced visualization tool enables virtual reconstruction of the SPURS "sensor web" and interactive exploration of detailed in-situ observations in relation to broader scale variability from remotely sensed fields.

## IV. Ocean Color Data Product Development

While undertaking the aforementioned data inventory and product selection activity, the absence of a NASA level 4 (L4), gap free ocean color dataset was noted. We thus subsequently endeavored to address this gap during our Phase B project by development of a chlorophyll-A (CHLA) product complementary to the widely used, NASA flagship L4 sea surface temperature (SST) dataset, "MUR-SST" [5]. Both implement the Multi-resolution Variational Analysis (MRVA/MUR) wavelet-based algorithm that ensures: 1) accurate preservation of input data geolocations without truncation, and 2) reliable output at one or more gridding scales. The methodology preserves the location of fronts resolved via satellite observations, critical for biological and other applications.

This work resulted in the implementation and validation of a COVERAGE L4 CHLA MRVA global 25km, daily, near real-time product with associated uncertainty fields [6]. A useful dT flag variable is also included that indicates the latency (in hours) between the observation and synoptic times for each pixel value. This allows users to segregate the pixels based on timeliness. When this flag is set to zero and, what is returned is essentially Level 3, non-interpolated data from the same product, which essentially are only the observed cloud free data values. The complete 25km CHLA MRVA series has been processed and integrated into the range COVERAGE services as an ongoing forward stream. Preliminary work undertaken also shows considerable potential to extend this approach to

high resolution (1-2km) based on L2 source MODIS and VIIRS data inputs as opposed to the 4km L3 data currently being used in our 25km evaluation product.



| Datasets - Season | Correlation | Bias (mg/m**3) | RMSD (mg/m**3) |
|---|---|---|---|
| MRVA/CMEMS - Winter | 0.94 | -0.0575 | 0.144 |
| MRVA/CMEMS - Spring | 0.94 | -0.0484 | 0.147 |
| MRVA/CMEMS - Summer | 0.95 | -0.0215 | 0.135 |
| MRVA/CMEMS - Fall | 0.94 | -0.0485 | 0.149 |
| MRVA/MSL12 - Winter | 0.97 | -0.0189 | 0.094 |
| MRVA/MSL12 - Spring | 0.97 | -0.0191 | 0.112 |
| MRVA/MSL12 - Summer | 0.93 | -0.0439 | 0.167 |
| MRVA/MSL12 - Fall | 0.95 | -0.0210 | 0.132 |

Fig. 2. L4 CHLA MRVA product outputs and validation results: (Clockwise from top left) 1. gap-free CHLA output for the Baja, CA region; 2. Baja saildrone campaign along-track CHLA measurements versus colocated MRVA CHLA values; 3. MRVA CHLA versus collocated in-situ CHLA vlaues from NASA/SEABASS. 4 .summarized L4 CHLA product intercomparison results (NOAA-MSL12, CMEMS-Globcolour, COVERAGE MRVA).

## V. Distributed Architecture

COVERAGE implements a data system architecture with the ability to integrate both satellite and in-situ oceanographic data from remote sources, served via a range of common data distribution service technologies. Figure 3 illustrates schematically COVERAGE's high-level distributed data architecture developed during the Phase B prototype system implementation activity in support of our ecosystem/fisheries thematic pilot demonstration.

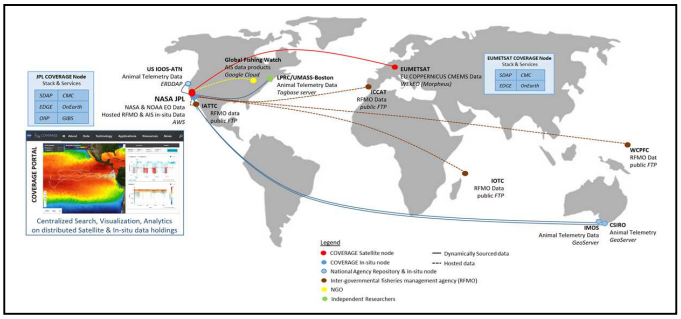


Fig. 3. High-level schematic of the prototype distributed data architecture implemented in support of our pilot ecosytem/fisheries thematic application.

The red symbols signify the 2 COVERAGE data nodes hosted within the NASA/JPL Amazon Web Services (AWS) and EUMETSAT WEkEO cloud environments running our software stack and serving US (NASA, NOAA) and European (Copernicus Marine Environmental Monitoring System – CMEMS) satellite ocean datasets respectively. Dynamic connectivity to in-situ data providers at IMOS and CSIRO in Australia and US-IOOS (blue symbols) serving relevant animal telemetry data via GeoServer and ERDDAP respectively is supported. COVERAGE also integrates extensive historical pelagic fisheries catch/effort data for highly migratory species publicly available from the 5 tuna commission regional fisheries management organizations (RFMOs), and high volume AIS data on fishing vessel movements provided by Global Fishing Watch.

## VI. Web-Portal

Integrated access to these datasets is though COVERAGE data services and tools available on our web-portal (https://coverage.ceos.org/) and via associated web-service APIs. The website additionally provides descriptive information on the project, a Resources Area with documentation, pointers to our open source software, tutorial demonstrations, and other project media online via our YouTube channel [7]. There is also a News section on the website, listing events and announcements on COVERAGE activities that is integrated with our project Twitter feed [8]. Incorporation of Google Analytics enables capture of site traffic and the detailed monitoring of usage metrics.

## VII. Web-based Data Visualization

COVERAGE provides a unique and advanced web-based tool for integrated, dynamic visualization of satellite and in-situ data of various types, including station point, profile, and trajectory series. Illustrative examples ranging from research cruise data, to autonomous saildrones, to animal telemetry data, and RFMO species catch and effort data in relation to key satellite data layers are shown in figure 4.

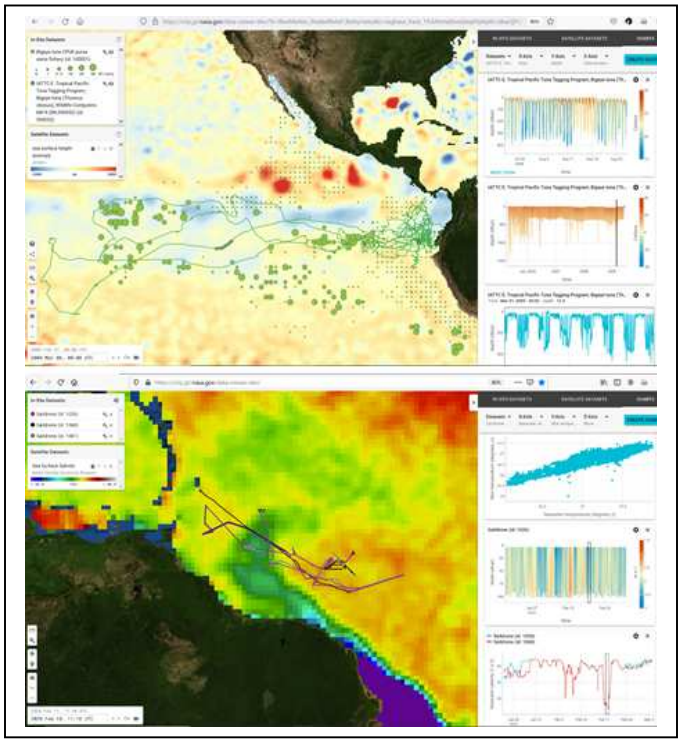


Fig. 4. COVERAGE's web-based data visualization tool illustrating interactive mapping and charting capabilities for a representative range of satellite and in-situ datasets within the system: IATTC Bigeye Tuna archival tag and spatial catch distribution data relative to AVISO-SSHA and animal telemetry environmental measurements (top); Saildrone ATOMIC cruise ADCP and CTD data overlaid on Sea Surface Salinity data from SMAP (bottom).

The COVERAGE Data Viewer is based on JPL's open source Common Mapping Client (CMC) software framework [9] and provides a well-rounded visualization capability for ocean data. Novel features include tightly synchronized horizontal and vertical views of data and their evolution over time, with flexible control of time stepping intervals. A dataset Search capability – allowing filtering by spatio-temporal criteria, plus project, platform, sensor, and measurement variable criteria – is also included, as is a "One-stop" Data Sub-setting capability for both satellite and in-situ data. There is integrated online help, and a useful URL sharing feature that allows exact reproduction and exchange of detailed data map overlay and charting views that is particularly helpful in collaborative and instructional settings.

## VIII. Imagery Services

The Data Viewer leverages a set of back-end imagery services for dynamic provisioning of a range of satellite imagery layers spanning multiple core parameters for rendering in our browser-based visualization tool (Figure 5).

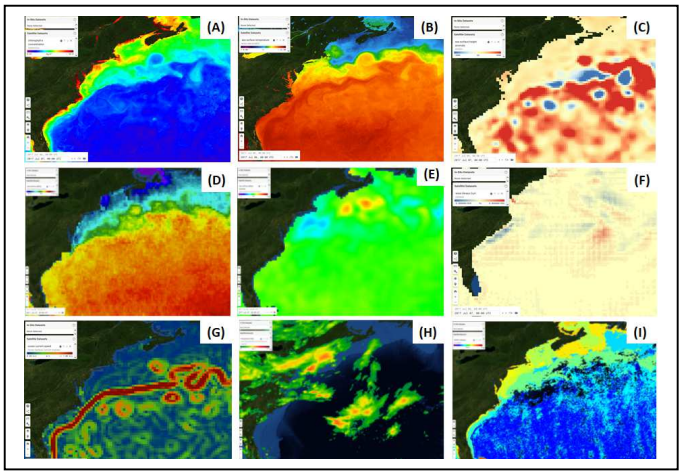



Fig. 5. Imagery layers of key ocean parameters provisioned via the COVERAGE OnEarth WMTS server for a range of global, multi-mission satellite remote products as rendered by our Data Viewer: (A) - L4 Globcolour CHLA Daily, near real-time (NRT), (B) L4 JPL MUR-SST Daily, NRT, (C) L4 AVISO Sea Surface Height Anomaly Daily, NRT, (D) L4 ESA-CCI OI Sea Surface Salinity Weekly, Reannalysis, (E) L4 U.Hawaii/RSS OI Sea Surface Salinity Anomaly Monthly, Reannalysis, (F) L4 IFREMER/CERSAT Wind Stress Curl 6-hour, NRT, (G) L4 ESR OSCAR Ocean Current Velocity, 5-day, (H) L4 IMERGE Precipitation Rate, 0.5-hour, NRT, L4 NOAA/MBON Seascapes Habitat Classification, Monthly.

These imagery services are also publicly accessible from COVERAGE via application service interfaces (APIs) for consumption by third party software client applications. Services include Open Geospatial Consortium (OGC)-compliant imagery delivery protocols through the JPL-developed OnEarth Web Map Tile Service (WMTS) [10] server hosted by COVERAGE on both our AWS and WEkEO cloud nodes. OnEarth is used to both serve and undertake routine imagery processing of the majority of our datasets (Figure 6). NASA's Global Imagery Browse Service (GIBS) [11] provides complimentary access to some select layers. COVERAGE also provisions imagery for all our US and European datasets via WMS services from our THREDDS servers deployed on both COVERAGE cloud nodes.

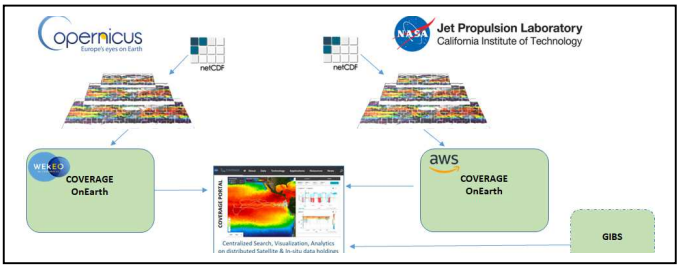



Fig. 6. Schematic of OnEarth WMTS server deployments on COVERAGE AWS and WEkEO cloud nodes processing and serving imagery data for consumption by the Data Viewer and third party clients.

## IX. Analytics Services

COVERAGE provides analytics service support for included remote sensing datasets via the "Science Data Analytics Platform" (SDAP). SDAP is an open source Apache.org incubator software project that enables big data science without the need for voluminous data download [12]. It facilitates on demand processing closer to the source data hosted in the cloud. A range of built in, parallel computing functions for spatio-temporal series summarization are supported, with SDAP extensions to implement a generalized satellite to in-situ matchup capability currently under development with NASA support.

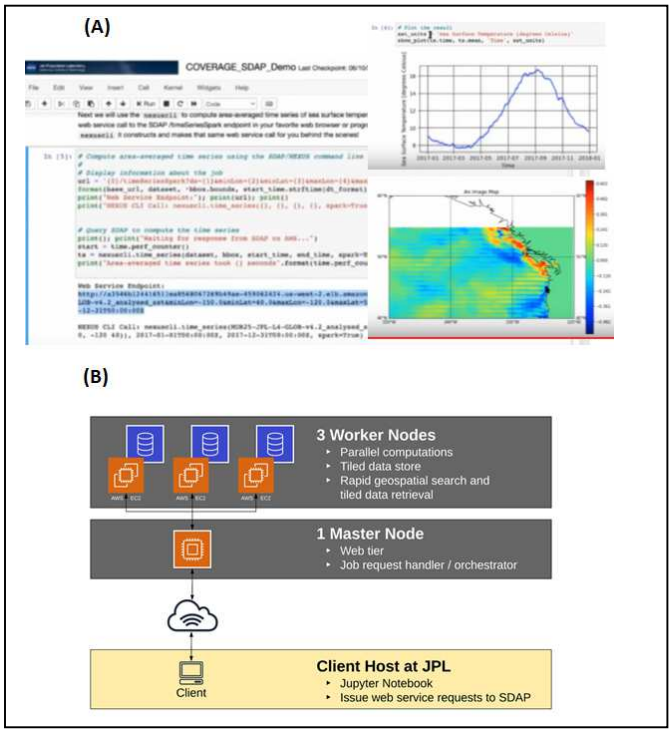



Fig. 7. Illustration of Science Data Analytics Platform: (A) Jupyter notebook interface, and (B) schematic of SDAP software stack deployement on WEkEO.

For COVERAGE, Kubernetes deployments of the SDAP software stack are in place at both our NASA/JPL AWS and EUMETSAT WEkEO cloud computing environments. This provides analytics capabilities for the range of interagency datasets integrated in COVERAGE (Figure 7). Interfaces to these SDAP instances are programmatically via SDAP APIs

and Jupyter notebooks. Enhancements enabling future use of the COVERAGE Data Viewer for interactive geospatial analytics query definition and result display are being implemented.

## X. Data Access Services

COVERAGE provides a range of familiar data access services available via Thematic Real-time Environmental Distributed Data Services (THREDDS) data servers that we host. While THREDDS is an established technology, what is novel here is that we are running a version of THREDDS that is cloud-enabled. This is capable of exposing collections of satellite data series maintained as native netCDF files within COVERAGE AWS S3 buckets and WEkEO object stores (Figure 8). Both THREDDS standard web-form interfaces and Jupyter notebooks tapping THREDDS APIs programmatically can be employed to access a variety of standard services powered by THREDDS 5 instances deployed on both our cloud nodes. Supported capabilities include several familiar and widely used THREDDS services and APIs: NCCS (subsetting), OPeNDAP, WMS, WCS, and ISO metadata plus a couple of interesting new services such as CDM-remote. The open source THREDDS 5 technology, developed by UCAR/Unidata [13], now additionally provides support for in-situ, discrete spatial geometry data in addition to gridded data types . Integration of forward data file streams on an ongoing basis into COVERAGE THREDDS, SDAP, OnEarth Imagery and Visualization services is achieved by the generalized "Data Harvester" software developed by our project and soon to be made available open source.

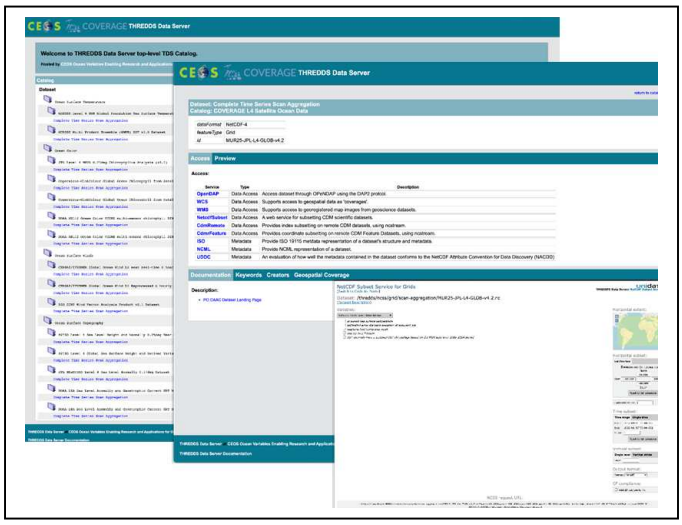



Fig. 8. COVERAGE THREDDS catalogs and services on AWS and WEkEO cloud nodes providing access to our interagency satellite data collection.

## XI. Thematic Applications

COVERAGE seeks to demonstrate the utility of both its technologies and data in the context of thematic regional spinoff applications that build upon our current ecosystem prototype. The first is in partnership with the Sargasso Sea Commission in support of its GEF and FFEM projects with UNDP and IOC-UNESCO involvement for the Sargasso region. It relates to conservation of biodiversity in areas beyond national jurisdiction. This is a complex, multi-faceted project involving development of a Ecosystem Diagnostic Analysis (EDA) for the region, identifying trends and impacts from available environmental, biological and socio-economic data. This in turn will support the development of an action plan towards adoption of an ecosystem-based stewardship approach for the Sargasso Sea. COVERAGE's role as a partner of the Commission within the scope of Phase C is to facilitate access to key satellite data and provide the integrative data system capabilities enabling the development and communication of ecosystem indicators for scientific and policy assessments.

A second regional application focusing on the Eastern Tropical Pacific (ETP) is in partnership with the Inter-American Tropical Tuna Commission (IATTC). IATTC is a 21 Nation Intergovernmental Regional Fisheries Management Organization (RFMO) responsible for the scientific assessment and management of large pelagic fisheries in the ETP. There is considerable interest and ongoing work involving the use of ocean remote sensing and model data to inform analyses of species habitat under environmental regime shift, spatial catch forecasting, by-catch mitigation, fishery closed area and MPA designation. As part of a collaboration between COVERAGE and IATTC during Phase C, the objective is to build upon the prototype application that integrates publicly accessible RFMO data and augment these with select, detailed fisheries observer datasets and higher resolution remote sensing datasets to support IATTC quantitative statistical analysis workflows via COVERAGE analytics capabilities. As an initial use case, we will seek to recreate a published analysis to demonstrate proof of concept.

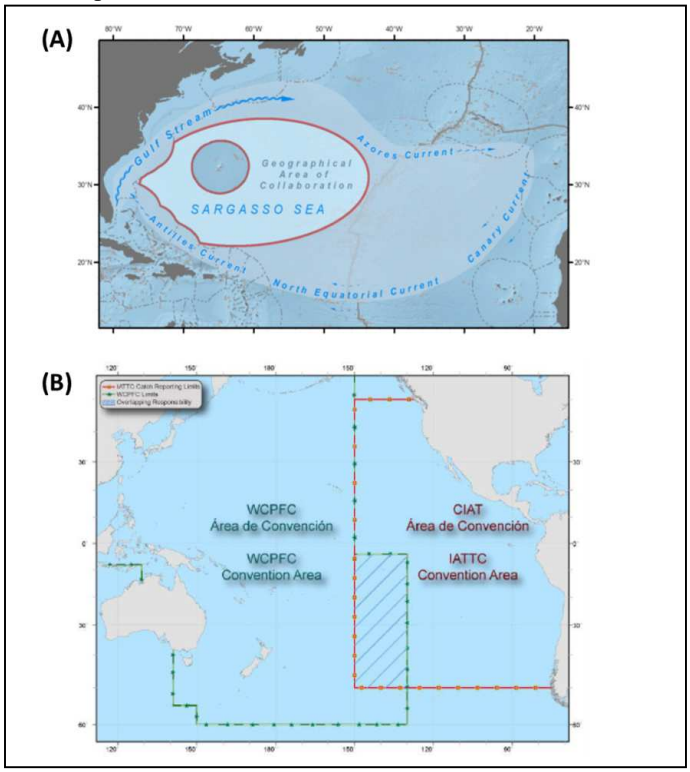



Fig. 9. Regioanl COVERAGE ecosystem/fisheries application pilots: (A) Sargasso Sea in collaboration with the SSC, and (B) E. Tropical Pacific in partnership with IATTC.

## XII. Community & UN Ocean Decade Engagement

COVERAGE places considerable emphasis on community and stakeholder engagement as a user-driven development. Over time we have presented at a range of relevant conferences [14], and regularly are involved in consultations with key stakeholders, including CEOS, agency partners and collaborators. We have also held 3 workshops over the past couple of years, the latest being at Ocean Sciences 2022 [15].

COVERAGE is additionally serving as a cumulative contribution and point of contact on behalf of CEOS to the UN Decade of Ocean Science for Sustainable development. Our "ocean shot" concept submission to Ocean Decade U.S. relating to next generation data infrastructures for a digitally integrated ocean observing system [16] is resonating and has been highlighted in the recent National Academies Ocean Studies Board report on Cross-cutting Themes for U.S. Contributions to the U.N. Decade of the Oceans [17]. We are committed to advancing this concept further under COVERAGE and are exploring alignment and collaborations with other Decade program initiatives of relevance and that we have connections to. This includes: DITTO (Digital Twins of the Oceans), Marine Life 2030 (Smithsonian National Museum of Natural History), and Basin Scale Events to Coastal Impacts (BECI).

## Acknowledgments

The research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). We acknowledge funding support for the COVERAGE project from the NASA Physical Oceanography and NASA ESDIS programs. We also thank our CEOS and other agency partners, including EUMETSAT in particular for their WEkEO cloud-computing resource contributions to the project.

## References

[1] https://coastwatch.noaa.gov/cw/satellite-data-products/multi-parameter-models/seascape-pelagic-habitat-classification.html

[2] https://optics.marine.usf.edu/projects/saws.html

[3] Special Issue: SPURS: Salinity Processes in the Upper-ocean Regional Study. Oceanography, vol. 28(1), pp. 1-115, March 2015

[4] https://podaac.jpl.nasa.gov/spurs

[5] Chin,T.M., Vazquez-Cuervo, J., Armstrong,E.M. 2017. A multi-scale high-resolution analysis of global sea surface temperature. Remote Sensing of Environment, vol.200, pp.154-169, https://doi.org/10.1016/j.rse.2017.07.029.

[6] Chin, T.M, V.M Tsontos, J. Vazquez. 2021. Technical report and product guide for the NASA/JPL LEVEL-4 MRVA CHL-A 0.25 Degree Spatial Resolution Dataset (version 3.0) developed under the NASA/COVERAGE project. p31. doi:10.6084/m9.figshare.13003928.v3

[7] https://tinyurl.com/coverage-channel

[8] https://twitter.com/coverage_ceos

[9] https://coverage.ceos.org/cmc/

[10] https://coverage.ceos.org/onearth-api/

[11] https://coverage.ceos.org/gibs/

[12] https://incubator.apache.org/clutch/sdap.html

[13] https://www.unidata.ucar.edu/software/tds/

[14] https://coverage.ceos.org/news/

[15] Tsontos,V.M., Vazquez-Cuervo,J., Huang,T., Chin,T.M., Roberts,J.T., Jacob,J.C., Platt, F., Yathirajam, V. 2022. Synopsis of a Consultative Workshop at the Ocean Sciences Meeting 2022 on the CEOS Ocean Variables Enabling Research and Applications for GEO (COVERAGE) Initiative. Bulleting of Limnology & Oceanography, vol.31(2), pp.53-55. 16 April 2022, https://doi.org/10.1002/lob.10495.

[16] Tsontos, V.M. 2021. COVERAGE: Next Generation Data Service Infrastructure for a Digitally Integrated Ocean Observing System in Support of Marine Science and Ecosystem-Based Management. Marine Technology Society Journal, vol.55(2), pp. 70-71(2), May/June 2021,https://doi.org/10.4031/MTSJ.55.3.45

[17] Cross-cutting Themes for U.S Contributions to the U.N. Decade of the Oceans. 2022. Ocean Studies Board, Earth and Life Sciences. A Concensus Report of the National Academy of Sciences, Engineering & Medicine. The National Academies Press, p128. https://doi.org/10.17226/26363