2021 Competition: COM, CVP and GOMO - **Innovative Ocean Dataset/Product Analysis and Development for support of the NOAA Observing and Climate Modeling Communities**

Competition Number: 2864453

Funding Opportunity Number: NOAA-OAR-CPO-2021-2006389

In FY21, this multi-program competition solicited proposals that would increase the usefulness of NOAA observations to the NOAA observing, modeling and/or satellite communities, and/or the broader scientific community by one of the following:

1. Developing new observation-based (in situ, satellite) ocean synthesis datasets or products (physical and/or biogeochemical) for climate monitoring or modeling applications through applying existing methods or developing new, state-of-art, innovative methods and approaches (e.g. ocean state estimation, data assimilation, and quantification of observational uncertainty).
2. Evaluating current methods and approaches for ocean observing and modeling, and the ability of observed and modeled data/products to reproduce physical or biogeochemical processes, climate phenomena, or interactions between Earth System components on different timescales. Proposals should examine bias(es) in observed and modeled data/products and advance understanding of the cause(s) for large differences between observed and modeled ocean data/products. Collaborations between the ocean observing and climate modeling community are strongly encouraged for Type 2 proposals.

***projects that received COM funding

Full information sheet for Competition Number: 2864453 (link here)

NOAA-OAR-CPO-2021-2006389

An air-sea flux, SST, wave database from the ATOMIC field program

Elizabeth J. Thompson (NOAA/PSL), Darren Jackson (CU CIRES / NOAA PSL), Christopher W. Fairall (NOAA/PSL), Dongxiao Zhang (CU CIRES / NOAA PSL) 4

Air-Sea Fluxes Derived from Global Surface and Field Campaign Observations: A Key Product for Improving Earth System and Climate Models

Shuyi S. Chen, Professor (University of Washington), Chris Fairall (NOAA/ESRL/PSD) 5

***__Biogeochemical Argo synthesis products of oxygen, nitrate, and pH for increased community utilization of autonomous profiling observations__

Seth Bushinsky (University of Hawai'i at Mãnoa), Nancy Williams (University of South Florida), Andrea Fassbender (NOAA/PMEL) 6

Developing a capability for the real-time comparison of near surface ocean with operational analysis products

Arun Kumar (NCEP/NWS/NOAA), Meghan F. Cronin (NOAA/PMEL), Jieshun Zhu, Caihong Wen (NCEP/NWS/NOAA), Dongxiao Zhang, Jessica Masich (NOAA/PMEL) 7

***__Developing an in situ-satellite Blended Marine Air Temperature Dataset Using Artificial Intelligence__

Jessica L. Matthews, Yuhan (Douglas) Rao (North Carolina State University/North Carolina Institute for Climate Studies), Lei Shi, Meteorologist (NOAA/NCEI) 8

***__Developing PSSdb: a Pelagic Size Structure database to support biogeochemical modelling__

Jessica Y. Luo (NOAA/GFDL), Rainer Kiko, Lars Stemmann (Sorbonne Université), Charles Stock (NOAA/GFDL), Todd O'Brien (NOAA/NMFS) Fabien Lombard, Jean-Olivier Irisson (Sorbonne Université) 9

Evaluating Earth System Models using apparent relationships rather than spatial fields- using neural networks as model comparators

Anand Gnanadesikan (Johns Hopkins University) 10

***__Global Open Ocean Data-Products for Biogeochemical Argo Research__

Brendan R. Carter (CICOES/University of Washington), Richard A. Feely, Andrea J. Fassbender (NOAA/PMEL) 11

***__Gridded data products with uncertainties for 21st century in-situ oceanographic observations__

Donata Giglio (University of Colorado Boulder), Mikael Kuusela (Carnegie Mellon University) 12

***Innovative analysis of deep and abyssal temperatures from bottom-moored instruments

R. C. Perez, C. S. Meinen, S. Dong (NOAA/AOML/PHOD); Senior personnel: D. Volkov, M. Le Hénaff 13

***MarineFlux: A user friendly in-situ marine turbulent flux data service

Shawn Smith, Mark Bourassa (Florida State University, Tallahassee) 14

Uncertainty quantification, data quality, and observing network design in an ocean state of the Tropical Pacific

Patrick Heimbach (The University of Texas at Austin), Matthew Mazloff (University of California San Diego), Aneesh Subramanian (University of Colorado Boulder) 15

An air-sea flux, SST, wave database from the ATOMIC field program

Elizabeth J. Thompson (NOAA/PSL), Darren Jackson (CU CIRES / NOAA PSL), Christopher W. Fairall (NOAA/PSL), Dongxiao Zhang (CU CIRES / NOAA PSL)

Field campaigns have been designed to collect in-situ data that both improve process level understanding and form benchmark datasets for experiments with and improvements of reanalysis, satellite data, and models. However, the transition of observations to model, satellite, and reanalysis projects rarely happens in a timely fashion or with full participation between all expert parties involved. This proposed data synthesis effort will produce a more effective and lasting use of ATOMIC field campaign observations for evaluating satellite and reanalysis products and use in numerical models.

The ATOMIC 2020 field campaign in the northwestern tropical Atlantic Ocean collected a unique and diverse dataset of air-sea interaction, ocean properties, and clouds from drifting, shipborne, airborne, air-deployed, and uncrewed platforms. The variety of measurement heights and depths, platform motions, and instrument uncertainties is at the same time a strength and a usability barrier. To invite effective and lasting use by modeling, satellite, and reanalysis teams, the observation teams must account for these details and synthesize the measurements in a common, standard, open source database. In the first year of the project we will convene a joint conference to obtain guidance from modelers on critical variables and data formats.

The air-sea fluxes, sea surface temperature (SST), and surface waves are the three most important sets of information for coupled weather and climate models to resolve correctly to understand and predict environmental change. Our hypothesis is that the strengths in terms of model, satellite, and observational research applications of a combined, consistent ATOMIC database of SST, fluxes, and waves from multiple platforms far outweighs the strengths of any individual dataset. We will make use of periods of co-located data between in situ and airborne platforms plus satellite overpasses, with the NOAA Ship *Ronald H*. Brown serving as the standard for comparison. We will draw on ATOMIC and EUREC4 A data for the most complete picture to produce and archive a combined, consistently quality controlled, and consistently processed database from all platforms. Databases for SST, air-sea fluxes, waves, plus near-surface bulk variables that result from this effort will be posted to NCEI for public distribution.

By bringing observations to the modeling and satellite communities and facilitating partnerships across the external and NOAA communities, the proposed data synthesis directly responds to the COM/CVP/GOMO call to “develop an observations-based product for climate monitoring or modeling application” that “increases the use of NOAA’s historical field campaign data” and “enables improved climate modeling or monitoring (e.g., enables future climate model evaluation, validation, process-oriented diagnostics)”, contributing to the first two objectives of NOAA’s long- term climate goals as described in NOAA’s Next-Generation Strategic Plan.

Air-Sea Fluxes Derived from Global Surface and Field Campaign Observations: A Key Product for Improving Earth System and Climate Models

Shuyi S. Chen, Professor (University of Washington), Chris Fairall (NOAA/ESRL/PSD)

This proposed study will focus on the first program priority: ***Developing new observation-based (in situ, satellite) ocean synthesis datasets or products (physical and/or biogeochemical) for climate monitoring or modeling applications through applying existing methods or developing new, state-of-art, innovative methods and approaches (e.g. ocean state estimation, data assimilation, and quantification of observational uncertainty).***

Fluxes of mass, energy, and momentum between the oceans and the atmosphere have a profound impact on global weather and climate. The air-sea fluxes are of critical importance as they represent and control the energy and water cycle of the Earth system. It plays a key role the global hydrological cycle and precipitation, which are difficult for the current Earth system and climate models to accurately represent. However, our current ability of observing the air-sea fluxes globally is limited and remains as an unmet challenge. ***This proposed project will aim at 1) increasing the use of NOAA's field campaign data (e.g., DYNAMO, YMC, ATOMIC and others) and emerging observations from sustained observing networks and systems, which will use a combined with non-NOAA observations that are available, and 2) enabling Earth system and climate model evaluation, validation, process-oriented diagnostics, and/or satellite calibration and validation.***

We will **develop an observation-based air-sea fluxes data product**, which will integrate observations from existing surface-based network (e.g., moorings, ships and saildrones), selected satellite measurements, and field campaigns from **1990-2020**. This air-sea fluxes data product with an interactive user interface (leverage an existing software that will be adapted for the air-sea flux datasets in this project) will be particularly useful the broad modeling community, not only the modeling centers but also research community including students who are not familiar with field observations.

We will **engage and work in collaboration with Earth system and climate modeling community from NOAA, NCAR, and others.** We plan to help developing model evaluation methods that are based on the observations of a wide range of temporal and spatial scales that can address both model physical process and model performance.

The PI Chen and Co-I Fairall have worked with field campaign and ship and surface-based observation in producing air-sea sensible and latent heat fluxes (e.g., Fairall et al. 2003, 2010; Chen et al. 2016). The air-sea momentum flux will be derived using the observations and a high-resolution coupled atmosphere-wave-ocean model to access and refine the flux in high wind conditions (e.g., Fairall et al. 2009, Chen et al. 2013). The air-sea fluxes product produced in the project will be made available through a user interface already developed by PI Chen and a research scientist Dr. Brandon Kerns at UW (http://dynamo.ml-ext.ucar.edu/dynamo_legacy/). This will ensure a rapid and easy access for a broader user community.

***Biogeochemical Argo synthesis products of oxygen, nitrate, and pH for increased community utilization of autonomous profiling observations

Seth Bushinsky (University of Hawai'i at Mãnoa), Nancy Williams (University of South Florida), Andrea Fassbender (NOAA/PMEL)

Changes in ocean chemistry associated with anthropogenic pollutants are threatening marine ecosystems around the world. Our ability to quantify the rates of these transformations and forecast future conditions depends on comprehensive and accurate observations that hone our mechanistic understanding of ocean physics and biogeochemistry. The recent expanded use of profiling floats equipped with biogeochemical (BGC) sensors presents an opportunity to vastly increase the temporal and spatial resolution of observational datasets useful for studying ocean health and constraining global climate models. However, there is presently no unified BGC float dataset that integrates all U.S. and international float observations, identifies and adjusts any calibration inconsistencies or other biases, and provides the data in an easy-to-use format for observationalists, modelers, and prospective new users.

This proposal responds to the NOAA Climate Program Office FY2021 COM and CVP call: Innovative Ocean Database/Product Analysis and Development for support of the NOAA Observing and Climate Modeling Communities. Here we propose to comprehensively assess and adjust observations of oxygen, nitrate, pH and derived dissolved inorganic carbon from the international BGC float database to produce an internally consistent, global dataset for scientific use. The proposed adjusted dataset will rely on comparisons between float and shipboard data in the deep ocean, below the research of most anthropogenic changes and where we know these tracers are stable. Preliminary analyses suggest that regional and float-specific biases exist, which must be corrected prior to the use of BGC float data for regional assessments or for long-term trend analyses. We will create (1) an adjusted float dataset; (2) a monthly 1°x1° gridded, bin-averaged float dataset on depth and density intervals; and (3) an optimally interpolated dataset that incorporates both ship and float data. The produced datasets, code for making and using them, and our insights into persistent biases in float BGC measurements will be shared with the global BGC float community, in which the project PIs are heavily involved.

In addition to developing datasets, we will demonstrate their value through analyses of long-term trends in deoxygenation and anthropogenic carbon accumulation in the Southern Ocean - a high-density region for the BGC floats. The development of standard format float and combined float-shipboard datasets will increase the use of BGC data by observationalists and modelers. Additionally, the project methods and code developed will establish a framework for secondary and independent quality control of the global, international BGC float observing system. This tool is anticipated to be of significant use to NOAA, which now has multiple BGC Argo pilot projects underway. These synthesized datasets will further NOAA's goals of accurately monitoring ocean health while providing a comprehensive, accessible dataset necessary for improved numerical model assessment of marine biogeochemistry and ecosystem projections.

Developing a capability for the real-time comparison of near surface ocean with operational analysis products

Arun Kumar (NCEP/NWS/NOAA), Meghan F. Cronin (NOAA/PMEL), Jieshun Zhu, Caihong Wen (NCEP/NWS/NOAA), Dongxiao Zhang, Jessica Masich (NOAA/PMEL)

Ocean observations are used for initializing and validating sub-seasonal to seasonal forecasts using coupled models; creating synthesis products, for monitoring evolution of ocean conditions; developing climate data records to monitor the influence of slow trends; and for improving understanding of the physics affecting the climate system. With continued investments in ocean observations, and upcoming enhancements in the tropical Pacific observing system (TPOS), increasing the utilization of ocean observations has been identified as one of the key challenges. There is also a consensus in the community that there is a longstanding disconnect between the investments in observations and their utilization in model-based analysis and prediction systems.

A real-time comparison between observations and model-based analysis products will represent a direct use of ocean observations for which a need has been long perceived but remains to be realized. Towards bridging the gap between the observation and modeling communities, the goal of this project is to develop a capability for the real-time comparison of in situ ocean observations with operational analysis products. The scope of the project will focus on the observations from moorings in the tropical Pacific. This is because El Niño-Southern Oscillation (ENSO) is one of the most important modes of coupled variability with largest societal impacts, and further, because the TPOS is currently going through evolutionary changes. To accomplish the goals of the project outlined above, the following tasks will be completed:

(a) Identify ocean and atmosphere mooring data to be used in the real-time model assessment and set up procedures to update the observational database in real-time.
(b) For atmospheric and ocean analyses, set up corresponding procedures to update the model database in real-time.
(c) Develop procedures to compare time-series of model analysis with observations and develop a web interface to disseminate the information to the community.
(d) Utilization of assembled observational and modeled data bases to address science questions of relevance in understanding coupled climate variability in the tropics.

The goals of the proposal are highly relevant to the focus of the present call to "...develop an observations-based product for climate monitoring or modeling application" that "enables improved climate modeling or monitoring (e.g., enables future climate model evaluation, validation, process-oriented diagnostics)". The project will also address other foci of the call for "Evaluating current methods and approaches for ocean observing and modeling, and the ability of observed and modeled data/products to reproduce physical or biogeochemical processes, climate phenomena, or interactions between Earth System components on different timescales."

***Developing an *in situ*-satellite Blended Marine Air Temperature Dataset Using Artificial Intelligence

Jessica L. Matthews, Yuhan (Douglas) Rao (North Carolina State University/North Carolina Institute for Climate Studies), Lei Shi, Meteorologist (NOAA/NCEI)

Marine air temperature drives climate processes which have far-reaching downstream impacts including an increase in extreme cyclones, sea level rise, and coastal flooding. Monitoring of this important climate indicator is imperative for decision makers and stakeholders to plan climate adaptation and mitigation efforts on local and regional scales: a long-term research goal for NOAA. As a key contributor to the estimation of global mean surface temperature, MAT observations are essential for climate monitoring. High quality MAT data also play a crucial role in understanding air-sea fluxes and its impact on the physical and biological processes in the ocean and coastal system. Several gridded MAT data sets (i.e., CLASSnmat, UAHNMATv1) have been developed using historical observations by gridding nighttime in situ MAT data into coarse grid (i.e., 5°×5°). Although valuable for long term climate analysis before the satellite-era, the existing in situ only datasets are insufficient for global and regional climate studies because of the sparse data coverage, cross-platform differences, and intrinsic data variability. To address this challenge, we propose the development of an innovative global high-resolution MAT dataset based primarily on blending high quality High-resolution Infrared Radiation Sounder (HIRS) retrieved temperatures and height-adjusted International Comprehensive Ocean-Atmosphere Data Set (ICOADS) observations using advanced machine learning approaches.

In direct response to the **Innovative Ocean Dataset/Product Analysis and Development for support of the NOAA Observing and Climate Modeling Communities** competition call, the project's two major scientific objectives are the 1) *development of a blended daily global MAT dataset* and 2) *application of the blended MAT dataset for climate monitoring*. A machine learning based model will be created using an existing framework, previously applied for land surface air temperature, to develop the relationship between the available reference in situ temperature observations and satellite retrieved temperature data along with ancillary radiation data. A pilot study over the tropical Pacific Ocean demonstrates that the machine learning framework can accurately estimate marine air temperature. The final global HIRSMAT will be a daily dataset on the common grid of 0.5° × 0.5° with daily mean, daytime, and nighttime MAT. The stability of the HIRSMAT dataset will be evaluated by cross comparison with other existing data sets that have been used for climate monitoring studies . Finally, we will use the monthly HIRSMAT anomaly dataset for climate analysis of 1979-present by examining the high-resolution spatial pattern of marine air temperature.

This project will result in a new observation-based (*in situ* and satellite) ocean synthesis dataset for climate monitoring and modeling applications. The new high-resolution global MAT dataset will be developed using a state-of-art machine learning framework. The improved spatial and temporal resolution and coverages provided by HIRSMAT will contribute to the associated NOAA objective to improve scientific understanding of the changing climate system and its impacts. Furthermore, the project will increase the use and utility of field campaign data observations (e.g., DYNAMO and ATOMIC) by using them for independent evaluation. The resultant dataset will be provided in an obs4MIPS-compliant format for efficient, ready access by the modeling community to perform comparisons with relevant CMIP6 output variables.

***Developing PSSdb: a Pelagic Size Structure database to support biogeochemical modelling

Jessica Y. Luo (NOAA/GFDL), Rainer Kiko, Lars Stemmann (Sorbonne Université), Charles Stock (NOAA/GFDL), Todd O'Brien (NOAA/NMFS) Fabien Lombard, Jean-Olivier Irisson (Sorbonne Université)

Marine plankton are essential components of ocean ecosystems, forming the bottom of the food chain and serving as controls on large-scale biogeographic patterns in ocean carbon, nutrients, and oxygen. Earth System Model (ESM) projections suggest that ocean warming and stratification will drive decreases in net primary production (NPP) and shifts in plankton community composition and size structure. These changes to the plankton community have subsequent implications for decreasing the strength of the biological pump, as well as declining fisheries productivity through trophic amplification mechanisms. However, a critical underlying factor driving these shifts – the simulated size structure of plankton and particles – is difficult to validate, as global datasets are lacking. Fortunately, new data streams from plankton imaging systems are capable of providing 3-dimensional, broad-spectrum views on plankton and particle size spectra, provided data are properly harmonized and cross-calibrated to a single standard.

Following the structure of the World Ocean Database (WOD) and COPEPOD database, we propose to establish a methodological processing pipeline for the ingestion, calibration, and harmonization of imaging data on plankton and particles spanning five orders of magnitude (1 micron - 10 cm). The ultimate goal of our work is the development of the Pelagic Size Structure database, or PSSdb, as a resource for global gridded data. We will start off with data from five imaging systems that include in-situ samplers to tabletop imagers: Imaging FlowCytoBot (IFCB), FlowCam, Underwater Vision Profiler (UVP), ZooScan, and In-situ Ichthyoplankton Imaging System (ISIIS). The resultant data will be structured in four levels, ranging from low (level 1) to high (level 3) taxonomic resolution, and will be made publicly available through NOAA websites.

Our methodological framework and resultant database will be broadly extensible, and able to incorporate both historical data and future sampling efforts. As plankton imaging systems become increasingly utilized in national and international sampling programs, and are being developed for automated vehicles, drifters, and floats, we anticipate both a growing need for data consolidation and harmonization, as well as the opportunity for these datasets to inform ESM development. Ultimately, an improved understanding of the key ecological mechanisms driving marine ecosystem shifts, elucidated by models with strong empirical validation, will increase confidence in Earth System Model projections and associated links between the climate and fisheries.

Evaluating Earth System Models using apparent relationships rather than spatial fields- using neural networks as model comparators

Anand Gnanadesikan (Johns Hopkins University)

While Earth System Models often fail to reproduce biological fields like phytoplankton biomass and chlorophyll, the reasons behind such failures are complex. Because phytoplankton growth rates depend on environmental conditions like nutrients and light, and these in turn depend on the rates of mixing and upwelling, physical biases in models can produce biases in circulation such that a “perfect” biological model will still give imperfect results. For example, an Earth System Model in which the relationship between macronutrients and biomass matches that found in the real North Atlantic will still produce a spatial bias in the distribution of nutrients if the path of the North Atlantic Current is poorly simulated. If we knew that the model had the correct relationship between biomass and nutrients, we could unambiguously tie such an error to model physics. However, the actual apparent relationships (those seen in the real world between environmental drivers and phytoplankton biomass) are far from simple and may deviate from intrinsic relationships based on bench science which are often coded into models. For example, low phytoplankton biomass may be associated with low levels of nutrients in the presence of high levels of light, or high levels of nutrients in the presence of low levels of light. Simply plotting biomass against nutrients will then result in a maximum biomass concentration at intermediate levels of nutrient, capturing the asymptote of biomass at high levels of nutrient may require careful extrapolation.

Better constraining the drivers of phytoplankton change and variability is essential to NOAA’s mission to improve the prediction of the Earth System in order to build resilience to changes. The proposal directly addresses the call within the competition to “examine biases in observed and modeled data/products and advance understanding of the causes for large differences between observed and modeled ocean data/products.” We aim to build on recent work showing that machine learning methods (in particular, Neural Network Ensembles) can be used to extract biologically reasonable complex relationships from ESMs and also used to compare the similarity of the biological codes across models. We propose to examine whether such methods can find robust relationships between biomass and observed environmental parameters on regional and global scales, and use the resulting relationships as metrics for evaluating Earth System Model output. We will do this using combinations of remotely sensed data (chlorophyll, carbon biovolume) and in-situ data (phytoplankton biomass, nutrients, Ekman upwelling, light, mixed layer depth). We will also develop a toolkit whereby Earth System Models that are part of the current IPCC process can be compared with observational relationships and to each other.

***Global Open Ocean Data-Products for Biogeochemical Argo Research

Brendan R. Carter (CICOES/University of Washington), Richard A. Feely, Andrea J. Fassbender (NOAA/PMEL)

Biogeochemical (BGC) floats have the potential to revolutionize the collection of data needed to monitor the ocean carbon cycle; inform and validate fisheries and ecosystem models; and monitor ongoing climate changes such as global ocean warming, acidification, eutrophication, and deoxygenation. Recognizing these strengths, US and international funding agencies have increasingly supported the develop of a BGC float array including a recent US commitment to deploy 500 floats globally (representing half of the anticipated international array) as part of the GO-BGC project. However, there remain opportunities to meaningfully improve the quality and quantity of information coming from the growing array. To this end, NOAA currently funds carbon data synthesis activities, a portion of which have provided calibration and validation data products for BGC float efforts. However, with the expanding array, there is new urgency in pushing forward this development in service of the myriad biogeochemical measurements that will be made globally. We propose new and expanded data product creation and data synthesis activities that will help more quickly and completely realize the potential of the BGC float observing strategy and related efforts. Specifically, we propose to create the following products:

1. *Empirical seawater property estimation routines:* This effort will be aimed at developing and validating empirical seawater property estimation routines with a focus on properties that are not part of the marine carbonate system, improving the quantity and quality of information that can be obtained from observation efforts like GO-BGC.
2. *A homogenized seawater pH data and metadata product:* Research efforts have recognized the difficulty of creating internally consistent pH data products due to the variety of measurement strategies used for pH and the varying completeness of metadata. This effort collate cruise metadata, propose recommended adjustments for cruise pH, and release a companion product with these adjustments applied.
3. *An estimation algorithm-based data Quality Control assistance tool:* Disagreements between measured and estimated values can help call attention to anomalous measurements that warrant further attention, whether because the measurements are problematic or because they are unexpected and interesting. We will create a tool that will make these comparisons fast and easy, and that will put the results in the context of expectations from statistics. Using this tool when ingesting data into data products will improve the quality of the products and the efficiency of data quality control efforts.

Broader Impacts and Relevance: This research is directly relevant to the COM and CVP call. It aims to develop new observation-based *in situ* ocean synthesis data products including updated algorithms for nutrients, dissolved organic carbon, and oxygen; a homogenized global pH data and metadata product; and a semi-automated quality-control assistant tool. The products will aid climate monitoring, both by improving the quality of data from the Global BGC-Argo effort and by increasing the information that can be produced by this and other observing efforts from a given amount of funding. The proposed funds will primarily support the work and professional development of a postdoc and to a lesser degree the project management and postdoctoral mentorship by PI Carter. Postdoctoral advisory efforts will be provided by co-PIs Feely and Fassbender.

***Gridded data products with uncertainties for 21st century in-situ oceanographic observations

Donata Giglio (University of Colorado Boulder), Mikael Kuusela (Carnegie Mellon University)

Developing and sustaining the global climate observing system is key to advancing our understanding of climate variability and change. E↵orts in the past few decades have started a data revolution, providing oceanographic observations with unprecedented coverage in space and time. Yet leveraging available observations for climate studies and to analysemodel biases relies on constructing gridded oceanic fields along with quantified uncertainties from sparse data, and there are fundamental challenges associated with doing this. These challenges stem from various factors, including (1) data quality control and bias correction, (2) sparse distribution of data in space and time, (3) spatial and temporal correlation of the observations, and (4) the choice of statistical model used to estimate gridded fields and their uncertainties. As a result, while several observation-based gridded products are available for oceanic fields, they currently lack a robust uncertainty quantification (UQ), i.e., they lack UQ estimates that (1) provide information on spatial and temporal correlation of mapped fields, along with correlation across variables (e.g., across temperature and salinity); (2) rely on realistic decorrelation scales (e.g., obtained using data-driven decorrelation scales estimated from the observations); (3) have undergone rigorous validation before being adopted. Since accurate uncertainties depend on the details of the gridding method, we need to produce a new gridded product that enables us to obtain robust uncertainties.

This project will **develop state-of-the-art mapping methods**, effectively addressing challenges described above, and **deliver gridded data products with robust uncertainties** for global ocean temperature and salinity (T/S) in the upper 2000 m, mixed layer depth and properties, ocean heat content, and steric sea level. We will produce both monthly and 10-day fields since 2004 (1 x 1 latitude/longitude grid, 66 vertical levels) and include Argo profiles, ship-based observations and glider data. Our proposed method features the option of jointly considering more than one variable at a time, which is beneficial for mapping and UQ of T/S and essential for leveraging, in the future, larger observational networks like core Argo in mapping sparser observations, e.g. BioGeoChemical (BGC) Argo. Users will be able to both **import our products directly into their programming environment** of choice (via API) and **visualize fields on the web app Argovis**, developed in Dr. Giglio's group. The entire code pipeline from raw data to final gridded product with UQ **will be made public** on GitHub, enabling reproducibility and extendability by the community.

This project will increase the usefulness of NOAA observations to the broader scientific community by (1) producing gridded fields that can be leveraged for model evaluation, validation, and process-oriented diagnostics, as well as for climate studies, (2) developing methods to produce robust UQ (e.g., for estimates of ocean heat content variability), and (3) setting the base for leveraging larger observational networks when mapping sparser data, which will be a key technical requirement for mapping future BGC Argo observations. Our codebase will serve as a platform for future community e↵orts on developing gridded ocean data products. As an example, our codebase could be adopted by operational centers that are interested in producing gridded fields with robust uncertainties.

***Innovative analysis of deep and abyssal temperatures from bottom-moored instruments

R. C. Perez, C. S. Meinen, S. Dong (NOAA/AOML/PHOD); Senior personnel: D. Volkov, M. Le Hénaff

The ocean plays a major role in the global redistribution of heat within the climate system, however direct observations of temperature variability in the deep ocean (i.e., below 2000 dbar) are very sparse in the present Global Ocean Observing System. A recent study (Meinen et al., 2020) led by scientists at NOAA/AOML, which garnered interest in mainstream media including Smithsonian Magazine, The Guardian, and more, has demonstrated that an innovative use of an internal temperature sensor located within the sphere of a pressure-equipped inverted echo sounder (PIES) mooring can provide high-temporal-resolution (hourly) near-bottom temperature measurements over extended time periods. Two long-term (10-15+ year) arrays of PIES moorings have been maintained in the North and South Atlantic by NOAA and international partners. These long-term records represent ideal data sets to both characterize deep temperature variations with the existing observing system and to evaluate the realism of deep temperature variability in the present generation of numerical ocean models. Furthermore, these data can be used to quantify deep ocean warming in a manner never before accomplished on a broad scale, and can be used to investigate the mechanisms involved in the observed deep/abyssal ocean temperature variability, as well as their implications. For this proposal we will primarily focus on bottom temperature observations collected from the 34.5°S and 26.5°N trans-basin arrays, however as time permits, additional existing records at other latitudes within the Atlantic sector will also be obtained and analyzed. This innovative analysis will aid in quantifying observational uncertainties in the existing long-term observations of deep ocean temperature, as well as help inform decisions about where future observations of temperature are needed within this poorly observed portion of the ocean. The deep temperature records will also serve as an example/reference data set for interpreting future deep temperature variations as new deep observational platforms are brought online (e.g., global deep Argo). As a second component of this proposed work, outputs from a series of different numerical model simulations, state estimates, and operational analyses (e.g., GOFS3.1, OFES, ECCO2, and GFDL simulations) will be examined. The highly-temporally-resolved temperature records from the PIES represent a novel benchmark against which to evaluate the skill of the deep and abyssal temperature evolution in these models. We will use this data to assess the fidelity of the model mean temperatures, their variability, and their trends. Most importantly, once validated, these models will provide the necessary tools to determine the mechanisms driving deep temperature variability at key locations in the Atlantic Ocean.

This proposal is relevant to the *Multi-program (COM-CVP-GOMO) call competition 5* as the deep ocean temperature observations/analysis involved will help to: (a) develop a new observation-based product for quantification of observational uncertainty; (b) increase the use of NOAA's historical field campaign data/observations; (c) enable and/or improve future climate model evaluation; and (d) improve/point toward improvements in current methods for ocean observation. The proposed work is also consistent with *NOAA's Strategic Goal: Climate Adaptation and Mitigation - Improved scientific understanding of the changing climate system and its impacts, and OAR's Strategic Goal: Detect Changes in the Ocean and Atmosphere: Identify and address gaps in observation requirements needed to understand causes of variability and change.*

***MarineFlux: A user friendly in-situ marine turbulent flux data service

Shawn Smith, Mark Bourassa (Florida State University, Tallahassee)

Our project, *MarineFlux: A user friendly in-situ marine turbulent flux data service*, will establish a user-friendly web access point for the dissemination of turbulent air-sea fluxes and the essential ocean and climate variables (EOV/ECV) required to estimate the fluxes from surface in-situ platforms. The service will include fluxes derived from the International Comprehensive Ocean-Atmosphere Data Set (ICOADS), the Shipboard Automated Meteorological and Oceanographic Systems (SAMOS) initiative, and select flux datasets from NOAA expeditionary cruises. The proposal targets NOAA's *Innovative Ocean Dataset/Product Analysis and Development for support of the NOAA Observing and Climate Modeling Communities*. Air-sea fluxes, recognized as an EOV, are a key component of the Earth's energy and water cycles, impacting ocean and atmospheric weather on scales from pop up storms to interannual variability and long-term climate. Flux differences (gradients) on relatively small spatial scales can cause relatively strong and rapid coupling between the ocean and atmosphere, yet these scales (<25 km) are not well resolved by the buoy network or by satellites. Observations from ships and drones do make observations on these scales, but the data sets are relatively difficult to work with because of heterogeneous data formatting and the observations for flux-related bulk variables being taken at a wide range of heights. The older ship data also has substantial periods of systematic errors that need to be identified and, if possible, corrected. This situation is a powerful barrier to the use of these data in developing the next generation of coupled ocean/atmosphere models that are expected to support weather and climate forecasts. Our experience with data management and with observations from ships and research vessels has given us the knowledge to address these issues and to provide users with easily accessible fluxes and flux related EOV/ECVs from a wide range of NOAA and other marine platforms.

We proposed developing a web-based service that will serve both EOV/ECVs and turbulent fluxes from multiple marine data sets. Working with the international flux community we will ensure the service provides not only the observations, but the metadata required to support discovery, search, access, and application for a wide range of users. The service will provide both a web user interface, for human search and discovery, and RESTful web services to support machine-to-machine queries. Comparisons between fluxes from different observing systems on a single platform and between nearby platforms will allow development and distribution of uncertainties for the fluxes. Finally, we will engage the flux user community both within NOAA and across the international community in the development and testing of the proposed service.

The proposed activity focus directly on the program priority to increase the usefulness of NOAA marine weather and surface ocean observations specifically for air-sea fluxes. Applying uniform quality control and systematic error correction will result in a set of observations that will be provided at standard observing heights, facilitating their application to time series and spatial analyses. Comparing the expeditionary fluxes, both derived from SAMOS observations and directly measured on select NOAA cruises, and those estimated from operational weather observing platforms, will allow better quantification of the uncertainty in the operational flux estimates and systematic errors related to either the flux calculation or the direct flux observations (perhaps due to airflow distortion by the platform). Provision of both standardized input EOV/ECVs and fluxes across these data sources will enable product developers, modelers, and those exploring new satellite technologies to simply use these fluxes and not be burdened with the complexities of underlying inhomogeneities in operational and expeditionary datasets. Providing these data addresses NOAA's mission of "Science, Service, and Stewardship" by supporting development of NOAA data products and models which, in turn, should result in better forecasts on time scales that are important to weather-sensitive components of our economy (e.g., agriculture, fisheries, shipping, energy production) and the general public.

Uncertainty quantification, data quality, and observing network design in an ocean state of the Tropical Pacific

Patrick Heimbach (The University of Texas at Austin), Matthew Mazloff (University of California San Diego), Aneesh Subramanian (University of Colorado Boulder)

Problem Statement & Summary of Work: This proposal addresses NOAA-OAR-CPO-2021-2006389, Competition #5 "*Innovative Ocean Dataset/Product Analysis and Development for support of the NOAA Observing and Climate Modeling Communities*". Our focus is on assessing observational strategies for the Tropical Pacific Observing System (TPOS) in the context of data assimilation for reanalysis and model initialization. Core elements of the TPOS mission include observing ENSO, advancing scientific understanding of ENSO variability, and identifying effective observational strategies for skillful weather and climate prediction on subseasonal to seasonal (S2S) timescales. To support TPOS, we propose a fundamentally dynamics-based approach to observing system design within a dynamically and kinematically consistent adjoint-based (4DVar) data assimilation framework. Our research proceeds along three main thrusts: (i) We will leverage a now-existing medium-resolution (1/3°) Tropical Pacific Ocean State Estimate (TPOSE) which will be augmented to a higher-resolution, 1/6° product (TPOSE_HR). This state estimate will improve the representation of currents and Tropical Instability Waves (TIWs) and provide a baseline for three study periods, covering an Eastern Pacific El Niño (2015/16), a Central Pacific El Niño or Modoki (2009/10), and an Eastern Pacific La Niña (2010/11). These baseline TPOSE_HRs will enable us to address a number of questions regarding the impact and best practices of assimilating emerging TPOS sensors. Because our adjoint-based framework provides sensitivities to the atmospheric state, we will be able to assess how existing or hypothetical sensors will constrain air-sea fluxes. (ii) In a second thrust, we will conduct dynamical attribution studies for a range of quantities of interest (QoIs) by means of a Green's functions approach, with the detailed sensitivity kernels provided by the adjoint. The choice of QoIs is guided by TPOS2020 goals, in particular with respect to ENSO monitoring, understanding, and prediction. Recognizing that ENSO diversity is governed by a complex interplay of dynamical processes and that TPOS serves to inform a diverse range of weather and climate phenomena, we will assess the degree to which sensitivity patterns for diverse QoIs to be tested exhibit redundancies. Sensitivity maps, interpreted as "heat maps" quantify variables and regions that impact the QoI as a function of lag time. These maps provide valuable information regarding predictable anomaly propagation and instrument placement. (iii) In a third thrust, use of derivative information will be further refined by quantifying how existing or hypothetical observational assets may reduce uncertainties in the QoIs chosen under (ii). This will be achieved within the context of Hessian-based uncertainty quantification (UQ) and optimal observing network design. The ability to conduct such calculations sets our framework apart from conventional studies based upon observing system [simulation] experiment (OS[S]E) approaches.

Broader Impact & Relevance to the Competition: Our work addresses three interrelated themes of the competition: (1) It will produce a new high-resolution dynamically consistent ocean synthesis dataset, which will synthesize all heterogeneous satellite and in-situ data streams, and with extended use of TPOS new mooring data to be assimilated. (2) Dynamical attribution of origins of variability of a range of metrics will advance our understanding on the dynamical processes that may give rise to predictability, on the role of air-sea flux patterns, and on the discrepancies between observed and simulated ocean products. (3) Optimal observing network design within the framework of Hessian UQ provides a unique perspective on observation complementarity vs. redundancy and frames observing placement strategies within a dynamical context.