2020 Competition: COM - **Developing terrestrial-, marine-, and ice-atmosphere boundary layer datasets through collaborations between observation and modeling communities**

Competition Number: 2808170

Funding Opportunity Number: NOAA-OAR-CPO-2020-2006076

In FY20, the COM program is soliciting proposals that develop, or significantly improve upon, oceanic, atmospheric, sea-ice, and/or terrestrial, physical or biogeochemical datasets from existing observations that will advance modeling efforts of the terrestrial-, marine-, or sea ice-atmosphere boundary layer (BL). Significant observations for monitoring and process understanding, as well as model intercomparison project efforts, and dataset assessment, have been undertaken by international organizations and U.S. government agencies to advance understanding of BL processes on multiple scales. NOAA has invested significantly in a wide variety of ocean-, and land-atmosphere observations and process studies in areas or conditions previously understudied. These observations provide opportunities to complement on-going work and address BL knowledge gaps (e.g., analysis of time-series of multi-platform observations for climatological features, blending and repackaging of disparate datasets for model use) through targeted dataset development and analysis for climate model improvement. To advance NOAA's climate prediction capacity in priority climate risk areas through improved BL understanding and representation in models, projects should promote collaboration and coordination between the observation and modeling communities and include at least one named collaborator with modeling expertise.

Projects should focus on the development and analysis of high-quality, climate-relevant datasets that will be readily utilized by the earth system/climate modeling community for model development, assessment, and/or performance evaluation. Proposed dataset development should have high potential to meet the critical need of reducing error and minimizing uncertainty of boundary layer processes in models, and therefore advance our foundational understanding of the earth system, and improve our ability to detect and forecast a changing climate for societal application. Priority areas for dataset development and analysis focus on key challenges related to integrating observations with models, and include those that have persisted over the last decade and were highlighted as recent opportunities for BL observation-model integration. Proposed dataset development should focus on at least one of the priority areas:

- Explore and apply state-of-art techniques to upscale surface-based network data and

point data for model use.
- Integrate observations from different platforms for model use (e.g., in-situ, satellite).
- Enable implementation of new metrics, such as process-based model diagnostics

Projects should be developed and implemented in consultation with the modeling community over the lifecycle of the project. Proposals should explicitly identify an existing gap that the dataset will address for the climate modeling community (e.g. What will the new dataset enable modelers to do that they could not do before? What new types of information will be learned from dataset analysis and use of the dataset in models?). Proposals should identify research stakeholders and users in the modeling community and discuss how these stakeholders/users will be involved in the process by which the data products will be developed, shared, accessed, and readily utilized.

COM encourages projects to capitalize on the large volume of existing NOAA observations. We recommend collaborations across NOAA cooperative institutes, labs, and the broader external community.

Please see the information sheet for a detailed description of specific requirements.

NOAA-OAR-CPO-2020-2006076

Combining data and models for robust estimates of global air-sea CO2 fluxes

Galen A. McKinley, Tian Zheng (Columbia University) 3

Developing an irrigation dataset for assessment of anthropogenic impacts on terrestrial-atmosphere energy-water coupling using machine learning-based data fusion

Ruijie Zeng, Tianfang Xu (Arizona State University), Guo-Yue Niu (University of Arizona) 4

Development of an integrated ocean-atmosphere-sea ice dataset for validation of modeled Arctic boundary layer processes over an annual cycle

Taniel Uttal, Amy Solomon, Leslie Hartten, Matthew Shupe, P. Ola G. Persson (NOAA/PSL) 5

Employing a NOAA climate measurement network to create products for evaluation and improvement of boundary layer height (BLH) and cloud interactions in climate models

Laura Riihimaki, Joseph Sedlar (CIRES/University of Colorado Boulder), Joseph A. Santanello Jr. (NASA-GSFC), Kathleen Lantz (CIRES/University of Colorado Boulder) 6

Flux Products, long time-series, and diurnal cycle metrics from OceanSITES moored buoys as a baseline for climate monitoring and assessing model and satellite resolved air-sea interaction
Meghan Cronin (NOAA/PMEL), Dongxiao Zhang, Kevin O'Brien, Kenneth Connel (University of Washington/NOAA/PMEL/CICOES) 7

Improvements to the NOAA Optimal Interpolation analysis of sea surface temperature
James A. Carton (University of Maryland), Thomas M. Smith, Boyin Huang (NOAA/NESDIS), Travis Sluka (Joint Center for Satellite Data Assimilation) Huai-Min Zhang (NOAA/NESDIS), Andy Harris, Eileen Maturi (NOAA/STAR) 7

Modern Satellite/In Situ-Based Air-Sea Turbulent Heat Flux Products for the Arctic Ocean and Adjacent Seas
Lisan Yu, Robert Weller (Woods Hole Oceanographic Institution) 8

Combining data and models for robust estimates of global air-sea CO2 fluxes

Galen A. McKinley, Tian Zheng (Columbia University)

The ocean significantly modulates atmospheric CO_2, having absorbed the equivalent of 38% of industrial-age fossil carbon emissions. The evolution of climate in the coming decades to centuries is strongly dependent on how ocean carbon uptake changes in response to growing atmospheric CO_2 and the associated climate change. If earth system models are to make accurate predictions of future climate, it is critical that the mechanisms of variability and change in air-sea carbon exchange be accurately represented. Providing useful data-based constraints with which to test and improve models is challenging because surface ocean carbon data are sparse. Nonetheless, multiple methodologies have been proposed that can fill the gaps in these sparse data, and thus create estimates of global, spatially-resolved carbon fluxes at monthly timescales. However, there has been insufficient quantification of the errors and uncertainties in these gap-filled products, a situation that substantially reduces their utility to models. Our proposed project will apply an earth system model testbed to assess uncertainty in existing gap-filled products, and then develop improved products with spatially-resolved uncertainty. Our work will be guided by two central hypotheses:

- ***Hypothesis 1: Existing data-based pCO_2 reconstructions have low mean bias, but overestimate the amplitude of CO_2 flux variability, particularly where data are most sparse.*** We will use a suite of earth system model Large Ensembles as a testbed to evaluate the uncertainties in two leading gap-filling approaches. Through this work, we will develop a comprehensive understanding of the degree to which state-of-the-art interpolation approaches can robustly compensate for the sparsity of surface ocean carbon data.

- ***Hypothesis 2: Careful tuning of data science methods to this application can improve fidelity and provide spatially-resolved uncertainty.*** We will evaluate a suite of state-of-the-art data science methods so as to develop improved pCO2 and CO2 flux reconstructions that include spatially-resolved estimates of uncertainty.

This work is of value to NOAA's long-term research goals because it ***advances scientific understanding, monitoring, and prediction of climate and its impacts, to enable effective decisions.*** We respond to this solicitation by developing ***biogeochemical datasets from existing observations that will advance modeling efforts of the marine-atmosphere boundary layer (BL).*** The existing observations we will use are largely from NOAA. Thus, we will enhance the value of NOAA observations; this is evidenced in letters from NOAA PMEL and AOML. The lead PI has a long history of model development and analysis, and also works extensively with surface ocean data. Our product will be ***high-quality, climate-relevant datasets that will be readily utilized by the earth system/climate modeling community for model development, assessment, and/or performance evaluation.*** This utility is evidenced by a letter from NOAA GFDL. The work we propose will evaluate uncertainty in existing surface ocean carbon datasets and develop improved datasets, and thus has ***high potential to meet the critical need of reducing error and minimizing uncertainty.*** As required, we will both ***explore and apply state-of-art techniques to upscale surface-based network data and point data for model use*** and ***integrate observations from different platforms for model use (e.g., in-situ, satellite).*** Broader impacts of this work will be improved climate predictions. Further, we will train a postdoc in oceanography and data science, thus strengthening the US STEM workforce from which NOAA will draw in the future.

Developing an irrigation dataset for assessment of anthropogenic impacts on terrestrial-atmosphere energy-water coupling using machine learning-based data fusion

Ruijie Zeng, Tianfang Xu (Arizona State University), Guo-Yue Niu (University of Arizona)

Problem Statement

Current mechanistic understanding and prediction capability of the impacts of agricultural irrigation on terrestrial hydrologic cycle and land-atmosphere feedbacks have been limited. In the High Plains where irrigation is extensive and intensive, the Weather Research and Forecast (WRF) model coupled with land surface models (LSMs) generally produces significant dry bias in growing-season precipitation. We hypothesize that the dry model biases (about 200mm underestimation in precipitation) are mainly caused by the lack of representation of spatiotemporal agricultural irrigation processes in these models. Therefore, development of a tailored, value-added irrigation dataset has high potential to fill in the critical knowledge gap and enhance the representation of anthropogenic processes in climate and weather models.

Objectives

The goal of this study is to develop a quality-controlled benchmark dataset of irrigation location, timing and amount for the High Plain region by fusing multi-source, multi-resolution data to improve model representation of anthropogenic irrigation and thus enhance prediction capacity of regional atmosphere-terrestrial feedbacks. Our specific objectives are: 1) Blending multi-source and disparate in-situ, remote sensing and reanalysis datasets using machine learning-based data fusion to develop an irrigation dataset (including irrigation location, timing and amount) for the High Plains; 2) Applying diagnostic physical constraints to reduce the uncertainty of the fused irrigation dataset; 3) Incorporating the irrigation data into the Noah-MP LSM to investigate irrigation impacts on terrestrial water storage change, gross primary productivity and evapotranspiration as a use case of the developed irrigation data by the modelling community to diagnose gaps in anthropogenic process representation.

Broader Impacts
The irrigation dataset developed in this proposed research does not only target at resolving the precipitation biases in the High Plains due to the missing anthropogenic processes, but also serves as diagnostic benchmark to improve the representation of irrigation in climate/weather modelling and to guide future agricultural practice for improving irrigation efficiency. The project team will work closely with the modelling community to ensure the developed irrigation dataset will be readily accessible and usable by the modelling community.

Relevance to the COM Competition and NOAA's Research Goals
The goal of the proposed study closely fits into the priority area of the Climate Observations and Monitoring Program: "*Integrate observations from different platforms for model use (e.g., insitu, satellite)*". The benchmark irrigation dataset developed in this study will bridge the gaps between observation and modeling community in understanding anthropogenic interferences on automorphic boundary layer. The proposed study will advance our understanding and prediction capability of the Earth's climate system and is thus aligned with NOAA's long-term climate research goals.

Development of an integrated ocean-atmosphere-sea ice dataset for validation of modeled Arctic boundary layer processes over an annual cycle

Taniel Uttal, Amy Solomon, Leslie Hartten, Matthew Shupe, P. Ola G. Persson (NOAA/PSL)

Leveraging the recent international field campaign to the Arctic (MOSAiC), this project will develop a dataset that characterizes the Arctic boundary layer, surface state and energy budget, and clouds over an annual cycle. Modelers will guide the development of the observational dataset to facilitate future studies of boundary layer processes and development of metrics for evaluating these processes in models. The resulting dataset will have the potential to be used in improved modeling of the stable Arctic boundary layer and cloud feedbacks in the boundary layer.

Employing a NOAA climate measurement network to create products for evaluation and improvement of boundary layer height (BLH) and cloud interactions in climate models

Laura Riihimaki, Joseph Sedlar (CIRES/University of Colorado Boulder), Joseph A. Santanello Jr. (NASA-GSFC), Kathleen Lantz (CIRES/University of Colorado Boulder)

Abstract: Interactions between the land surface, the atmosphere, and clouds are poorly understood across the atmospheric boundary layer (BL), typically defined as the layer of the atmosphere between the surface and 2 km. A wide array of processes and feedbacks operate across a vast range of spatial and temporal scales within the BL. Difficulty in improving fundamental understanding of BL processes and feedbacks is often traced back to a lack of observing systems at the necessary spatial coverage and temporal frequency. Without sufficient observations, the physical parameterizations, typically derived from observational relationships, required to model these subgrid scale processes in global climate models (GCMs) lack the sophistication needed to be universally applicable. As a result, GCM hindcasts or climate projections often suffer from inadequate treatment of BL and cloud interactions and feedbacks.

The proposed work will develop three new data products of boundary layer height (BLH), cloud type classification, and a multivariate statistical summary product including BLH, cloud regime, surface radiation, and other meteorological parameters useful for evaluating GCMs at all stations in the existing NOAA CPO-funded SURFRAD network. These data products will fill an observational gap quantifying how clouds impact surface radiation and the growth of the convective mixed layer during the day in multiple climate regimes. Specifically we seek to build data products that address the question: ***How do certain cloud types (e.g., low stratiform, low cumulus, high cirrus) drive, or feedback onto, the BLH evolution?***

This proposal is specifically aimed to meet the COM proposal call to “develop terrestrial-atmosphere boundary layer datasets” using existing observational assets in the NOAA Earth System Research Laboratory Global Monitoring Division (ESRL GMD) SURFRAD network. These new BLH retrievals will be combined with cloud, radiation, and surface information to “enable implementation of new metrics, such as process-based model diagnostics”. The data products will be developed by those with observational expertise in BL, clouds, and radiation measurements and retrievals, in collaboration with climate modelers at NOAA GFDL, from the modeling and land-atmosphere diagnostic communities from NASA and DOE, and those actively involved with developing GEWEX Local Land-Atmosphere Coupling (LoCo) diagnostics. This project sits squarely in the mission space of the NOAA OAR COM program, as we will develop these new value-added products from the “existing observational assets” of the NOAA GMD SURFRAD network and will make all products publicly and freely accessible through that network’s website. The development of data products able to assess GCM process representation of the interactions between radiation, clouds, and BL will further NOAA OAR CPO’s mission of advancing “scientific understanding...and prediction of climate and its impacts”. With better representations of BL processes, future GCM predictions will allow better understanding of the impacts of community decisions on economic and ecosystem stewardship.

Flux Products, long time-series, and diurnal cycle metrics from OceanSITES moored buoys as a baseline for climate monitoring and assessing model and satellite resolved air-sea interaction

Meghan Cronin (NOAA/PMEL), Dongxiao Zhang, Kevin O'Brien, Kenneth Connel (University of Washington/NOAA/PMEL/CICOES)

Leveraging data and observations from OceanSITES moored buoys, this project will create four new product webpages to make existing data more accessible and discoverable. Webpages will highlight air-sea flux data, concatenated long time-series, deep ocean temperature and salinity data, and diurnal cycle metrics. In addition to a suite of monitoring applications, results of the project will be valuable and relevant to NOAA's routine evaluation of model-based analyses products, such as the NCEP/NCAR Climate Forecast System Reanalysis.

Improvements to the NOAA Optimal Interpolation analysis of sea surface temperature

James A. Carton (University of Maryland), Thomas M. Smith, Boyin Huang (NOAA/NESDIS), Travis Sluka (Joint Center for Satellite Data Assimilation) Huai-Min Zhang (NOAA/NESDIS), Andy Harris, Eileen Maturi (NOAA/STAR)

The gridded Daily Optimal Interpolation SST product (DOISSTv2) 1981-pres is a key tool in NOAA's response to the National Research Council's 2004 call for calibrated climate data records of SST spanning multiple decades. For the recent 4th National Climate Assessment DOISSTv2 continues to contribute by providing a multi-decade long record of the health of US maritime regions. This proposal to the Climate Observations and Monitoring Program by a group of University and NOAA scientists will explore introduction of new data sets and analysis techniques to help the DOISST project to connect to related NOAA efforts and extend its role as an internationally regarded, calibrated record of global SST spanning multiple decades.

The primary data used to construct DOISSTv2 is the set of Level 3C infrared radiance measurements from AVHRR instruments carried on NOAA and EUMETSAT polar orbiter satellites. These satellite data sets are calibrated and bias-corrected by comparison to a variety of in-situ observations obtained from the ICOADS release 2.1 archive. Statistical objective analysis with prescribed spatial covariance functions is used to map SST onto a uniform 0.25°x0.25° grid at daily intervals. In the decade since DOISSTv2 was developed new satellite instruments have become available and older satellite data have been reprocessed. New and reprocessed in-situ observation sets are also available to improve bias-correction. Improved understanding of the nearsurface temperature structure of the ocean and new objective mapping techniques, including application of neural networks, offer additional opportunities to improve accuracy and reduce bias. This project to support modernization of DOISST will also align its software base with the multi-agency Joint Effort for Data Assimilation Integration (JEDI) framework, thus allowing DOISST to benefit from future JEDI developments. Specific tasks include:

1) Explore the impacts of inclusion of recalibrated and new infrared and microwave satellite data sets including VIIRS and GOES-R.

2) Explore the impacts of use of an updated and expanded in-situ observation set including Argo profiles.
3) Explore the impact of benefitting from an improved model of the upper ocean with explicit consideration of the diurnal cycle and nearsurface temperature stratification
4) Explore improved gridding techniques to improve error modeling, reduce bias, and improve background state estimation.
5) Carry out a concentrated effort to improve the accuracy of the SST analysis at high latitude

Broader Impacts, Relevance to the Competition, and to NOAA's climate research goals: SST is a key variable for monitoring variability and trends of marine biogeochemical properties, coastal upwelling, marine heat waves, and global climate. It provides the lower boundary condition for the atmospheric boundary layer and for weather forecasts including hurricane forecasts, and it provides the upper boundary condition for ocean forecasts. This proposal to improve the DOISST –type climate data records, thus, directly supports the *Climate Observations and Monitoring Program's* call to 'develop marine boundary layer datasets through collaborations between observation and modeling communities'. By exploring new data sets and methods to improve DOISST the proposal also directly supports a long-term mission of NOAA to provide well-documented, accurate and statistically consistent Climate Data Reference data sets to monitor the climate and health of the marine environment.

Modern Satellite/In Situ-Based Air-Sea Turbulent Heat Flux Products for the Arctic Ocean and Adjacent Seas

Lisan Yu, Robert Weller (Woods Hole Oceanographic Institution)

This proposal is in response to NOAA FY2020 COM - Developing terrestrial-, marine-, and ice-atmosphere boundary layer datasets through collaborations between observation and modeling communities. We propose to develop air-sea variable and turbulent heat flux products for the Arctic and adjacent subarctic seas (poleward 50N) by capitalizing on recent observing advancements made at high-latitude environments and the machine learning applications for remote sensing. The diminishing sea ice has radically changed the seasonal patterns of turbulent exchange between the Arctic atmosphere and ocean. In regions of sea-ice retreat, increased absorption leads to increased storage of solar heating in summer and increased upward surface turbulent latent and sensible heat fluxes in fall/winter. The heat released to the atmosphere has been suggested as a key mechanism for the Arctic amplification. However, both climate model simulations and atmospheric reanalyses show difficulty in representing the surface heat fluxes associated with the growth and melt of sea ice. The large disparity in the predicted rates of Arctic sea ice loss highlights the urgent need for developing accurate satellite-based air-sea turbulent heat flux products for the Arctic Ocean to serve as validation for the modeling community.

Surface turbulent heat fluxes on basin-to-global scales are commonly constructed from bulk flux parameterizations using air-sea variables (e.g. wind speed, sea and air temperatures, near-surface specific humidity) that can be observed or modeled. Three distinct features separate the construction of Arctic surface fluxes from that for low-to-mid latitudes. First, the sea-ice surface is a heterogeneous medium, where the surface drag is a combination of skin drag at the horizontal surfaces and form drag associated with the vertical surfaces of floes. Second, the atmospheric boundary layer over the summer sea ice is characterized by near neutral or weakly stable conditions. The bulk flux algorithms that work well for low-mid latitudes are inappropriate for the Arctic environment. Lastly, the effect of sea surface temperature on satellite microwave emissivity increases as the amount of water vapor decreases toward high latitudes. As a result, the retrieval models for near-surface thermodynamic variables change with latitudes, which requires special training using high-latitude observations. Yet, the lack of sufficient in situ observations in the Arctic and Adjacent seas has hampered these efforts until recently.

The proposed new heat flux products will be developed from an innovative platform that features (i) a collection of contemporary air-sea variables and/or surface fluxes observed by unmanned saildrones and surface mooring at high latitudes - this is used for training retrieval algorithms and validating end products; (ii) a deep artificial neural network framework for improving satellite retrievals of near-surface thermodynamic variables; (iii) an optimization framework that includes the synthesis of satellite scatterometers and passive microwave radiometers (including L-band microwave passive radiometers to improve high wind speed retrievals associated with synoptic weather events; and (iv) a bulk flux parameterization algorithm suitable for representing the air-sea exchange processes in the Arctic environment. This proposal brings together leading researchers with complementary expertise in satellite and in situ observations of air-sea fluxes, artificial neural network, atmospheric reanalysis, and climate modeling. The proposed activities directly address the two priority areas solicited in the competition: (i) explore and apply state-of-art techniques to upscale surface-based network data and point data for model use, and (ii) integrate observations from different platforms for model use (e.g. in-situ, satellite). The proposed study supports the NOAA's long-term climate goal on providing the essential and highest quality environmental information vital to our nation's safety, prosperity and resilience.