Blue Waters Professor Allocation Annual Report: September 25, 2019 – October 12, 2020

Title: Satellite remote sensing and 3D radiative transfer modeling for improved weather and climate predictions

PI: Larry Di Girolamo, Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign

Collaborator(s): Matias Carrasco Kind, Dongwei Fu, Yulan Hong, Jesse Loveridge, Arka Mitra, Don Petravick, Puja Roy, Yizhe Zhan, Guangyu Zhao, *University of Illinois at Urbana-Champaign*; and Souichiro Hioki, Yi Wang, Ping Yang, *Texas A&M University*

Corresponding Author: Larry Di Girolamo, gdi@illinois.edu

Executive Summary:

Our multi-pronged approach for tackling key problems in weather and climate research using Blue Waters, satellite observations and 3D radiative transfer has had an incredible year of success: (1) dissemination of research in 3 peer-reviewed publications (2) extensive progress on the Terra Data Fusion project with continued support from NASA and Illinois; (3) continued Terra data analysis for meteorological studies in SE Asia as part of NASA's Cloud and Aerosol Monsoonal Processes Philippines Experiment (CAMP2Ex) field campaign; (4) continued enhancements to our understanding of global microphysical properties of water and ice clouds; and (5) continued development of training data from Terra to support ML (CNN) cloud detection algorithm development for NASA's MAIA mission. Our initial request of 200K NH was on target, but was significantly underutilized owing to the decommissioning of Nearline storage and the significant interruptions caused by the COVID-19 pandemic. The unused NH will be used by the end of 2020 for LES and 3D radiative transfer simulations in support of CAMP2Ex.

Description of Research Activities and Results:

Research in weather and climate has massive societal benefits, and indeed has been one of the leading drivers for advancing supercomputing infrastructures. One of least understood and most important aspect of the weather and climate system are Earth's clouds. Clouds cover about 70% of our planet. They are one of the most interconnected components of the Earth System, playing a key role in the Earth's hydrological cycle, regulating the incident solar radiation field more than any other atmospheric variable, and acting as the most important greenhouse constituent in our atmosphere. As such, they modulate a wide range of physical, chemical, and biological processes on Earth. The Intergovernmental Panel on Climate Change (IPCC) affirms that the role of clouds remains the leading source of uncertainty in anthropogenic climate change predictions. In addition, the role of cloud microphysics and cloud-radiation interactions in the timing and intensity of weather events remains an active area of research.

To make headway in reducing uncertainty in weather and climate predictions, the World Meteorological Organization and the IPCC defined a list of Essential Climate Variables (ECVs) requiring global satellite observations (http://www.wmo.int/pages/prog/gcos). It has been established that ~2/3 of the ECVs derived from satellite do not meet accuracy requirements, therefore calling for improvements in the algorithms and technologies used by satellites. For improving algorithms, one of the key recommendations from the NRC 2007 Decadal Survey on Earth Science and Applications from Space (NRC 2007) is clear: *"… experts should… focus on providing comprehensive data sets that* *combine measurements from multiple sensors."* This, in part, targets NASA's flagship of the Earth Observing System called Terra. Terra was launched in 1999 and continues to collect data for Earth sciences using five instruments: the Moderate-resolution Imaging Spectroradiometer (MODIS), the Multi-angle Imaging SpectroRadiometer (MISR), the Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER), the Clouds and Earth's Radiant Energy System (CERES), and the Measurements of Pollution in the Troposphere (MOPITT). Terra data is amongst the most popular NASA datasets, serving not only the scientific community, but also governmental, commercial, and educational communities.

While the need for data fusion and the ability for scientists to perform large-scale analytics with long records have never been greater, the challenge is particularly acute for Terra, given its growing data volume (>1 petabyte), the storage of different instrument data at different NASA centers, the different data file formats and projection, and inadequate cyberinfrastructure. We initiated the Terra Data Fusion (TDF) Project, supported under NASA Grant Number NNX16AM07A, to tackle two long-standing problems: 1) How do we efficiently generate and deliver Terra data fusion products; 2) How do we facilitate the use of Terra data fusion products by the community in generating new products and knowledge through national computing facilities, and disseminate these new products and knowledge through national data sharing services? Blue Waters provides the computational resources needed, in part, to solve these problems. Earlier this year, all project deliverables were met and made public, as described on the TDF project page:

https://digirolamo.web.illinois.edu/projects/terra-fusion/

Two scientific use-cases used to drive the TDF project were also completed in the past 12 months: (1) retrieving liquid water cloud drop size distribution, and (2) retrieving cirrus cloud ice crystal shapes. In the first use case, we fused the Terra MISR and MODIS data to characterize and correct the biases in cloud drop size datasets derived from MOIDS-only. The MODIS-only dataset is the primary satellite dataset used by the Intergovernmental Panel on Climate Change for providing observational constraint on cloud-aerosol interactions in climate change predictions. Our approach for building an improved product through MISR+MODIS fusion builds upon the approach we showed in Liang et al. (2015), which was the impetus for the TDF. We extended that study under NASA Grant Number NNX14AJ27G to examine the underlying causes of the biases. Ph.D. graduate student Dongwei Fu painted a new view of the global distribution of cloud drop sizes from Terra, which was published in Fu et al. (2019).

For the second use case, we are worked closely with Prof. Ping Yang at Texas A&M on a specialized MISR and MODIS fusion dataset designed for retrieving ice cloud microphysical properties. That work was supported under NASA Grant Number NNX15AQ25G. As in Fu et al. (2019) significant deviations from the MODIS-only ice cloud products were observed. We developed a new optimal surface roughness model for ice crystals that allowed for improved retrievals of ice cloud properties through MISR+MODIS fusion. We showed that the global distribution of ice crystal type depends very much on cloud regime. Results were published in Wang et al. (2019).

The fusion of different Terra instruments at scale required a very efficient method. We developed a new software tool, called pytaf, that resamples Earth observation data stored in swath, grid or point structures using a novel block indexing algorithm. This tool is specifically designed to process large scale datasets and extensively tested on Blue Waters as part of the TDF project. Results were published in Zhao et al. (2020).

With the onset of the COVID-19 pandemic, NASA provided us additional funding to extend the Terra fusion dataset to present day in order to examine the global impact of the human reaction COVID-19 as seen from the Terra record. Terra is our longest, most stable record of Earth from space, hence our most valuable in teasing out COVID-19 signals. The move of the Terra record off of Nearline and onto the storage condo proved to be a significant impediment to this research. We continue to work on how to best access and manage data on the condo.

Blue Waters, the Terra data, and other satellite datasets are also being used in analysis supporting the meteorological studies in SE Asia as part of the Cloud and Aerosol Monsoonal Processes Philippines Experiment (CAMP2Ex) field campaign, where we deployed NASA assets from the Philippines from August to October 2019. This research is being carried out under NASA contract 80NSSC18K0144. Our deployment was a great success, and was recognized by NASA with a Group Achievement Award. More about CAMP2Ex can be found here:

https://digirolamo.web.illinois.edu/projects/camp2ex/

We are in the process of setting up large scale LES simulations on Blue Waters to simulate the cloudy environment of CAMP2Ex to help in developing an understanding of aerosol-cloud interactions in the region. We are also using these simulations to run 3D radiative transfer simulations to simulate radiation at sensor in order to test next generation tomographic techniques for cloud and aerosol remote sensing, under NASA contract 80NSSC20K1633. These simulations are expected to use 104K NH from the remaining NH allocated for 2020.

Publications and Presentations Associated this this Work

Peer-Reviewed Journals and Dissertations

Fu, D., L. Di Girolamo, L. Liang, and G. Zhao, 2019: Regional Biases in Moderate Resolution Imaging Spectroradiometer (MODIS) marine liquid water cloud drop effective radius deduced through fusion with Multi-angle Imaging SpectroRadiometer (MISR). *J. Geophys. Res. Atmos., 124,* https://doi.org/10.1029/2019JD031063.

Wang, Y., P. Yang, S. Hioki, M.D. King, B.A. Baum, L. Di Girolamo, D. Fu, 2019: Ice cloud optical thickness, effective radius, and ice water path inferred from fused MISR and MODIS measurements based on a pixel-level optimal ice particle roughness model. *J. Geophys. Res. Atmos., 124,* https://doi.org/10.1029/2019JD030457.

Zhao, G., M. Yang, Y. Gao, Y. Zhan, H.-K. Lee, and L. Di Girolamo, 2020: PYTAF: a python tool for spatially resampling Earth observation data. *Earth Sci. Informatics,* https://doi.org/10.1007/s12145-020-00461-w.

Conference, meeting, and seminar presentations

Di Girolamo, L., et al., 2019: Terra Level 1 data fusion at scale: from production to applications. *American Geophysical Union Fall Meeting*, Dec. 8 - 13, San Francisco, CA.

Di Girolamo, L., et al., 2020: Improved cloud microphysical properties through MISR and MODIS fusion. *MISR Science Team Meeting*, February 12-13, Pasadena, CA.

Di Girolamo, L., et al., 2020: Terra Level 1 data fusion at scale: from production to applications. *MISR Science Team Meeting*, February 12-13, Pasadena, CA.

Levis, A., Y. Shechner, J. Loveridge, A. Davis, and L. Di Girolamo, 2020: Multi-view polarimetric scattering tomography and three-dimensional retrieval of cloud microphysics. *MISR Science Team Meeting*, February 12-13, Pasadena, CA.

Di Girolamo, L., 2020: NASA's Earth Observing System. Center for Astrophysical Survey, University of Illinois, February 27, Urbana, IL.

Wuebbles, D., A. Sharma, and L., Di Girolamo, 2020: COVID-19, Climate Change, and Health, National Center for Supercomputing Applications, University of Illinois, May 19, Urbana, IL.

Di Girolamo, L., 2020: COVID-19, Air Quality, and Climate, Illinois Foundation, University of Illinois, May 27, Urbana, IL.

Wuebbles, D., A. Sharma, and L., Di Girolamo, 2020: COVID-19, Climate Change, and Health, Illinois Climate Working Group, Illinois State Government, June 9, Urbana, IL.

Plan for Next Year:

Our work for the upcoming year extends much of the work described above and supported by NASA grants NNX16AM07A, 80NSSC20K1633, and 80NSSC18K0144. In addition, we have a new project with NASA (80NSSC20K0902) to examine deep convection updraft populations in organized and unorganized systems across the globe using, in part, the Terra dataset. For the upcoming year, we anticipate 70,000 node hours on Blue Waters for the processing and analysis of the Terra data over all these projects. For the LES and 3D radiative transfer simulations noted above for the CAMP2Ex project, we estimate 100,000 node hours based on tests on Blue Waters carried out over the past few weeks.

We therefore request 170,000 node hours for next year. We expect the usage break down by quarter to be the following:

Q1: 30% Q2: 30% Q3: 20% Q4: 20%

The storage requirement across projects is being handled externally at NCSA and the Department of Atmospheric Sciences. We can work within the current 50 TB allocation on Blue Waters.