



Center for Western Weather  
and Water Extremes  
SCRIPPS INSTITUTION OF OCEANOGRAPHY  
AT UC SAN DIEGO



SCRIPPS INSTITUTION OF  
OCEANOGRAPHY

UC San Diego

**Big Data and the Earth Sciences: Grand Challenges Workshop**  
**May 31 to June 2, 2017 Atkinson Hall Auditorium at Calit2/Qualcomm Institute, UCSD**

Sponsored by the Pacific Research Platform (PRP), Center for Western Weather and Water Extremes (CW3E) & Scripps Institution of Oceanography (SIO)

Grand Challenge Lectures:

- Dr. Larry Smarr, Calit2/University of California, San Diego
- Dr. Vipin Kumar, University of Minnesota
- Dr. Padhraic Smyth, University of California, Irvine
- Dr. Michael Wehner, Lawrence Berkeley National Laboratory

##### Workshop Day 1 #####

**Wednesday, May 31st:**

- 8:00am-8:45am      Registration**
- 8:45am-9:00am      Welcome, Scott L. Sellars, University of California, San Diego**
- 9:00am-9:50am      Grand Challenge Lecture: Larry Smarr, Calit2/University of California, San Diego**  
 Title: *“Using the Pacific Research Platform for Earth Sciences Big Data”*
- 9:50am-10:00am      Break**
- 10:00am-11:15am    Big data collaborations? Technologies for next generation big data transfers, storage, and modeling**
  - 10:00-10:20: Rachana Ananthakrishnan, Globus  
 Title: *“Globus for research data management”*
  - 10:20-10:40: John Graham, Qualcomm Institute

Title: *"PRP Multi-Institution, Hyper-Converged ScienceDMZ using Kubernetes"*

- 10:40-11:00: Ilkay Altintas, San Diego Supercomputer Center  
Title: *"Workflow-Driven Geoinformatics Applications and Training in the Big Data Era"*
- 11:00-11:20: V. Nandigam, San Diego Supercomputer Center  
Title: *"Distribution and processing of high resolution topography data via the OpenTopography portal"*
- 11:20-11:40: David Minor, Research Data Curation, University of California, San Diego  
Title: *"Research libraries and big data: the current landscape"*

**11:40am-1:00pm      Lunch and/or Tour The Big Wall, Calit2/UCSD**

**1:00pm-1:50pm      Grand Challenge Lecture: Michael Wehner, Lawrence Berkeley National Laboratory**

*Title: "Opportunities with very large high resolution climate model datasets."*

**1:50pm-2:00pm      Break**

**2:00pm-3:15pm      What is needed to advance our understanding in Earth Sciences?**

- 2:00-2:15: Douglas Schuster, National Center for Atmospheric Research  
Title: *"Accessing Multi-TB Sized Datasets at NCAR's Research Data Archive"*
- 2:15-2:30: Ross Towe, Lancaster University  
Title: *"The role of emergent technologies in understanding the natural environment and the complex web of interconnected ecosystems"*
- 2:30-2:45: Margaret Hurwitz, National Aeronautics and Space Administration & Science Systems and Applications, Inc.  
Title: *"The role of big data in applying NASA Earth science for societal benefit"*
- 2:45-3:00: Daniel Crichton, National Aeronautics and Space Administration  
Title: *"Enabling a Data-Driven Ecosystem for Earth Science Research at NASA and JPL"*
- 3:00-3:15: William S. Cleveland, Purdue University

Title: *"Divide & Recombine with DeltaRho for Bigger Data and Higher Computational Complexity"*

**3:15pm-3:30pm      Break - Coffee and snacks**

**3:30pm-5:00pm      Big Pixel Initiative Panel**

Chair, Gordon Hanson, School of Global Policy and Strategy, UCSD

- Ran Goldblatt, School of Global Policy and Strategy, UCSD
- David I. Kline, School of Global Policy and Strategy, UCSD
- Mehrdad Yazdani, Qualcomm Institute, UCSD
- Jen Burney, School of Global Policy and Strategy, UCSD
- Jessica Block, Qualcomm Institute, UCSD

##### Workshop Day 2 #####

**Thursday, June 1st**

**8:00am-9:00am      Registration/Breakfast**

**9:00am-9:50am      Grand Challenge Lecture: Vipin Kumar, University of Minnesota**

*Title: "Big Data in Climate: Opportunities and Challenges for Machine Learning"*

**9:50am-10:00am      Break**

**10:00am-11:15am      Data mining, machine learning & predictive modeling**

- 10:00-10:15: William W. Hsieh, University of British Columbia  
*Title: "Online learning with extreme learning machines, with applications to environmental sciences"*
- 10:15-10:30: Michael Bianco, Scripps Institution of Oceanography  
*Title: "Regularization of geophysical inversion using dictionary learning"*
- 10:30-10:45: David John Gagne, National Center for Atmospheric Research  
*Title: "Learning High-Level Representations of Spatial Atmospheric Data with Generative Adversarial Networks"*
- 10:45-11:00: Matthew Bowers, Purdue University,

Title: *“Deep Analysis of Large Complex Satellite-Based Cloud and Precipitation Properties”*

- 11:00-11:15: Ata Akbari Asanjan, University of California, Irvine  
Title: *“Short-Term Precipitation Forecasting using Recurrent Neural Network”*

**11:15am-1:00pm      Lunch**

**1:00pm - 2:15pm      Pattern Detection, Segmentation, and Object Recognition for Earth Sciences**

- 1:00-1:15: Stefan Liess, University of Minnesota  
Title: *“A Teleconnection between the West Siberian Plain and the ENSO Region”*
- 1:15-1:30: Qi Liu, Purdue University  
Title: *“Divide & Recombine (D&R) with DeltaRho Visualization and Modeling of TRMM Big Data”*
- 1:30-1:45: Haiqiang Niu, Scripps Institution of Oceanography  
Title: *“Underwater source localization using supervised machine learning”*
- 1:45-2:00: Santosh Nannuru, Scripps Institution of Oceanography  
Title: *“Multi-frequency sparse Bayesian learning with uncertainty models”*
- 2:00-2:15: Sahoo, Sasmita, Pennsylvania State University  
Title: *“Estimation of future groundwater stress in a changing climate using machine learning methods”*

**2:15am-3:00pm      Break - Coffee and snacks**

**3:00pm-4:30pm      Predictability Challenges in the Earth Sciences**

- 3:00-3:30: Jeanine Jones, California Department of Water Resources  
Title: *“Another Grand Challenge – Improving Sub-seasonal to Seasonal Precipitation Forecasting – Can Big Data Help?”*
- 3:30-3:45: Tiantian Yang, Deltares USA  
Title: *“Delft-FEWS: A Decision Making Platform to Merge Model, Data, and Optimization for Water Resources Management”*
- 3:45-4:00: Aneesh Subramanian, Scripps Institution of Oceanography  
Title: *“Exploring frontiers in seamless prediction: a journey towards exascale”*

*computing*”

- 4:00-4:15: Minghua Zheng, Scripps Institution of Oceanography  
Title: “*Scenario Identification and Forecast Verification of U.S. East Coast Winter Storms in Multi-Model Ensembles*”
- 4:15-4:30: Bruce Cornuelle, Scripps Institution of Oceanography  
Title: “*A less grand challenge: How can we merge machine learning with data assimilation?*”

**5:00pm-7:30pm      Reception, Martin Johnson House - Scripps Institution of Oceanography (Buses depart at 5:00pm outside Calit2)**

##### Workshop Day 3 #####

**Friday, June 2nd**

**8:00am-9:00am      Registration/Breakfast**

**9:00am-10:00      Structuring Unstructured Data in the Earth Sciences**

- 9:00-9:15: Peter Gerstoft, Scripps Institution of Oceanography  
Title: “*Using graph clustering to locate sources within a dense sensor array*”
- 9:15-9:30: Dimitri Kondrashov, University of California, Los Angeles  
Title: “*Data-adaptive Harmonic Decomposition and Real-time Prediction of Arctic Sea Ice Extent*”
- 9:30-9:45: Phu Nguyen, University of California, Irvine  
Title: “*UC-Irvine CHRS’s global satellite precipitation products and tools*”
- 9:45-10:00: Jamon Van Den Hoek, Oregon State University,  
Title: “*The Refugee Archipelago: A global assessment of the enviro-climatic marginality and geographic isolation of 922 refugee camps*”

**10:00am-10:30am      Break**

**10:30am-11:30pm      Grand Challenge Lecture: Padhraic Smyth, University of California, Irvine**  
Title: “*Machine Learning, Statistical Modeling, and Data-Driven Science*”

**11:30am-12:30pm Panel Discussion: Meeting the Challenge –paths forward for big data in Earth Sciences**

- **Padhraic Smyth, University of California, Irvine**
- **Vipin Kumar, University of Minnesota**
- **Larry Smarr, Caltech/University of California, San Diego**

**Adjourn**