White Paper Call for the Coordinated Open Watershed Science Workshop (due Tuesday Jan 15th)

- Workshop title and linked summary: Leveraging Distributed Research Networks to Understand Watershed Systems
- Workshop dates: Jan 28-30, 2019
- We encourage anyone interested in physical, biological, or chemical processes occurring throughout **any component** of watersheds to submit a white paper, regardless of affiliation, status, funding sources, background, etc.
- Submit your white paper as a Word document to the email addresses listed immediately below by January 15th, 2019.

  Copy and paste these addresses into your address bar:
  James.Stegen@pnnl.gov; Kelly.Wrighton@colostate.edu; elbrodie@lbl.gov; mbriggs@usgs.gov; jesus.gomezvelez@vanderbilt.edu; cvaradharajan@lbl.gov; David.Lesmes@science.doe.gov; Jessica.Moerman@science.doe.gov; Sujata.Emani@science.doe.gov
- Questions on this process should be addressed to Jessica.Moerman@science.doe.gov

A key outcome of this workshop will be the articulation of how the environmental science community can best use coordinated, open network science (a.k.a. 'distributed research') to fill knowledge and data gaps needed to understand and predict watershed function across the continental U.S. In this case, understanding is focused on a distillation of the underlying physical, chemical, and biological processes that that control system function across scales. The concept of watershed function can emphasize various factors ranging from ecosystem services associated with water/habitat quality and flood protection to global carbon/nutrient cycling.

A goal of the questions provided below is to consult with the science community regarding the aspects of watershed function that are most critical to understand and predict in the context of ongoing and future perturbations. An additional goal is to capture guidance on the primary challenges, solutions, and opportunities associated with deploying distributed research programs aimed at generating knowledge, data, and models needed to improve predictions of selected watershed functions.

Below you will find a set of questions intended to help structure your responses. These represent key considerations to help the workshop team understand the rationale for your interest in the specific topic of your white paper and provide a framework for synthesis across diverse topics. However, you do not need to respond to all questions or all parts of questions, they are meant to guide your white paper, not dictate the exact topics. These white papers will be included as appendix to the workshop report, with attribution to those that contribute.

It is expected that the white papers will be relatively short, targeting a maximum of 3 pages. Shorter white papers are also encouraged (i.e., it is fully acceptable to write only 1-2 focused paragraphs). As you read through the questions below, we encourage you to find aspects that inspire you and write to those topics with no concern about covering all topics, and to feel free to stray into spaces outside topics covered in the questions.

Guiding questions:

- **What aspect of watershed function** do you feel is most critical to understand and predict, in particular with respect to changes in that function due to ongoing and future perturbations?

- **Why** is that function so critical to predict?
    - If we had improved capacity to predict this function, what would the implications be to society and stewardship of the environment?
    - Who/what are the stakeholders, decision makers, aspects of society, etc. that would be impacted by and interested in the associated fundamental knowledge and predictive capacity?
    - What would they do with that knowledge and the associated predictions?

- What are the **essential processes** that must be understood to enable prediction of the selected aspect of watershed function?
    - At what spatial and temporal scales must these processes be understood?
    - Where within watersheds must we understand these processes (e.g., within hydrologic exchange zones, hill slopes, rooting zones, surface water, etc.)?
    - When in time must we understand these processes (e.g., during perturbation events, under steady-state conditions, within a particular season, etc.)?

- **How will mechanistic and/or data-driven models benefit** from new data, concepts, and/or mechanistic understanding of the processes you described above?
    - Are there existing computational codes that are well-suited for modeling/predicting key processes/scales you identified above, and if so, why? (e.g., do they integrate the necessary mechanisms, run at the right scales, have existing parameterizations, HPC compatible, open source, etc.)?
    - What new model developments are required, and why? What would their essential elements be (e.g., what would they predict, what disciplines would they be built from, what scales would they run at, etc.)?
    - What would the data-model integration strategy be (e.g., formal data assimilation methods, direct parameterization, etc.)?
    - Would the approach provide opportunities to repeatedly iterate between data acquisition and model refinement, and if so, how would that iterative (i.e., ModEx) approach be pursued?

- Are there **opportunities to use models *a priori*** to guide data generation?
    - How would these modeling efforts guide the type and scales of data generated and where/when those data would be generated (i.e., how would an iterative ModEx approach be implemented?)
    - What would be the approach to doing this (e.g., which models, what spatial and temporal domains, parameter and structural sensitivity analyses, etc.)?

- How would a broad, **distributed science approach** be used to generate the knowledge and data needed to **inform key processes and models**?
    - What are the logistical and technical challenges associated with generating data from across many different watershed systems (e.g., sensor deployments and maintenance, costs of sample processing, etc.), and how would they be overcome (e.g., DIY sensors, leveraging user facilities, crowd sourcing, etc.)?
    - How would new distributed research efforts leverage existing investments in physical infrastructure (e.g., SBR watershed test beds, National Ecological Observatory Network, USGS, etc.), data archiving (e.g., ESS-DIVE, DataONE, etc.), and computational tools (e.g., Kbase, CyVerse, Jupyter Notebooks, simulation codes, etc.).

- **What strategies would be key to energizing** the science community and other stakeholders to contribute to the efforts (e.g., providing resources, public data, recognition of individual efforts, etc.)?
    - What strategies would promote rapid and open sharing of data and ideas by individuals as part of distributed research efforts and how do we ensure fair-use and appropriate recognition (e.g., dataset DOIs, conceptual idea attribution, white paper DOIs, encouragement from funding agencies)?
    - What would the data management and sharing approach and policies be?
    - How do we curate and manage data so they are easy to use across networks?