Poster #161

LBNL Watershed SFA Data Management and Assimilation

Deb Agarwal[1], Charuleka Varadharajan[1], Roelof Versteeg[2], Boris Faybishenko[1], and Matthew Henderson[1]

[1] Lawrence Berkeley National Laboratory
[2] Subsurface Insights

Contact: Deb Agarwal [daagarwal@lbl.gov]

The objective of the Data Management and Assimilation (DMA) component of LBNL's Watershed Function SFA is to enable science by addressing the priority data science needs of the SFA team. The heterogeneous data sets collected at the East River site include over 80 different data types and includes hydrological, geochemical, geophysical, microbiological and remote sensing data. This data is being collected from a number of different sources. This includes data collection directly funded by the SFA (e.g. data from instrumented sensors, geochemical sampling efforts and remote sensing campaigns), data collected by a broad range of collaborators, and data from 3rd party sources (e.g. USGS streamflow and NRCS SNOTEL data). This data needs to be used as inputs for analytical and numerical modeling of the meteorological and hydrological water balance, including infiltration and groundwater recharge, as well as an evaluation of long-term trend of climatic conditions and nutrients balance.

The DMA team has developed a number of tools for automated data QA/QC, data ingestion, data exploration, data preservation, and data distribution. Automated data QA/QC is performed using statistical to identify and flag issues in the datasets. Data integration is achieved via a brokering service that dynamically integrates data from distributed databases via web services, based on user queries. The integrated results are presented to users in a web portal that enables intuitive search, interactive visualization and download of integrated datasets. Data preservation and distribution are being enabled by a web portal that allows authorized users to upload and download data files as packages. The DMA team is also building an interactive map of data collection sites run by the SFA and its collaborators, to inform the broader community about SFA field activities (the map is featured in the Watershed SFA Community Observatory page of the SFA website). Sites can be filtered by their key measurements and other metadata, leading to detailed site landing pages.

The data are used for building crosscutting data products needed for the hypothesis testing and numerical modeling of hydrological and biogeochemical conditions of the East River watershed by the SFA project teams. For example, we are analyzing a set of meteorological model drivers (temperature, rainfall, solar radiation, snow depth, snow water content, etc.) from a network of meteorological stations, and the river discharge from field observations along the East River.