RESEARCH STATEMENT: IMPORTANCE OF BIG DATA

Data has become the central aspect in all of science. In fact, most of the ten big ideas for future NSF investments (https://www.nsf.gov/about/congress/115/10bigideas.jsp) deal with big data. Many of them also have connections to water and remote sensing making my research ideas ripe for the future. These include harnessing the data revolution, human technology frontier and the rules of life. In all of these areas, the use of data from all sources – observations, model outputs and remote sensors are used in analysis. In the rules of life, the scaling up along the organism – population – community – ecosystem – biosphere pathway (a) Requires data at various spatial and temporal scales – scaling up on the pathway (b) Water plays a central role in each component of the pathway.

Research Problems

In order to study questions related to the variability of water in space and time in global watersheds, we need to use observations and models. The research areas are listed in the next two pages and all of them involve use of large data sets.

It is important to point out that use of satellite data for water storage and movement has many implications. Extreme hydrological events – droughts and floods are spatio-temporal problems that have societal applications for provision of disaster relief. These are addressed by federal agencies who use systems analysis to stage relief efforts in a manner and planners who use the spatially and temporally distributed information for future land use planning and layout of roads and design of evacuation corridors.

The movement and storage of water using spatially and temporally distributed data sets from observations and models is the key for many societal problems.

RESEARCH AREAS

Global Water Resources and Monitoring of Hydrological Extremes

Water is the most important natural resource we possess today on this planet. Too much of water in the form of floods can cause unspeakable damage destruction of property, loss of lives and have a huge impact on the economy of entire nations when they occur on a large scale. The other end of the water spectrum is droughts. Drought is defined as lack of water. Droughts result in loss of agriculture when pumped irrigation sources are unavailable and rainfall is the only source of water for the crops. The time scales for the inception and propagation of floods and droughts are completely different. Whereas floods, especially flash floods can happen in an instant whereas droughts take months and sometimes years to form. In the case of flooding, there is some advance warning in the precipitation forecasts using atmospheric models. In cases of downstream floods, the flow time from the upper reaches of the catchment to the lower reaches offers some advance lead-time. In the case of droughts, the lack of precipitation coupled by high evapotranspiration is a prescription for the upcoming disaster and this could take long periods of months and even years. With rapid advances in computer modeling and observing systems, floods and droughts are studies with greater precision today than ever before. Land surface models (especially over entire Continental United States) can map the hydrological cycle at kilometer and sub-kilometer scales. In the case of smaller areas (not the entire Continental United States) there is even higher spatial resolution and the only limiting factor is the resolution of input data. In-situ sensors are automated and the data directly relayed - to the world-wide-web for many hydrological variables such as precipitation, soil moisture surface temperature and heat fluxes. In addition, satellite remote sensing has advanced to providing twice a day repeat observations at kilometer to ten-kilometer spatial scales.

Improving spatial resolution of radiometer derived soil moisture

Soil moisture is an important variable in land surface hydrology as it controls the amount of water that infiltrates into the soil and replenishes the water table versus the amount that contributes to surface runoff and eventually to stream flow. However, observations of soil moisture at point scale are very sparse and observing networks are expensive to maintain. Satellite sensors can observe large areas but the spatial resolution of these is dependent on microwave frequency, antenna dimensions and height above the earth's surface. Soil moisture is generally observed at the L and C bands. Higher altitude in space of the sensor results in lower spatial resolution and conversely at low elevations the spacecraft would use more fuel due to increased atmospheric resistance and the mission life would be shorter. Large antenna diameter would provide higher spatial resolution but large antennas require more fuel to maintain in space. Given these competing issues most passive radiometers have spatial resolutions in 10s of kilometers. However, these spatial resolutions are too coarse for catchment hydrology applications. Most local applications require kilometer or sub-kilometer scale soil moisture data sets. We use radar data as well as visible and near infrared satellite data sets along with model products to downscale radiometer based soil moisture.

Optimal integration of models, point observations and satellite remote sensing data

Prediction of land surface and subsurface states by hydrological models is limited by (a) Model use of physical prognostic equations which may not represent reality, (b) Parameters that "lump" our lack of understanding of the system into a few physical but mostly non-physical numbers and (c) Inputs to land hydrology that may be incorrect in spatial variability and/or temporal description. Therefore, it is proper to merge observations and model outputs using a weighting scheme that assigns weights in proportion to the reliability of the estimates (observations versus model). Observations are classified as point-based (ground measurements: river discharge; measurements of soil moisture and temperature at a point) and spatial averages (remote sensing:

temperature, soil moisture, precipitation). A proper merging of these two estimates (at any specific time step) will help in re-configuring the model for the next time step, thereby helping to remove some of the discrepancies in the parameter inputs and/or model physics. With greater amount of ground and satellite data available to us today and in the foreseeable future, we should be able to optimize hydrological predictability. Work on validation of satellite retrieval algorithms and studies using satellite data at continental scales have yielded encouraging results.

Feedbacks: Land-Atmosphere Relationships using Satellite Remote Sensing

The atmosphere influences the land surface to a significant degree. Soil moisture and surface temperature are a result of coupled water and energy balance. However, this coupling is in two ways. The land surface influences the atmospheric processes through transpiration by the vegetation and the evaporation from bare soil. Increased soil water and radiation causes an increase in evapotranspiration that could lead to condensation and eventually precipitation (at the same location or elsewhere depending on advection). The presence of vegetation is important in governing the evapotranspiration (latent heat) component of water (and energy) budget. We currently have numerous weather-scale meso-scale meteorological models for studying these phenomena. However, some of these models have inadequate representation of land surface effects that can be improved using the hydrological modeling advances and satellite description of the vegetation.