ExaSheds: Advancing Watershed System Science using Machine Learning and Data-Intensive Extreme-Scale Simulation

Dipankar Dwivedi, (LBNL), Ethan Coon (ORNL)
Xingyuan Chen (PNNL), Ben Brown (LBNL), David Moulton(LANL),
Scott Painter (ORNL), Carl Steefel, (LBNL)

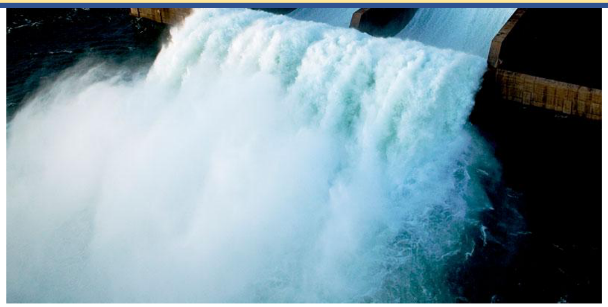
ExaSheds

U.S. DEPARTMENT OF ENERGY | Office of Science
BERKELEY LAB
OAK RIDGE National Laboratory
Pacific Northwest NATIONAL LABORATORY
Los Alamos NATIONAL LABORATORY EST. 1943

Healthy watersheds are critical to water security

- Watersheds funnel rain and snowmelt to rivers where they can be used by municipalities, agriculture, and energy producers.

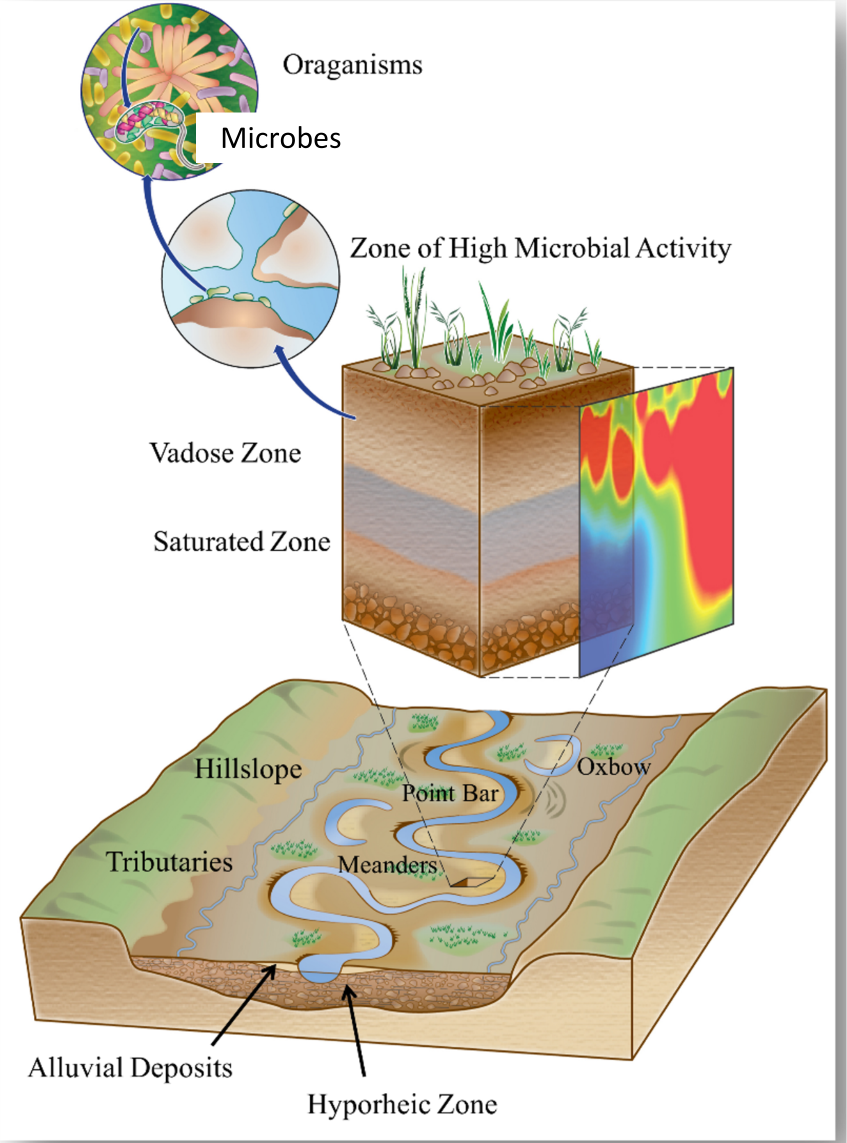
A predictive understanding of watershed function that is both predictive and mechanistic is required to ensure water security.



  - Increases in contaminant and nutrient inputs
  - Changing precipitation patterns and temperature

Gaps

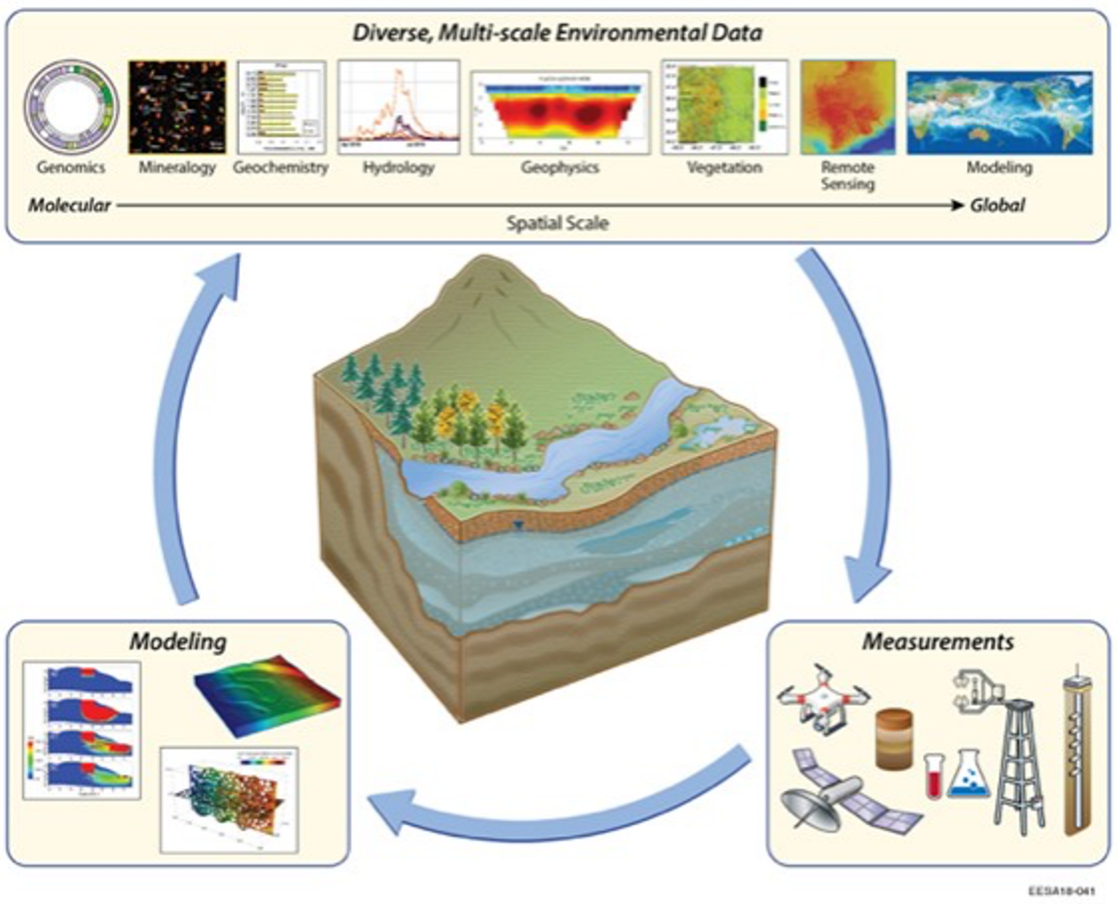
- Current generation **water quality modeling** tools for watersheds do not fully incorporate coupled multi-physics processes
- **Multi-scale character** of watersheds poorly represented
- Real **time data-model integration** is not presently considered, and rigorous UQ is neglected
- Mechanistic representations come with **large computational** and **characterization burdens**



ExaSheds: Advancing computational watershed science with machine learning and advanced simulations

5-year vision: Hyper-resolution, process-explicit hydrobiogeochemical simulations **at river basin scales** taking full advantage of diverse and spatially extensive data and providing feedback to design of distributed networks

Approach: Combine modern data-driven approaches with advanced integrated surface-subsurface hydrological and biogeochemical models leveraging leadership-class computing facilities



Varadharajan et al. (2018)

Learning-assisted multiscale simulations

Interoperable Design of Extreme-scale Application Software (IDEAS)

ATS + Alquimia
(code related development)

SAWASC (SFA)

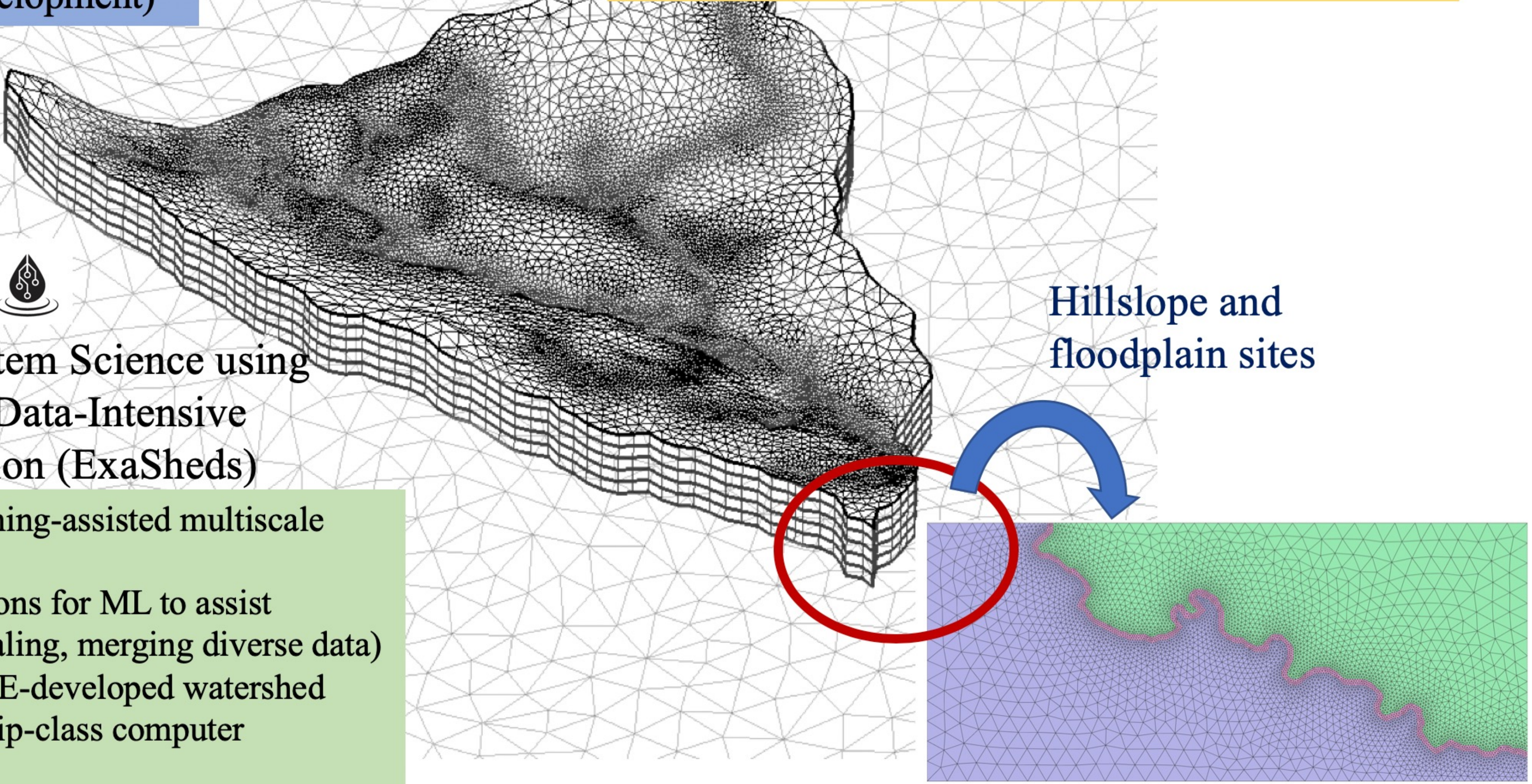
1. Develop Variable Resolution Meshes
2. ML-assisted bi-directional exchanges
3. Make use of zones (zonation approach)
4. Integrate subsystems (e.g., intensive and satellite sites) to get the watershed response

One Year Plan (Seed Project)

ExaSheds

Advancing Watershed System Science using Machine Learning for Data-Intensive Extreme-Scale Simulation (ExaSheds)

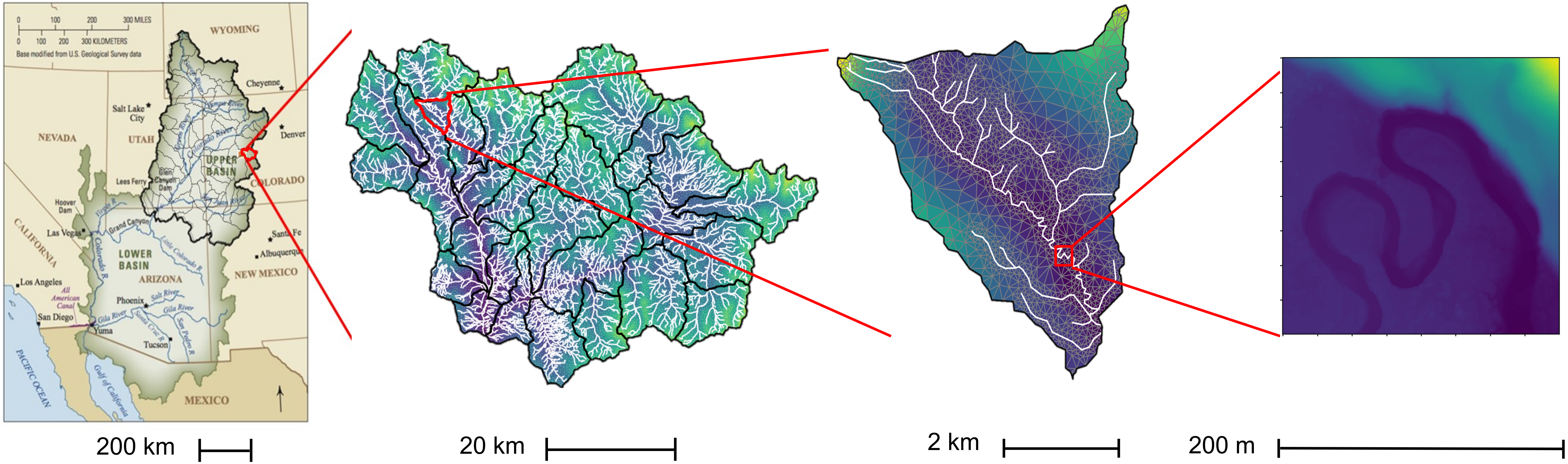
- Explore strategies for learning-assisted multiscale simulation
- Explore example applications for ML to assist simulations (e.g., Downscaling, merging diverse data)
- Preliminary porting of DOE-developed watershed modeling tools to leadership-class computer architectures



Learning-enabled Data Model Integration Across Scales

Data driving models across scales, from catchments to full-river basins.

- Digital elevation models: LIDAR → NED
- Soil databases: Geophysics → NRCS Soils
- Hydrography: LIDAR → NHD
- Meteorological data: Weather Stations → DayMet
- Land cover/vegetation: Mapping → NLCD
- Hyporheic exchange flux: Tracer experiments → NEXSS



Machine learning approaches for integrating between and across scales.