CI-WATER: Cyberinfrastructure to Advance High Performance Water Resource Modeling

Introduction: Population growth, shifting land uses, and climate variability are altering the magnitude and timing of water fluxes, stores, and availability in the arid Intermountain Western U.S. These alterations are driven by coupled human-natural system interactions at spatial scales ranging from farm plots and buildings to entire river basins and temporal scales from seconds to centuries. These pressures produce interconnected responses in atmospheric, surface, and subsurface processes, threatening the sustainability of natural water systems supporting fragile ecosystems and the resiliency of constructed water systems on which tens of millions of people depend. Numerous researchers have investigated the individual atmospheric, hydrological, geological, environmental, economic, and sociological components of this complex human-natural water system. However, components are only infrequently integrated to characterize the overall system sustainability. Furthermore, it is rare for comprehensive evaluation of integrated system behavior to inform or guide land use or water system planning or management by individual users, cities, water conservancy districts, or states.

This project, which we call **CI-WATER**, will allow a consortium of Utah and Wyoming researchers to acquire and develop hardware and software cyberinfrastructure (CI) to support the development and use of large-scale, high-resolution computational water resources models to enable comprehensive examination of integrated system behavior through physically-based, data-driven simulation. Successful integration requires data, software, hardware, simulation models, tools to visualize and disseminate results, and outreach to engage stakeholders and impart science into policy, management, and decisions. The computational requirements of stochastic methods to consider uncertainties, fine spatial and temporal resolutions to improve accuracy, and representation of dynamic processes that include feedbacks among system components demand use of state-of-the-art high-performance computing (HPC). We propose a robust and distributed CI consisting of integrated data services, modeling and visualization tools, and a comprehensive education and outreach program that will revolutionize how computer models are used to support water resources research in the Intermountain West and beyond.

Our proposed CI developments leverage and extend the shared expertise in Utah and Wyoming and align with the focus areas presented in both the current Wyoming Science and Technology plan and the draft Utah plan.

Project Goal and Objectives: The scientific problem that this project will address is*: How are the quality and availability of water resources sensitive to climate variability, watershed alterations, and management activities*? The CI challenge that we will address is: *How can we best structure data and computer models to address this scientific problem through the use of high-performance and data-intensive computing by discipline scientists coming to this problem without extensive computational and algorithmic knowledge and experience*? The project thus aims to broaden the application of CI and HPC techniques into the domain of integrated water resources modeling. To meet these goals, we propose four primary components:

1. Enhance cyberinfrastructure facilities. The demands of the proposed data-intensive modeling require additional hardware in the form of data systems and compute clusters to support the work proposed and enhance the research competitiveness of the Consortium states. This CI will include servers and disk farms connected to the four participating institutions via an ultra-high-speed network. The servers will host the data services, regional models, and computing clusters used to access and run simulations. It will also include a new 100+ TeraFLOP class computer system and a complementary data storage system.

2. Enhance access to data- and computationally-intensive modeling. We will develop easy-to-use model and data interfaces that link integrated system models running within an HPC environment to multiple data sources to enhance the use of computational simulation and data-intensive modeling to better understand water resources. We will establish a cybercollaboratory based on tools developed from NSF-funded projects and blend hydrologic data and information system capability with HPC gateway functionality into a platform for scientific discovery. These tools will reduce the time and expertise required to define, populate, and execute regional models on HPC. This will include new data services to provide hydrologic, geophysical, demographic, socio-economic, and legal water rights data needed to inform and interpret output from large-scale models run in stochastic predictive form.

3. Advance high-resolution multi-physics watershed modeling. We will revolutionize water resources modeling by integrating new physical process descriptions and linking with existing widely-used codes to enable long-term simulations with high spatial resolution to answer location-specific questions related to water availability in the face of growing demands, uncertain future meteorological forcings, and existing prior-appropriations water rights. The model will include engineered watershed features and water transfers, enabling water management and economic analyses.

4. Promote STEM learning and water science engagement. We will foster scientific and cyber-literacy and improve educational and research capacity within the Consortium through dissemination and communication activities reaching four audiences across the STEM learning pipeline: 1) the CI-WATER partners and other EPSCoR jurisdictions; 2) Utah and Wyoming K-12 students and their teachers; 3) higher education faculty, students, and water agency professionals; and 4) adult learners. Using water resource issues as the basis for educational and diversity-building activities, the Consortium will reach all levels of the STEM pipeline.

Management, Coordination, and Evaluation: We will manage the project in a manner that will result in cross-disciplinary and cross-jurisdictional scientific integration. We have identified metrics that will be used throughout the project to **evaluate and assess** the adoption and effectiveness of our cyberinfrastructure at enabling cyber-related discovery through the use of high-performance water resources models. This information will be used by the state committees and EPSCoR offices, an external evaluator, and an advisory committee to inform project direction following best industry practices.

Comprehensive Diversity Plan: Our project is ideally suited for engaging under-represented groups in STEM learning and bringing to them the excitement of cyber-related discovery and innovation that is important for retention. We will hold annual summer workshops for underrepresented students. The proposed web user interface to our system aids in **broad dissemination**, draws upon the capabilities in the Utah and Wyoming university systems, and will be integrated into curricula at the respective universities to provide a tight coupling between education and research components.

Intellectual Merit: The integrated data-intensive modeling enabled by the proposed CI will lead to better understanding of coupled natural and human water resources systems and their response and sensitivity to alterations across space-time scales. Advances in data and modeling systems that enhance HPC usability and access by non HPC specialists will transform the way hydrologic knowledge is created and provide broader informatics applicability beyond the field of water resources.

Broader Impacts: This project provides CI that improves access to data and sophisticated models, combines models and model components from different sources, enables scientists to populate models with readily accessible data, harnesses high-performance computing resources to perform multi-decadal simulations over large spatial areas with high space-time resolution, and transforms the way hydrologic knowledge is created and used in water resource planning and management. The CI enhancements in this project will be integrated into a robust education program focused on improving cyber-literacy throughout the region. Outcomes will be disseminated broadly to water users, scientists, planners, managers, regulators, industries, municipalities, engineers, state and federal agencies, non-profit entities, and the public. The measure of our success will be a transformational change in how students, researchers, practitioners, and citizens in our states use data and models to make meaningful predictions regarding influencing factors, choices, and future availability of ground and surface water. The adjacency of Utah and Wyoming provide compelling case studies related to water, energy, urban growth, climate, and ecosystems that will foster natural collaboration at all levels of education and across disciplines. Open access to the data and models combined with enhanced usability and active outreach will extend capability to those groups currently underrepresented in water resources.

4.1 Status

Within Utah and Wyoming, ongoing cyberinfrastructure (CI) efforts have created a unique environment and opportunity for the proposed research through existing network connectivity, existing and soon to be available high-performance computing (HPC) capacity, and ongoing NSF-supported CI development efforts. Although network connectivity and HPC access limitations remain at research institutions within the two states, **in this proposal we are addressing the significant challenge of creating access to data, models, and HPC resources in a way that scientists who are not computer scientists can use them**. Related to this is the development of a new generation of CI and HPC-literate students and researchers throughout the consortium.

In addition to their proximity in the mostly high deserts of the Intermountain West, Utah and Wyoming share a number of similarities in their CI opportunities and challenges. For large-scale data center development, both states offer relatively inexpensive electric power. Their cool, dry climates promote the use of ambient air most of the year for data center cooling, thus greatly reducing the operational costs of these facilities. The State of Wyoming is a major investor in the new National Center for Atmospheric Research (NCAR)/Wyoming Data Center (NWSC) outside Cheyenne and has committed $40M in state funds to the facility (see letter from W. Gern, University of Wyoming VP for Research). Concurrently, the University of Utah is making an investment in excess of $25M to convert a former industrial building in downtown Salt Lake City into a campus and statewide HPC and Data Center with availability slated for January 2012; half the facility is initially devoted to HPC, including 1.25 MW of power and space for 100 racks of HPC and storage hardware.

With NSF support, both States have also made or are currently implementing major investments in optical networks to connect the four research university campuses and the two new data centers to proximate national research and education network nodes in Salt Lake City and Denver. Networking investments have provided the University of Wyoming (UWYO) with excellent connectivity via redundant 10-Gbps connections along the Bi-State Optical Network (BiSON) to the Front Range GigaPOP.

4.1.1 Utah

Existing HPC resources in Utah are listed in Table 1.

Table 1. High Performance Computing resources in Utah

Institution	Resource/Initiative	Description
Brigham Young University	Fulton Super Computing Laboratory	Nine computing clusters with a total of 911 compute nodes that have 9,592 CPU cores with 27.1 TB of memory. The computing resources are supported by 100 TB of on-line storage. The largest cluster has 512 nodes, 6,144 cores, and 12,288 GB of memory. Other clusters are configured for different purposes such as large memory jobs (256 GB/node) and special uses, such as compiling code or software testing
University of Utah	Center for High Performance Computing	Four computing clusters: Ember 262 nodes with 3,144 cores, Updraft 256 nodes with 2,048 cores, Telluride 72 nodes with 576 cores and Sanddune Arch 156 nodes, with 624 cores. The Center also provides high-performance networking across the campus
Utah State University	Center for High Performance Computing	Three computing clusters for a total of 130 nodes

CI network connectivity is well established in Utah. The Utah Education Network (UEN) has established unprecedented connectivity among all higher institutions and K-12 districts throughout the state. Utah is fortunate as UEN already provides sustained network connectivity and collaboration tools for Utah researchers at a level above many other states, particularly in the Intermountain region. In addition, the ongoing EPSCoR RII C2 Cyberconnectivity project will deliver enhanced optical connectivity via the Research@UEN network to Brigham Young University (BYU), and additional planned UEN upgrades – with anticipated funding from a NTIA BTIP broadband stimulus award - will result in optical network upgrades to the University of Utah (UU) and Utah State University (USU). The C2 effort brings together the statewide networking capabilities of UEN with the respective CI and campus network leadership at UU, USU, and BYU. Research@UEN has been under development for two years, and its earlier design focused on (a) the Salt Lake City optical network to connect UU, a new state-of-the-art, off-campus data center, and the major national node in the Intermountain West and (b) a fiber-based extension to Logan in support of USU. In addition, the C2 project will make strategic investments in local campus networking capabilities to better enable EPSCoR research and science, technology, engineering and mathematics (STEM) activities at UU, USU, and BYU. The ongoing RII C2 project and UEN efforts are enhancing the level of collaboration and computational engagement among the three institutions. Combined with resources described for Wyoming below, there will be unprecedented potential to expand our ability to collaborate with researchers nationally and internationally through high-speed access to Internet2.

UEN's Distance Learning infrastructure hosts over 550 video-connected classrooms in Utah and licenses a statewide web conferencing tool, Wimba, and learning management system, Instructure Canvas. These resources are available for online meetings and courses in support of this project. CI-WATER goals and objectives are attainable because this fundamental education infrastructure is already in place, widely adopted, and proven at scale.

4.1.2. Wyoming

At the University of Wyoming, the Information Technology Center (ITC) Computer Room opened in late 2008 with a 6,000 ft^2 raised-floor machine room with space, cooling, and electrical power for at least 200 standard computer racks with traditional CPU-based blades. Hybrid CPU/GPU has higher power and cooling requirements, which may change the number of racks that can be supported without modification to the machine room infrastructure.

There are 18 dedicated clusters on campus, ranging from a few dozen connected PCs to a 512 CPU cluster to a 64 GPU compute cluster. Most of the clusters are small and only the Institute for Scientific Computing's clusters (a blade server and the GPU cluster) are generally available to anyone on campus. The GPU cluster has a few applications running at 3 TFLOPS on all of the GPUs. Like many campuses nationwide, UWYO has reached the limits of sustainability of the current research computing model in which individual researchers acquire, house, and maintain their own high-performance computing environments. Areas that are using supercomputing include earth and environmental sciences, energy, engineering, water resources, economics of mineral usage, complex ecosystems, physics, chemistry, bioinformatics, and general computational sciences. These groups, as well as groups studying subsurface flow, will be able to do useful research with the proposed CI.

A key aspect of UWYO's developing cyberinfrastructure landscape is its partnership with the State of Wyoming, NCAR, NSF, and local agencies in the NCAR-Wyoming Supercomputing Center (NWSC), which is being built just 40 miles away from UWYO. The NWSC will begin operation in June 2012, and 20% of the NWSC compute and storage resources are allocated through agreement for UWYO faculty and student research each year for the next 20 years. The NWSC will support geological, and earth-system sciences computing. A requirement for using the NWSC facilities is that an application be proven to scale on a computer system with at least 1% of the peak speed of the NWSC machine. The modeling activities described in this proposal will meet the criteria for research eligible to run on the NWSC (see commitment letter from J. Reaves, Associate Vice President of UCAR).

4.2 Results from Relevant Prior Support

GeoInformatics: CUAHSI Hydrologic Information Systems, PIs: D.R. Maidment, **D.G. Tarboton**, I. Zaslavsky, J. Goodall, and D.P Ames, EAR 0413265, $1,156,059, 2004 - 2008 and EAR 0622374,

$4,500,000, 2007-2012. **Tarboton** and **Horsburgh** are involved in the design and implementation of the CUASHI HIS cyberinfrastructure, which will serve as a foundation for portions of the work proposed here. The CI components developed [1-4] have been deployed at WATERS Hydrologic Testbeds [5], and at numerous other sites (40+ observation networks networks). The central HIS catalog holds metadata for web-accessible hydrologic time series from over 1.8 million measurement points in the US, while the WaterML schema [6] has been adopted by USGS and NCDC to publish hydrologic data. HydroDesktop [7] is a client application for accessing HIS data.

DataONE (Observation Network for Earth), PIs: W. Michener, T. Vision, S. Hampton, and R. Cook, OCI 0830944, $20,000,000, 2009 – 2014. **Horsburgh** is a member of the DataONE Core Cyberinfrastructure Team, which is tasked with the architectural design and implementation of the DataONE CI. The DataONE CI is designed to enable the long-term preservation of diverse and complex multi-scale, multi-discipline, and multi-national science data by providing open, persistent, robust, and secure access to well-described and easily-discovered earth observational data. Horsburgh is also a co-lead of the DataONE Semantics and Data Interoperability working group, which is working toward the integration of semantic technologies for data discovery and integration within DataONE.

Tools for Environmental Observatory Design and Implementation, PIs: D.K. Stevens, **D.G. Tarboton, J.S. Horsburgh**, and N.O. Mesner, CBET 0610075, $355,936, 11/2006 – 10/2009. **Horsburgh** and **Tarboton** developed several software applications that have now become part of the current CUAHSI HIS HydroServer Software stack [8-11].

Wyoming EPSCoR Research Infrastructure Improvement, PI: Randolph Lewis, S. Jackson, EPS 0447681, $8,604,167, 05/01/05 – 06/30/11. The current award focuses on improving infrastructure and educational outreach in ecological systems from molecular to global scales. Three core facilities needed for ecological research were established or enhanced using EPSCoR funding including equipment, renovations, and research salary for the Nucleic Acids Exploration Facility (NAEF), which provides low-cost DNA sequencing and genotyping and training for graduate students and teachers. Since its start in 2005, 29 funded projects (>$2.8 million total) involving 14 faculty members and 32 graduate students have been conducted at the NAEF. The Stable Isotope Facility (SIF) was originally funded by an NSF Major Equipment award, but EPSCoR support enabled the hiring of a manager and purchase of additional equipment. Since 2005, the SIF has supported 33 funded projects (>$5.5 million total) for 26 faculty members and 42 graduate students. The Wyoming Geographic Information Sciences Center (WyGISC) was developed under EPSCoR funding, and in 2009 supported more than 200,000 downloads by more than 10,000 different users. Under the current EPSCoR award, an ecoinformatics specialist was hired to provide expertise and training to researchers across campus. All three staff positions are now state funded. Additionally, four new faculty positions in ecology and related areas have been supported. All four positions were filled by outstanding female scientists. Wyoming EPSCoR has fostered research as an important educational tool for students at UWYO and community colleges, supporting approximately 100 students and faculty members per year under a competitive Fellowship Program, increased by 50% since the start of the program.

4.3 Active NSF RII Awards

In this section, we highlight the active NSF RII C2 project in Utah, which is directly relevant to the proposed CI-WATER project. The current NSF Track 1 RII project in Wyoming is slated for completion on June 30, 2011. With a focus on ecological research, it established essential core facilities that have provided foundational infrastructure for the advances proposed here.

Extending Campus Networks and Research @ UEN Optical Network in Support of the Utah EPSCoR Initiative, PI: **Steven Corbató**, University of Utah, $1,176,470 (09/01/10 – 08/31/12)
Scope of Work: This RII Cyber Connectivity (RII C2) award is leveraging the facilities and statewide reach of the UEN to expand the optical networking capabilities of the research and education communities to more effectively engage faculty and students across Utah in STEM fields. The project is

extending network capabilities (initially provisioned at an aggregate bandwidth of 30 Gbps) of the *Research@UEN* optical network to BYU. This segment is complementing capabilities already under development USU and UU through separate funding, but closely coordinated with the RII C2 project. *Research@UEN's* Phase 1 development includes a Salt Lake City metropolitan optical network and a fiber-based spur to Logan in support of UU and USU, respectively. Combined with a southward extension to Provo for BYU, these advances will enhance the level of collaboration and computational engagement among the three institutions, and greatly expand their capabilities to collaborate with researchers nationally and internationally through high-speed access to the national research & education network, the Internet2 Network. Another component of this RII C2 is targeted improvements in campus networks in support of HPC and EPSCoR-related, computationally-based research at the three Utah research universities. ***Alignment and Integration with Proposed Work:*** The RII C2 award is significantly improving the baseline bandwidth and capacity-provisioning capability for BYU and the other two research universities in this project. In particular, the associated Salt Lake City metro project will establish optical connectivity to new downtown data center hosting the Utah data store and the telecommunications node in western Salt Lake City hosting the regional node for the Internet2 Network, which in turn provides the high-speed connectivity between the Utah and Wyoming research components.

4.4 Cyberinfrastructure-Enabled Science and Engineering Program

Limited water supplies and increasing demands for agricultural, urban, industrial, recreational and environmental uses make long-term water resource planning and management critical to the sustainable growth and economic vitality in the western United States. Water suppliers are increasingly forced to identify and develop alternative sources and implement conservation measures. With uncertain future supplies, effective planning and management requires a comprehensive understanding of the complex interactions among different parts of the natural and managed water systems. Successful planning must also predict future needs and identify the sensitivity of the system to changes in inputs and uses.

This project, which we call **CI-WATER**, will allow Utah and Wyoming researchers to **acquire and develop required hardware and software CI to support the development and use of sophisticated, large-scale, high-resolution computational water resources models.** These models will link surface water, ground water, and water management components to evaluate hypotheses about the functional behavior of water resources systems in the western U.S. that includes Utah and Wyoming.

In the western U.S., snowmelt is the primary source of water, and water availability is sensitive to changes in climate, runoff and infiltration flow paths, land cover and land use, water resources infrastructure, irrigation practices, trans-basin diversions, water rights, environmental policies, and urban development. Because most western U.S. rivers are already over-allocated, anticipated future decreases in flows due to climate change will pose significant problems for water managers [12].

Understanding, predicting, and mitigating impacts of future land use change, energy development, and climate variability at key points in western rivers will require compiling information from multiple sources and using integrated, data- and computationally-intensive modeling approaches. Data collection and model integration must include socio-economic factors, environmental policy, and water rights to predict the effects of water management strategies and projects on water availability in a region where every drop of water is allocated. The uncertainty regarding the impacts of non-hydrologic factors on water storage and management in Lakes Powell and Mead were noted by the National Resource Council [12]:

> *"climate change implications for streamflow and reservoir management of the many individual upper basin tributaries upstream from Lakes Powell and Mead may vary considerably from those for the mainstem because of seasonal, topographic, legal, and physical infrastructure constraints particular to each specific sub-basin."*

Organizing data and integrating system components will require acquiring, developing, and deploying new CI, and executing models with high spatial and temporal resolution over large geographic areas on

HPC systems. The CI and models must include representations of natural hydrologic processes and the engineered environment and prior-appropriation water rights systems that are central to allocating water resources in western U.S. states such as Utah and Wyoming. This CI will allow us to accurately simulate hydrological processes over small spatial and time scales, and use the model to integrate cumulative effects over large river basins and multi-decadal time periods considering the effects of legal constraints, socio-economic factors, and environmental policies.

The CI and supported knowledge-based simulation capabilities proposed will enable scientists, researchers, water managers, state officials, students, and others to test hypotheses related to hydrologic processes, demand forecasts, infrastructure development, policies, socio-economic conditions, and component interactions. The focus areas of water and CI directly align with the published Wyoming Science and Technology (S&T) Plan (2010) [13] and are consistent with ideas being incorporated into the plan being developed for Utah. The Wyoming S&T plan has three focus areas: water, energy, and computation. The research programs, computational capacity and predictive capability to determine water availability are key areas that will be advanced by this proposed multi-state effort, and will contribute directly to Wyoming's S&T efforts.

The problems we propose to address are multi-disciplinary, and CI-supported integrated modeling provides a way to formalize communication across the disciplines involved. Utah and Wyoming are ideally positioned to collaborate due to both the commonality of hydrological challenges and complementary research capabilities. The team has an excellent working relationship and history of prior collaboration. Our proposed CI infrastructure developments and acquisitions will provide ready access to required data, link easily to models, run quickly, and promote data- and computationally-intensive science as a new paradigm for discovery and cyberlearning. The CI and computing resources will be configured so that researchers, water managers, and others can effectively use the HPC resources without detailed knowledge of CI or HPC. The following sections describe our project components, broader impacts, and the nature of the intellectual partnership formed by our collaboration

4.4.1 Component #1 – Enhance Cyberinfrastructure Facilities

Challenges: At each university, we have identified new CI facilities required to address the scientific needs posed by Components 2 and 3. Accordingly, we will acquire and deploy appropriate development and staging environments that will enable the project teams to develop and prototype the needed software and modeling components of this project. In addition, we plan to purchase and deploy production web servers and enhanced storage resources within the UU data center to host project websites, data services, and datasets developed by this project. The enhanced storage resources in the UU's data center will serve as a staging area for assembling large data packages to be sent to HPC resources at the time of model execution.

Proposed Activities: Currently, the largest scientific computers at UWYO are a 512 CPU cluster and a 64 GPU compute cluster. Through this project, UWYO will acquire a 100+ TFLOPs hybrid multi-processor computer system and associated storage resources dedicated to supporting the combined HPC modeling needs of our Consortium. We expect the UWYO computer system to be lighter on disk storage than normal since we intend to use the proposed data facilities at UU to house large data sets that will be read at the beginning of simulations and saved after the simulations are completed. A significant portion of the disk storage at UWYO will act as a large disk cache where all consortium members can stage large datasets and model packages for timely execution on the large machine. The NWSC machine also will have a GPU component so we easily can migrate code developed on the proposed UWYO machine to NWSC.

At the UU's new data center, we will deploy a Utah/Wyoming atmospheric and hydrologic data repository (CI-WATER STORE) with 180 Terabytes (TB) of usable storage initially and at least one 96 TB couplet added annually thereafter (roughly 0.4 Petabyte overall). This will provide the CI foundation to meet the data storage, sharing, analysis, and curation requirements of the high-resolution, data-intensive modeling proposed. If the pending Utah and Wyoming Track-1 EPSCoR proposals are funded, additional funds will be available for storage acquisition, which will be fully integrated with these resources. We plan to utilize the capabilities of the UU's new data center, including reliable power, engineering and operational support, and an existing HPC-based storage array that can accommodate the proposed growth with minimal impact on users and system administrator time. CI-WATER STORE will be managed centrally and deployed as dedicated or shared storage resources. File system allocation policies can be tuned at the name space, directory, and individual file levels to meet the mixed usage requirements. This system granularity also will enable snapshots, auto-tiering (usually used to move unused files to slower but more economical disks, drives, or tape), replication, and quotas. Additional storage resources and development server capability will be established at USU and BYU to support development work at those institutions.

Timelines and Milestones

CI-WATER Milestones and Timeline	Year 1				Year 2				Year 3			
(assumes 9/1/11 start)	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3
Component 1. Enhance Cyberinfrastructure Facilities												
Establish development servers, storage resources, and workstations	■	■										
Finalize Utah-based storage system architecture and procure system	■	■										
Establish production storage resources at UU, USU, and BYU		■			■				■			
Vendor bidding ready specifications for HPC resources	■				■				■			
Functioning new UWYO supercomputer+associated storage resources			■				■				■	
Operational UWYO HPC resources			■	■	■	■	■	■	■	■	■	■

Long-term Impacts and Outcomes:

The hybrid, high performance cluster at UWYO will:

1. Serve as a gateway to the NWSC machine, which is projected to approach 2 PetaFLOPs in peak computational speed, for running simulations of large watersheds, such as the upper Colorado River basin.
2. Run at a sustained speed of 1% or greater of the peak speed of the NWSC machine in order to gain access to NWSC. Our CI will provide a platform in order to achieve this level of performance, which has never been seen before in hydrology codes.
3. Enable and support to the development of a new hydrology model specifically targeted at HPC resources that are fast enough to evaluate 'what if' scenarios, especially important to the public and government officials who may require these results in a very short period of time.

The CI-WATER STORE data system at UU will:

1. Support and provide access to large volumes (initially hundreds of Terabytes and scalable to Petabytes) of hydrologic and atmospheric data while leveraging existing storage infrastructure.
2. Provide a database and data modeling infrastructure to support integration and assimilation of local, regional, and national data tied to the research projects in this Consortium.

4.4.2 Component #2 – Enhance Access to Data- and Computationally-Intensive Modeling

Challenges: Currently, there exists a digital divide among most hydrologic researchers (experimentalists and modelers who have process-based modeling expertise) and HPC experts who have technical hardware and programming knowledge to efficiently use HPC resources and the resource managers who need to translate the outcomes to practical planning. This divide limits comprehensive modeling and the greater understanding of integrated hydrologic-human processes that we believe it would provide. Additionally, modelers spend much of their time finding and organizing required model input data, rather than engaging in analyses that lead to scientific discoveries. Thus, the significant CI challenges tackled by the work proposed under this component are to: (1) provide hydrologic researchers, modelers, water managers, and users access to HPC resources without requiring they become HPC and CI experts and (2) to reduce the amount of time and effort spent in finding and organizing the data required to execute the models. This challenge is really the CI to elevate water resources modeling to fully exploit the "fourth paradigm" use of computational simulation and data-intensive modeling to gain understanding from massive datasets that have emerged as avenues for scientific discovery [14].

Proposed Activities: To address these challenges, we will develop easy-to-use model and data interfaces that link integrated system models running within an HPC environment to multiple data sources, allow users to efficiently and effectively explore alternative physical representations of study systems, and describe the sensitivity to and consequences of alternative representations on the integrated system represented by the models. Additionally, we will develop user interfaces that summarize and visualize data and model outputs to support research, hypothesis testing and inform water management decisions. This combination of data, modeling, analysis, and visualization capability will result in a collaboratory in support of data- and computationally-intensive water resources modeling (Figure 1). In this context, collaboratory refers to a community-specific computational environment for research and education that provides HPC services, data and information services, knowledge management services, human interface and visualization services, and collaboration services, all of which are essential to facilitating high scientific productivity [15].

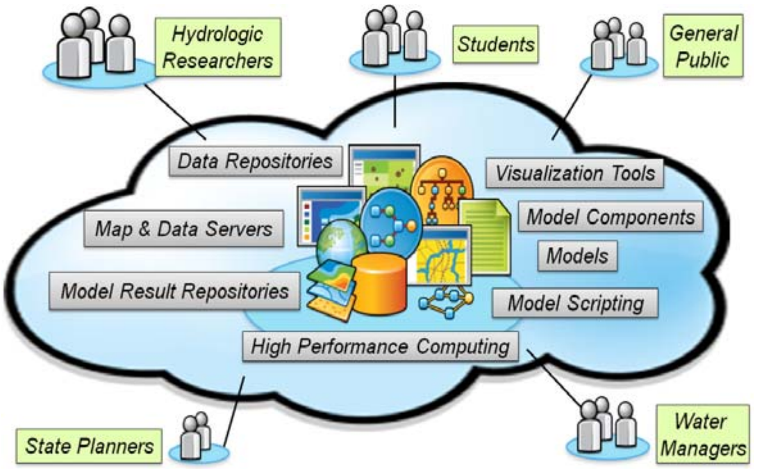

Figure 1. Data- and Computationally-Intensive Modeling Collaboratory.

We will build a Science Gateway CyberCollaboratory based on established tools developed associated with the NSF-funded TeraGrid project [16, 17]. Our CyberCollaboratory will blend hydrologic data and information system capability developed as part of our work on the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) Hydrologic Information System (HIS) [2, 3, 18] with internet portal and content management capability developed in support of high performance and grid computing by the NSF-supported HUBZero project [17]. We will draw upon open source ezHPC and ezVIZ components developed at the US Army Engineer Research and Development Center (ERDC) Information Technology Laboratory (ITL). These provide graphical user interface and visualization elements for HPC. Our collaborative partnership with ERDC on projects such as GSSHA (http://chl.erdc.usace.army.mil/gssha), TauDEM (http://hydrology.usu.edu/taudem) and WMS (http://chl.erdc.usace.army.mil/wms) (see commitment letter) will facilitate access to these capabilities.

The focus of this component is the CI framework that enhances access to data- and computationally-intensive modeling. It is complementary to Component 3 below, which focuses on enhancing multi-physics modeling capability. Teams working on Components 2 and 3 will be closely coordinated and have considerable overlap in membership.

Data Services. We will build upon the data sharing capabilities of the CUAHSI HIS to develop the data services required to meet the input data requirements of the models we develop and for sharing model outputs. Where possible, we will use existing HIS data services, such as those that have already been created for the USGS National Water Information System. When important datasets are identified for which data services do not exist, we will establish, either ourselves, or in collaboration with partners, data services that provide access to these datasets through standards-based web services that support data sharing and interoperability.

It is anticipated that in supporting the data needs of this project, we will need to go beyond the capabilities of the existing CUAHSI HIS by creating functionality to support additional classes of data, including: (1) water resources infrastructure data; (2) data characterizing water uses and demands; and (3) urban water systems data. We will develop this additional capability by extending the HIS data model, drawing upon other existing domain data models such as those that USU Co-PI Rosenberg and others have developed as part of a modular decision support system for network-based (node and link) water management models (www.hydroplatform.org) [19]. Indeed, we have already identified a number of data sources with data that fall into these categories and whose availability needs to be enhanced to support data-intensive modeling. Work involved here includes acquisition of the needed data or arranging for its online availability in a useable format from its primary source. Important datasets include: (1) National Weather Service Colorado Basin River Forecast Center operational hydrologic data such as components of the river basin water balance model, reservoir storage levels, estimated diversions and losses, meteorological fields (rainfall and temperature) and hydrologic fluxes and storages (e.g., soil moisture and snow water equivalent) (See Schmidt commitment letter); (2) Central Utah Water Conservancy District Central Utah Project operational decision data (see Shawcroft commitment letter); (3) Utah Division of Water Resources water supply data (see Strong commitment letter); and (4) gridded model inputs such as soil properties, land cover, and topographic properties from several sources. Publishing gridded data is presently possible using the CUAHSI HIS; however Unidata capability (e.g. NetCDF, OPeNDAP, THREDDS, http://www.unidata.ucar.edu/) may be better suited in some cases. We will investigate their addition to the data services framework we establish.

This work will include the identification and design of data representations for engineered water resources infrastructure (such as on- and off-stream reservoir storage, detention basis, hydropower generation facilities, diversions, canals, irrigated areas, pipelines, water supply treatment plants, wastewater and stormwater collection systems, wells, wastewater treatment plants, and outflows) and the physical attributes that describe them, such as capacity, storage-elevation-area curves (for reservoirs and basins), release/outlet capabilities, among others. Infrastructure information will also include operations such as how storage is partitioned among uses, release rules, water rights, and observations of operations. It will also involve the design of data service capabilities to describe water uses and demands – including the location, spatial extent, use type, and timing disaggregated to the scale for which data is available or needed, and the design of data models and data service capabilities for data related to water systems in urban environments with links to regional-scale models. Quantities to include are representations of: metropolitan area water demands for potable and non-potable uses, wastewater discharges, conservation and reuse information, and rainwater harvesting. It is anticipated that model outputs will be of similar spatial and temporal structures and scales as model inputs and that the same data models and data sharing capabilities developed for model inputs can be used for model outputs.

Modeling Services. Modeling services will be built around the functionality of HUBZero (http://hubzero.org/). At HUBzero's core is a compilation of open source tools and systems including: GNU/Linux, Apache Web Server, OpenLDAP, PHP, MySQL and the Joomla content management system (CMS). These software and server tools support sharing of models and information and enable interactive simulation and modeling from a web browser interface [17]. We will create a HUBZero instance and will configure it to serve as a portal for access to the data and HPC modeling resources we have proposed.

Integrated modeling and model interoperability is an area of active research [20-23]. We will build on capability for model interoperability established using standards such as OpenMI [24] and systems such as CSDMS (http://csdms.colorado.edu/). We will develop a suite of tools for processing model inputs, defining model scenarios, and post-processing model results. These tools will be integrated with our portal and will be developed as independent building blocks compatible with HPC platforms (e.g., the portal tools will enable users to assemble model input data packages, define model scenarios, and then ship them off for execution on an HPC platform). This approach allows the blocks to be re-configured into tools to construct more complex workflows based on the needs of a particular modeling study. We will initially develop a set of low-level tools for processing the input data associated with modeling scenarios. This will include tools for harnessing web services for accessing input data. Hydrologic modeling frequently requires data pre-processing, including terrain analysis and watershed delineation. This can be burdensome for large areas at fine resolution. We will build on the parallel processing capability developed for terrain analysis for this work [25].

When it comes to running models, multiple runs are frequently required for calibration and as part of stochastic simulations required to quantify uncertainty. Our HUBZero portal will be configured to facilitate the use of HPC resources and for running ensembles on complex integrated models and organizing and visualizing the output.

Implementation. Our primary operational Water Resources Data and Modeling Collaboratory will be located in the UU data center (CI-WATER STORE). USU and BYU will both establish development platforms that allow virtualized model and data system development. These will be used to develop and test the data and model services functionality prior to deployment at CI-WATER STORE.

Timelines and Milestones

CI-WATER Milestones and Timeline	Year 1				Year 2				Year 3			
(assumes 9/1/11 start)	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3
Component 2. Enhance Access to Data- and Computationally-Intensive Modeling												
Water management components for data services												
Develop data model design	■	■										
Implement prototype data services			■	■								
Test and refine prototype data services					■	■						
Implement operational data services							■	■				
Urban data components for data services												
Develop data model design	■	■										
Implement prototype data services			■	■								
Test and refine prototype data services					■	■						
Implement operational data services							■	■				
Develop and deploy data services for datasets required by models	■	■	■	■	■	■	■	■	■	■	■	■
Develop a HubZero instance that interfaces with Utah/Wyoming HPC resources	■	■	■	■								
Develop HubZero functionality			■	■	■	■						
Develop appropriate user interfaces within HubZero to enable user access to models					■	■	■	■				
Develop model output post processing and visualization tools							■	■	■	■		
Develop community model collaboration capabilities									■	■	■	■

Long-term Impacts and Outcomes

This component will result in unprecedented access to data for water resources modeling in the Western U.S. as well as unprecedented capability for non-HPC specialists to take advantage of these data and perform data- and computationally-intensive model simulations to evaluate hypothesis, quantify uncertainty and embrace computational thinking in their research. Implementing the data services, modeling tools, and model integration in a HPC environment will provide the ability to execute full-detail watershed models in long-term continuous simulation mode with incorporated uncertainty. The new data available on water resources infrastructure will facilitate previously impossible detailed simulations of water resources systems sensitivity properly accounting for the presence of human alterations in the system. This will enable a fundamentally deeper understanding of water resources in Utah and Wyoming. Specific key contributions from this component include:

1. Establishes a collaboratory in support of data- and computationally-intensive water resources modeling by providing access to the massive databases and models to a variety of users;
2. Allows for the creation and execution of new simulation scenarios with minimal effort, including stochastic analyses;

3. Supports simplified creation of model input derived from data accessed via web services;
4. Provides post-processing, visualization, and communication of the massive output produced by HPC models, and
5. Provides opportunities to share and collaborate through use of community modeling tools.
6. Extends legacy code functionality to the HPC environment.

4.4.3 Component #3 – Enable High-Resolution Multi-Physics Modeling

Challenges: The Upper Colorado River Basin provides a natural setting for this proposed collaboration between Utah and Wyoming since continued trends towards increases in temperature will likely result in reduced runoff from upper basin States [12, 26-28]. Effective water management requires watershed-scale integration of data and simulation models within a socio-economic and prior-appropriations water rights framework that seeks to minimize the negative impacts of management decisions upon stakeholders and the environment [12]. In the western U.S., watershed management is most often performed at the scale of river basins, with key points where water must be delivered on demand according to in-state water rights or interstate compacts. These points consist of irrigation, industrial, or municipal diversions, are identified at the boundaries between States and other points defined by interstate compacts. Current models use grids of >1 km^2 and run overnight on a single CPU, whereas important processes such as snowmelt occur at scales of 100 m^2. Furthermore, existing models do not represent the built environment and water rights with sufficient detail to evaluate infrastructure and land-use changes at scales where they predominantly occur. Effective management with uncertain forcing and inputs requires consideration of large numbers of scenarios. The CI requested will allow results to be obtained in a timely fashion so as to be useful to government agencies, water managers and planners and the community in general.

Current State of Modeling: Hydrological modeling of large river basins in the western United States has been possible for decades. However, detailed, high-resolution research hydrologic models [29-31] either do not run on HPC systems or simulate the complexity of the built environment or prior-appropriation water rights. Two models that have been used with significant success in the Colorado River Basin are the Variable Infiltration Capacity (VIC) model [32], and Riverware [33]. These models have been successfully applied at large scales, and rely on empirical relations or statistics to account for subgrid features that cannot be resolved within the model spatial discretization. Neither of these models contain fully-coupled hydrodynamic processes necessary to simulate all feedbacks associated with land-use and climate changes that make management of over-allocated rivers in the western U.S. challenging.

The large-scale semi-distributed VIC model has been used primarily as a research tool to assess climate change impacts on stream flows. The VIC model uses a conceptualized infiltration parameterization on large grids. Land surface heterogeneities in topography and land cover are treated using statistical distributions. Grids do not interact with each other and are simulated sequentially, with river routing performed using linearized open channel flow equations of motion. The VIC model does not simulate groundwater in a way that is suitable for simulating the effects of groundwater withdrawals on surface flows [33].

Riverware is used by the U.S. Bureau of Reclamation to set monthly operations targets for the large reservoirs in the Colorado River Basin. Riverware uses a topologic scheme to link different land-surface features with a river network, and is optimized for reservoir operation and hydropower production. Hydraulic routing within channel reaches is performed using different conceptualized flow routing schemes or approximations of the open-channel momentum equations. Riverware has an optional sequential link to the U.S. Geological Survey MODFLOW groundwater code to compute exchanges between surface and groundwater.

The conceptualizations employed in both VIC and Riverware make explicit simulation of land-use change effects difficult because of uncertainties regarding these changes on their parameter values. Land-use changes due to both human (e.g. conversion of agricultural lands to subdivisions) and natural (e.g. pine bark beetle) are very important factors in the Colorado River basin [34].

Proposed Activities: There is a need for high space-time resolution models that can simulate the effects of changes at the scale of pine stands and agricultural fields on the response of the Colorado River basin. This model will have broad applications for hydrological studies, watershed management, economic analysis, as well as evaluation of changes in environmental policy and law. To be useful, the results must be available in minutes and we must be able to perform "what if" inquiries using Monte-Carlo techniques [35, 36] and ensemble Kalman filtering methodologies [37, 38] for data assimilation without re-running the entire simulation.

We propose creation of a new model to fill this important need that is specifically developed for HPC environments, not single PCs. Our goal is to include all hydrologically important processes and anthropogenic influences at scales of relevance using a physics-based approach wherein parameter values are measurable and identifiable and provide useful results for non-researchers. The proposed CI-WATER HPC model will operate at grid sizes as small as 10 m, and ingest data from remote sensing sources, atmospheric models and geophysical techniques. The model will include structured- and unstructured-grid options, as well as adaptive grid capabilities [39, 40]. Different model domains (e.g., surface water, groundwater) may use different space-time numerical discretizations where appropriate [29, 41] and will be capable of multi-scale resolutions using uncertainty quantification methods [42].

The new model will include the following features: three-dimensional groundwater simulation capabilities for predicting mountain-front infiltration and the effects of groundwater pumping on stream flows. The vadose zone will be simulated using a one- and two-dimensional multi-physics algorithm [43] that can simulate variably-saturated matrix, macropore, crack, and film flow. Surface flow algorithms will include two-dimensional overland flow routing, and one-dimensional full-dynamic channel flow routing on a dendritic (vector) channel network. The channel network will interact with lakes, reservoirs, detention basins, irrigation diversions, and trans-basin transfers.

The proposed model will include engineered features. In urban areas, the influence of subsurface storm drains will be simulated [44], as will urban water applications and consumptive use. Agricultural water use will be simulated in accordance with prior-appropriations water rights using records from State Engineers. Irrigation canals will be explicitly simulated as will irrigation. Irrigation water not consumed by evapotranspiration will be partitioned into groundwater recharge and return flows to streams. Evapotranspiration will be calculated using meteorological inputs and standard equations or input from remote-sensing estimates.

The accumulation, re-distribution, sublimation, and melt of snow above 2400 m (8000 ft) elevation provide the dominant source of water in the western U.S. [45]. The CI-WATER HPC model will include published snowmelt routines and facilitate testing of advanced routines to simulate these processes. High spatial resolutions employed will allow simulation of snow redistribution in mountainous terrain [46, 47], while above-ground LiDAR returns that quantify the location and distribution of above-ground canopy will allow explicit simulation of stand-level effects of vegetation on snow accumulation and melt [48]. Large eddy simulations will allow explicit simulation of snow/topography/vegetation interactions [49], which is very important given the pine bark beetle epidemic underway in the region [34].

The model design will include development of application programming interface (API) standards to allow easy substitution of alternative process-level simulation routines [50]. Existing widely-used and successful codes will be employed to avoid "re-inventing the wheel". The modeling system will include ensemble Kalman filtering [51] to improve forecast prediction accuracy while quantifying the effect of model formulation errors on predictions. Monte-Carlo techniques [35] will allow evaluation of the effects of parameter, data, and model uncertainties when used in a predictive mode. We will use finite volume and finite element discretizations on various grid types. Explicit numerical solvers will provide highly scalable parallelism to take advantage of the UWYO CI. Fluxes are needed on multiple processors and we will employ an efficient, highly scalable database approach providing a virtual shared memory using either commercial or freeware versions of a similar system [52-54].

As smaller discretizations are employed and/or larger watersheds are studied, the parallelism will grow. At the 10 m land-surface resolution, the memory requirement for each ensemble member for a demonstration scale problem (Green River basin in Wyoming) is about 200 GB. For the Upper Colorado River basin above Lake Powell, it is about 5200 GB. While sensors and mapping databases would have to be improved significantly, in the future a 1 m LiDAR-derived land-surface resolution will require 100 times the memory. For a single ensemble member, at some land-surface resolution and number of state variables we will have to move to the NWSC facilities since we will be running simulations too large for the UWYO CI. Further, we need 20-100 ensemble members so the memory requirements might outstrip the capacity of the first computer at NWSC by the end of our three-year project and might be a candidate for the NWSC in the 2015-2016 timeframe.

Timelines and Milestones

CI-WATER Milestones and Timeline	Year 1				Year 2				Year 3			
(assumes 9/1/11 start)	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3
Component 3. Advance High-Resolution Multi-Physics Watershed Modeling												
Evaluate existing model codes for compatibility with project objectives	■	■										
Design a high-resolution, multi-process, linked regional and urban hydrology model	■	■										
Evaluate existing HPC API's, select and/or modify	■	■		■			■					
Define model input data structures	■	■										
Develop model interfaces that enable model subcomponents to be linked & substituted	■	■	■	■	■							
Adapt existing model codesusing CI-WATER interfaces and in an HPC environment			■	■	■	■	■	■	■	■	■	
Develop new hydrologic process solvers using CI-WATER interfaces for HPC			■	■	■	■	■	■	■	■	■	
Build model component coupling capabilities			■	■	■							
Develop, populate, and execute model instances and case studies				■	■	■	■	■	■	■	■	■
Evaluate and deploy parameter estimation routines for HPC environment					■	■						
Develop ensemble Kalman filtering scheme for forecasting in HPC environment								■	■	■	■	■
Model performance benchmarking							■	■				
Transition to larger scales on NWSC								■	■	■	■	■

Long-term Impacts and Outcomes

The proposed CI-WATER HPC model promises a significant advance that will revolutionize the way in which water is managed in the western U.S. The model also provides a bottom-up HPC modeling system for hypothesis testing and model rejection within a multi-university collaboration. This work will provide a strong and direct complement to but does not directly overlap with proposed Wyoming Track I activities. Track I is focused on understanding fundamental land surface-atmosphere interactions in the context of changing vegetation, changing climate and complex terrain. CI-WATER HPC could provide a framework for incorporating results from Wyoming Track I efforts and other ongoing studies in the western U.S. into a foundation for improved hydrologic prediction and the operational use of water resource modeling. Once validated, this model will provide sensible and latent heat fluxes that would allow effective coupling as a lower-boundary condition for coupled climate models (see letter of collaboration from NCAR Research Applications Laboratory).

4.4.4 Component #4 – Enhance STEM Learning and Water Science Engagement

Challenges: An update to the report *Rising Above the Gathering Storm* [55] indicates an immediate, high priority to increase America's talent pool, research, and incentives for innovation to assure U.S. security and competitiveness. The most pervasive concern is the state of U.S. K-12 education, which on average lags behind other industrial economies - while costing more per student than any other developed country. Utah and Wyoming educators face similar obstacles to engaging learners in environmental science. This project comes during a period when educators report a severe lack of timely science content for their formal and informal learning environments. According to a recent assessment by the Utah Society for Environmental Education, only 31% of Utah teachers received any pre-service training in environmental education, and the majority of elementary school teachers say they are not comfortable teaching environmental and natural history or environmental science issues [56]. Wyoming teachers report similar needs for environmental science content. Further, only 23% of Wyoming's adult population holds a college degree, pointing to the need to educate at all levels of the population.

Proposed Activities: We will advance NSF Core Priorities, particularly science and technology education in formal and informal learning environments. We will foster scientific cyber-literacy and improve educational and research capacity within the CI-WATER consortium through education activities designed to reach three audiences across the STEM learning pipeline: 1) High School Students and Teachers; 2) Higher Education and Water Professionals; and 3) Adult Learners. STEM learning and community engagement partners include the Utah Museum of Natural History (UMNH), the UU's Genetics Science Learning Center (GSLC), Wyoming PBS, and The Science Zone in Wyoming, each of which has a strong track record in STEM learning with educators and citizens. See Figure 2 for details of the activities proposed.

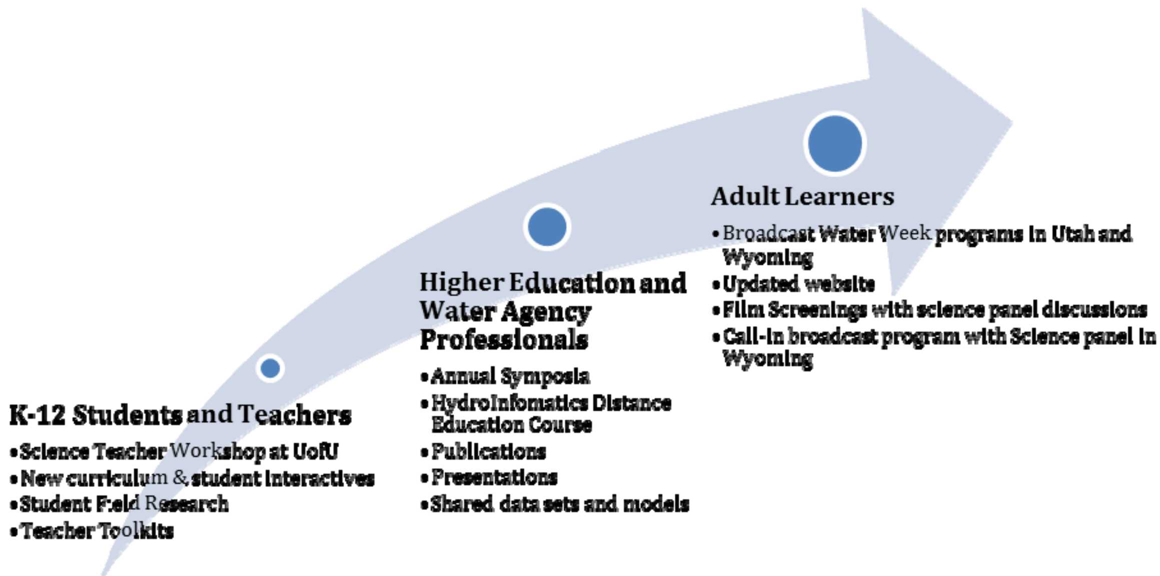

Figure 2. Audiences in the STEM learning pipeline.

Each participating organization is committed to building scientific literacy and strengthening the educational and research capacity throughout and beyond the Consortium. We specifically focus on cyber-literacy with a goal to improve an understanding of quantitative and computationally-intensive scientific methods, and water scarcity issues specific to the West. To accomplish this goal, the UEN will serve as the lead education partner working with the WY EPSCoR office to create STEM learning opportunities designed to develop a diverse, well-prepared, globally engaged, and scientifically literate public for our specific target groups, as described below.

1) K-12 Students and Teachers. We will foster STEM learning through coordinated K-12 activities between experts in Wyoming and Utah. The Genetics Science Learning Center at the UU has world-renowned expertise in curriculum development of scientific concepts, and will conduct a curriculum development summer institute that brings together 15 Utah and Wyoming secondary science teachers, project scientists, and staff to learn about and discuss the module topics, develop learning objectives, and create highly engaging online learning experiences. UEN and Wyoming PBS will broaden the distribution of these resources nationally through our partnerships with the PBS Digital Learning Library.

Wyoming and Utah will partner to offer a field workshop experience for up to ten high school students from both states. The workshop will held in Wyoming and will follow the model of a successful high school summer program for minority and first generation college students. Using the workshop template for this ongoing program, Wyoming will offer one-week intensive workshops to introduce computer technology for studying water resources management problems. Partnering with The Science Zone in

Casper, WY, these high school students will participate in a five-day field-based environmental education program. Students will be selected to represent geographic, cultural, and ethnic diversity.

During the second CI-WATER project year we will collaborate with the GK-12 "Science Posse" project at the University of Wyoming, which provides STEM education assistance for middle and high-school students and their teachers in Wyoming. The Science Posse has committed one Ph.D. fellowship to the CI-WATER project (see letter of commitment from R. Mayes) to assist in identifying a compelling cyberlearning opportunity, developing an example curriculum, and educating teachers from around Wyoming and Utah about the CI-WATER project with an example application. Instructional Toolboxes focusing on CI-WATER topics will be created by the Utah Museum of Natural History and The Science Zone in Casper, Wyoming. Toolboxes typically contain physical specimens and print materials to support engaging, relevant classroom instruction and preparation for field experiences. A sample water modeling toolbox will include water-based experiments and hydrology modeling computer programs in an age-dependent game format. The toolboxes will vary in complexity with age of the target group. Over the duration of the project, eight toolboxes will be built using national science standards. Toolboxes will be available for check-out through each museum's teacher lending program, with four placed in rural regional education centers, offering direct in-classroom access to water modeling resources.

Participating institutions will host evening or day-long Code Challenge camps for high school undergraduate students where participants will have the opportunity to tour HPC facilities, meet faculty and graduate student project staff, and tackle small coding and programming challenges. Camps will include refreshments (pizza, soda), offer prizes for the most elegant, efficient, or creative programming solutions, and result in paid summer-in-residence opportunities for students at participating institutions.

2) Higher Education and Water Agency Professionals. We will foster collaborative research and strategic partnerships among institutions, junior faculty, students and industry professionals. New educational programs will be initiated in both states. In Utah, a new graduate-level 3-unit class on Hydroinformatics and water resources modeling will be jointly and simultaneously offered by USU, BYU, UU, and UWYO as an online course through the shared UEN Instructure/Canvas LMS system to Master's and Ph.D. students at all of the associated campuses and taught jointly by faculty from all involved Universities. Curriculum will focus on conventional and spatial database design and querying, web services, model population, and script programming.

UWYO has recently approved an interdisciplinary Ph.D. program in hydrology. The goals of this program are strongly tied to CI-WATER outcomes and will enhance research, educational capacity, and hydrology infrastructure. USU will leverage existing graduate research training activities to include new Ph.D. CyberInformatics fellowships. Additionally, in years 2 and 3 of the project, USU will sponsor a summer coding challenge that will provide graduate students with the opportunity to submit proposals for paid summer coding internships. Two to three students will be selected each year to work on real CI challenges facing the project team in development of the proposed CI. Combining efforts, these hydrology Ph.D. programs and coding opportunities will improve graduate student research and educational opportunities both during the duration of the grant and after its completion.

State agency interactions are particularly important to encourage and maintain a communication pipeline between researchers and water professionals. These interactions are mutually educational because agency needs must be well understood by the researchers and model outcomes must be conveyed by the researchers. These activities will be promoted through university partnerships with state and federal entities, such as the Water Development and State Engineers Offices, Game and Fish Departments, and the Bureau of Reclamation offices in both Wyoming and Utah.

Each collaborating campus will support part-time student assistants at the undergraduate or graduate level to work directly with Consortium partners under the supervision of faculty advisors with specific recruiting for these positions of students from underrepresented groups.

3) Adult Learners. By partnering with public broadcasters UEN-TV and Wyoming PBS, this project literally reaches into every home in both Utah and Wyoming. Both stations will share water-themed local productions and documentaries, many produced by local PBS station KUED. The award winning films highlight water issues particular to the West, such as *A Desert Sea*, and *Desert Wars: Water and the West*. UEN-TV and Wyoming PBS will broadcast Water Week programming as a coordinated PR effort, during all three years of the project and produce short spots highlighting the project and directing viewers to the project website WaterWiseWest.org for additional information.

Screenings of these films will be held at each consortium location for scientists, citizens, and stakeholders. Panel discussions and demonstrations of new modeling tools will be presented at these events as a way to engage general audiences and increase cyber-literacy and regional water understanding.

Timelines and Milestones

Activities to accomplish Component 4 will occur throughout the project, with most events for educators occurring in the summer months, and activities for the general public coordinated around Utah's annual Water Week in May.

CI-WATER Milestones and Timeline	Year 1				Year 2				Year 3			
(assumes 9/1/11 start)	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3
Component 4. Promote STEM Learning and Water Science Engagement												
Teacher curriculum development institute at GSLC and curriculum launched								X	X	X	X	X
Wyoming 5-Day field based environmental education program								X	X	X	X	X
Instructional toolboxes for teachers complete and deployed to schools/museums							X	X	X	X	X	X
Annual One-Day Water Symposium at UMNH			X				X				X	
Online HydroInformatics course is complete and taught ea/semester				X		X		X		X		X
Select and license water tv programs; produce on-air PR spots	X	X										
Broadcast Water Week on UEN-TV and Wyoming PBS			X				X				X	
Film screenings and panel discussions			X				X				X	
Project website Water Wise West launches and is updated	X	X	X	X	X	X	X	X	X	X	X	X
Summer workshops for HBCU students				X				X				X
Code camps for high school and undergraduate students				X				X				X
Summer of code program				X				X				X

Long Term Impacts and Outcomes

The education partners for CI-WATER were carefully selected based on their outstanding prior work and success in STEM learning. The project will produce tangible outcomes in the form of curricula, and age-relevant computer games. Teacher workshop partner GSLC is world-renowned for their ability to make complex science engaging and understandable. Each year, their website researches over 7.1 million unique visits from over 180 countries. Likewise, UEN and Wyoming PBS have wide reach and impact, with a recent Roper poll of Americans ranking public TV as the most trusted public institution [57]. Key contributions from this component are:

1. New collaborative online course in Hydroinformatics
2. Greater public awareness of water issues and water science through screenings, broadcast, panel discussions and new Water Wise West websites
3. Programming challenges, code camps, and summer of code opportunities for high school and college students
4. Production and national distribution of new curricula for high school classes that includes computer modeling games to build quantitative skill and computer literacy methods of hydrology research, and improve analytical reasoning skills
5. Production of CI-WATER teacher toolkits that offer specific instruction in computer literacy and quantitative problem solving to fill a materials gap in rural Utah and Wyoming schools
6. New annual Water Symposium resulting in increased networking among water researchers, students, water industry professionals and the general public

4.4.5 Intellectual Partnership

The members of the CI-WATER Consortium bring a wealth of experience on water resource modeling, HPC and data management. The top part of Figure 3 shows the areas of expertise of the participating institutions. UU is a leader in urban watershed modeling and CI and leads the Utah outreach effort through UEN and other programs. UWYO has world-class HPC resources and is a leader in hydrologic modeling, remote sensing, and GRID computing. The Wyoming EPSCoR office leads the Wyoming outreach and educational efforts through collaborations with the Science Math Teaching Center and The Science Zone. USU and BYU bring significant expertise in water resource modeling and modeling tools. This includes model interfaces, data services, and decision support systems. More extensive information is provided in the research biographical entries.

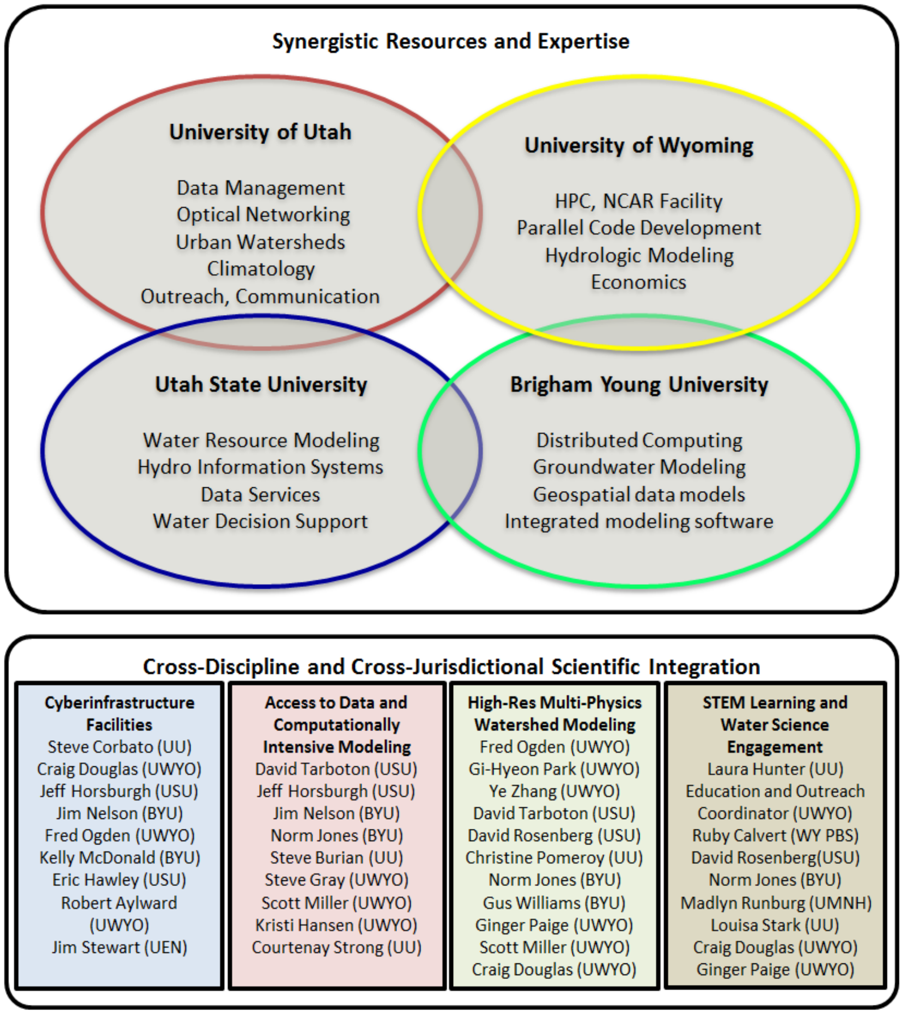

Figure 3. Institutional expertise and multi-disciplinary teams.

We have created cross-disciplinary and cross-jurisdictional research teams to take advantage of these diverse skills as shown in the bottom of Figure 3. Each team is assigned a primary project component and includes members from each institution. The team members will use a monthly development cycle with frequent team meetings to collaborate and manage progress (Agile management – see Section 4.9).

Table 2 illustrates how the consortium intellectual partnership leverages the skills and resources of each institution to meet the proposal objectives.

Table 2 Consortium partnerships to achieve project objectives

Component	USU	UU	BYU	UWYO
Enhance cyberinfrastructure facilities	HPC resources, development servers, storage	HPC resources; servers; disk farms	HPC resources	New 20-50 TFLOP class computer
Enhance access to data- and computationally-intensive computing	Data services; basin modeling; water management	Host production portal, data services, and data curation	Model integration; software development	HPC; hydrologic modeling

	modeling	environment		
Enable high-resolution multi-physics modeling	Specification of model inputs via data services; data models for model inputs; hydrologic process representations	Urban watershed modeling; verify urban hydrology dynamics	Data models for model inputs; groundwater process representations	Lead to develop regional modeling code, NCAR access
Enhance STEM Learning and water science engagement	PhD fellowships; code camps; summers code challenges	UEN, K12 engagement, outreach symposia	Code camps, undergraduate mentoring	Science Zone, Wyoming PBS

4.5 Diversity Plan

The CI-WATER partners are committed to developing programs that support diverse learners, particularly students and junior faculty who are women, minorities, first generation college students or located in rural areas. Currently, Utah and Wyoming have relatively little ethnic diversity, as compared to other states. Wyoming has 88% Caucasian population as compared to the national average of 65.1% Caucasian. Successful diversity recruitment at UWYO, reporting 82% Caucasian students attending UWYO, represents a higher ethnic diversity than the state overall. The leading ethnic groups in Wyoming are Hispanic 7%, followed by American Indian 2.5%, and African American 0.9%. Utah ethnicity data is similar, with 81.2% Caucasian, 12.3% Hispanic, and 2.1% Asian [58]. We will focus recruitment and research activities toward underrepresented groups, while also recognizing that disabled and rural students in both states face unique challenges, particularly a lack of access to advanced science and technology education, and will make a particular effort to support their engagement as well. We have targeted three major levels for diversity enhancement.

1) Increasing Diversity in Higher Education. Students from underrepresented groups will be encouraged to apply for the new undergraduate and graduate research jobs included in this proposal, in order to foster meaningful school-to-work experiences. This recruitment will be supported by the Diversity office on each campus and coordinated by Lisa Cohne in Utah and the EPSCoR office in Wyoming. Other specific target groups include:

a. American Indian students: UU will increase American Indian participation in the computational sciences and HPC system engineering by 10%. This will be accomplished by recruiting American Indian students from a joint Research Experience for Undergraduate (REU) program based in Pittsburgh, Richland, and Salt Lake City. CI-WATER will partner with the American Indian Resource Center (AIRC) and the campus chapter of the American Indian Science and Engineering Society (AISES), under the sponsorship of Professor Otakuye Conroy-Ben (Civil & Environmental Engineering) at University of Utah. In addition, Utah collaborators will participate in the AIRC's annual high school outreach event to provide early visibility for these opportunities to prospective American Indian undergraduates, who are drawn primarily from the Navajo, Ute, Paiute, Goshute, and Shoshoni tribes.

UWYO has ongoing education and research projects with Northern Arapaho, Eastern Shoshone students in Wyoming and Crow students at the Wyoming Montana border. Under the direction of Ginger Paige, senior project personnel, the high resolution database and the model tools that will be developed from CI-WATER will be further developed for on-the-ground water resource and watershed planning efforts. Along with integration into education curriculum, these tools will be used for on-ground outreach extension programs in watershed and water quality monitoring and management, which are currently ongoing in Wyoming. Track 2 funding will support two additional workshops and up to 10 American Indian students from the region. Outcomes of the activity will include water quality and watershed assessment curricula for the Tribal Colleges as well as water resource workshops.

Utah State University will coordinate fellowships for Ph.D. students in Cyberinformatics targeted towards minority students and will also encourage underrepresented groups to apply for summer code camps, summer coding challenges, and part time computer programmer positions that will be sponsored by this project.

b. Interactions with HBCUs: UWYO will hold one-week summer workshops for students from Historically Black Colleges and Universities (HBCUs) with whom UWYO has current memoranda of understanding. These workshops will be overseen by project co-PI, Craig Douglas, and the UWYO Education and Outreach Coordinator (to be hired). The UWYO Diversity Office at Wyoming supports such workshops via a competitive funding opportunity through the Strategic Diversity Initiatives Committee (SDIC). The one-week workshop will recruit up to 10 students from partnering HBCUs and will include hands on use of CI and instruction in hydrology and HPC technology. Student training will be provided by faculty in the departments of mathematics and civil engineering. The workshop will coordinate with ongoing atmospheric science workshops at UWYO operated by Bart Geerts, coordinated through the Diversity Office, and facilitated by the WY EPSCoR office.

2) Increasing Diverse Learners in High Schools. Greater diversity will also be achieved through our summer workshop exchange and Code Challenge camps, described in more detail in section 4.4.4. Utah will host a teacher workshop with teachers from both states, and Wyoming will host a workshop for students from both states. For the summer teacher workshop in Utah, science teachers from high schools with high ethnic diversity will be given priority enrollment, and 1/3 of the workshop seats will be open only to rural science teachers. High school students who apply for the summer workshop in Wyoming will be selected to represent gender equality and ethnic diversity among the applicant pool, with particular recruitment efforts toward highly diverse and rural schools. Each institution will host a Coders Challenge. Participants will be invited to tour HPC facilities and develop innovative solutions to small programming challenges. Diverse students will be recruited for this activity as a path toward eventual college enrollment and pursuit of a STEM career.

3) Reaching Diverse Learners in the Community. Finally, our broad informal learning activities include partnering with WY and UT Public television and local museums (described in more detail below in section 4.6). These activities will reach rural populations in Utah and Wyoming through museum and community film screenings, panel discussions with scientists, broadcast Water Week programming, multimedia websites, high school lesson plans and toolboxes, and email mentoring. All programs aired during Water Week are close captioned for deaf or hearing impaired accessibility. The web site development team has advanced training on accessibility and builds to ADA standards, including software readers for the blind or visually impaired.

4.6 Dissemination and Communication Plan

Both Utah and Wyoming have distinctly rural and isolated populations that will benefit from greater CI improvements proposed here. The CI-WATER programs will facilitate cyber-enabled learning, greater understanding of water scarcity issues, and will bridge our local, state and regional boundaries. Significant synergy will result from collaborative communication among the partners, and more broadly with our constituents and peers.

1) Partner Communication. The Consortium will establish effective internal communications among partners and other EPSCoR jurisdictions to enable efficient sharing of data and information. Leveraging the existing network and videoconferencing capabilities of UEN, the Consortium will deploy a web-based collaboration system to enable on-demand meetings with desktop sharing, video, and audio. Monthly newsletters will be produced by the consortium and distributed to other EPSCoR jurisdiction leaders and posted on the project website. Each partner location will have a student communications/outreach assistant dedicated to this project who will support project-related communication between the partners.

We have already identified several CI-WATER collaborations that will be useful to communicate with other EPSCoR jurisdictions and peers. For example, our high-performance computing infrastructure is deployed at each institution, but will be jointly employed for collaborative data sharing and modeling. Our new Hydroinformatics classes will be jointly taught by project personnel from two or more consortium institutions. Undergraduate, MS, or PhD student advisory committees include project personnel from two or more consortium institutions working collaboratively.

2) Dissemination of Scientific Results. The CI-WATER team will increase the overall knowledge base of our respective disciplines by publishing model input data packages, model simulation results, models, and code as open resources to the public within one year of being generated or deemed production ready. Published data products, models, and code may also include a data access/use agreement specified by the data publisher that conforms to established practices and policies already in place within the Consortium Universities. Faculty will weave the results of this EPSCoR opportunity into new collaborative proposals and journal articles, from two or more consortium institutions.

3) Dissemination of Education Tools. Educational products will be disseminated broadly through the unique partnership between the UWYO Science Math Teaching Center (SMTC) and the UT Math Science Consortium-facilitated through each state's EPSCoR offices. The partnership provides a shared exchange program between Utah and Wyoming for high school teachers and their students in the areas of cyber-learning and water resource management. STEM media programs developed by UEN will be further developed through a partnership between UEN-TV and Wyoming Public Television, with an ultimate goal of national dissemination of relevant water curricula to the national PBS Digital Learning Library. Additionally, toolkits advanced as successful teaching modules for museums and the general public in Utah will be distributed and used throughout Wyoming.

4) Communication with Water Industry Professionals. Both Wyoming and Utah will focus additional communication efforts toward the water planning staff at key state agencies (e.g., Wyoming Water Development and State Engineers Offices) and resource managers within state and federal agencies ranging from the Wyoming Game and Fish Department to regional Bureau of Reclamation offices, the Central Utah Water Conservancy District, and the Utah Department of Natural Resources. This communication will provide a seamless transition from research and formal education to industry and STEM careers. The agencies benefit from the development of modeling and application frameworks, and researchers benefit from input and advice on critical needs and will also enrich the activities of the CI-WATER. Topics of communication for these professionals and the CI-WATER team include initial scoping work, the role of modeling in long-range (i.e., seasonal and longer) runoff forecasting, and the generation of runoff and water availability scenarios for use in planning. Findings will be published in industry journals and presented at technical conferences to increase their reach.

5) Communication via Annual Symposia. A one-day annual symposium will be held at the Utah Museum of Natural History's new LEED-certified building, opening in Fall 2011, to communicate project results, benefits, and processes to faculty, graduate students and water agency professionals. The Museum location has garnered national attention for its ecological and aesthetic design, and will serve as a fitting setting for an event at which scientists, citizens, and stakeholders will meet to discuss water resource modeling and issues of water scarcity specific to the West.

6) Online Communication via Web Sites. Effective engagement also depends on presenting CI-WATER information to learners of all ages through a website; therefore, UEN will update the website WaterWiseUtah.org, a well-established resource which includes water conservation and awareness resources targeted to children, educators, and the general public. The updated site will be renamed Water Wise West to include Wyoming and surrounding regions and include new information from the proposed project and curriculum material. To increase website use as a learning resource, UEN will also produce embeddable widgets for key site features that Consortium partners, other EPSCoR jurisdictions, and social media clients can embed on their own websites to drive additional traffic and learning.

4.7 Evaluation and Assessment Plan

Project evaluation and assessment will provide quantitative and qualitative data that helps the project team and Steering Committee to: 1) refine and improve project implementation at both State and Consortium levels; 2) measure project progress in meeting objectives; 3) assess the impact of the project in developing collaborations that address inter-jurisdiction water resources modeling challenges; and 4) quantify the project's impact on the use of data-intensive learning in education and outreach activities. Our evaluation and assessment is comprised of: (a) Internal evaluation and reporting; (b) Independent external evaluation by an AAAS panel on an annual basis; and (c) an External advisory committee.

Internal evaluation and reporting: The Steering Committee will implement procedures to collect and report progress according to the metrics listed in Table 3 on a monthly basis. These metrics are designed to help answer formative and summative questions aimed to identify project effects and impacts. These questions include: what strategies are most effective to engage represented and underrepresented students in STEM activities? What strategies are most effective to boost collaboration among Utah and Wyoming consortium members? How has enhanced hardware and software infrastructure improved access to hydrologic data and model use? How has enhanced access to data and models improved cyber learning and discovery by students, faculty, and stakeholders? How has enhanced access to data and models improved understanding and management of water resource systems? What wider water research, management, and educational activities have students, faculty, and stakeholders in Utah, Wyoming, and beyond undertaken using project-developed CI resources?

Table 3. Proposed metrics to monitor achievement of project objectives.

Broad Goal Metrics. Enhance Utah and Wyoming collaboration in the broad area of Water Resource Modeling - Overall numbers of system users disaggregated demographically by level (faculty, students, stakeholders, STEM K-12), group (women, underrepresented minorities), affiliation (project members, consortium universities, or outside consortium universities), and capabilities of the system used (Data storage, Data access, HPC, multi-physics modeling, educational material) - Numbers of publications and graduate theses where the underlying research utilized the capabilities enabled by this project, authors demographically disaggregated as above, and itemizing publications from two or more consortium institutions - Number of classes jointly taught by project personnel from two or more consortium institutions - Number of undergraduate, MS, or PhD student advisory committees that include project personnel from two or more consortium institutions - Numbers of collaborative proposals awarded to project personnel from two or more consortium institutions - Reasons project personnel undertook or did not undertake collaborative student advisory committees, proposals, and papers as determined from an annual questionnaire
Component 1. Enhance cyberinfrastructure facilities - Capacity (in TB) of additional data storage online at each participant institution available to the Water Resources Modeling community served by the project - Capacity (in TeraFLOPs) of additional compute cycles available to the Water Resource Modeling community served by the project - Utilization (%) of available compute cycles at the new UWYO facility - Utilization (%) of the Utah data storage facility - Advanced network capacity (Gigabits per second) available between the participating universities

Component 2. Enhance Access to Data- and Computationally-Intensive Modeling

- Number of data sites accessed by users through project HPC portal
- Size (MB), spatial, and temporal scales of data accessed by users through HPC portals
- Location, prominence, and popularity of data sites accessed by users through project HPC portals
- Reasons users accessed data through project HPC portal (from online response form)
- User suggestions to improve the HPC data access portal (from online suggestion form)
- Number of legacy hydrologic modeling codes ported to HPC environment

Component 3. Advance high-resolution multi-physics watershed modeling

- Number of hydrologic process solvers written for HPC environment
- Numbers of model runs made
- Reasons users made model runs (from online response form)
- User suggestions for improvements (from online suggestion form)
- Useable watershed scale for multi-decade single- and ensemble hydrologic forecasts with resolution below 100m

Component 4. Enhance STEM Learning and Water Science Engagement

- Number of new courses developed by consortium members that use CI-WATER infrastructure
- Number of students enrolled in and courses that use CI-WATER infrastructure
- Student ratings of courses and instructors using CI-WATER infrastructure
- Number of stakeholder workshops held
- Numbers of associated PhD fellowship applicants, awards, and graduates
- Numbers of proposals submitted by high school, undergraduate, and graduate students to participate in project computer programming challenges
- Numbers of underrepresented groups involved in various STEM activities
- Numbers of research experiences awarded to computer programming challenge participants
- Subsequent college and employment activities of computer programming challenge participants
- Numbers of PhD dissertation, undergraduate and MS theses that comprise student project work
- Satisfaction indicators in survey responses from participants at Annual Water Symposium

AAAS Evaluation and Assessment: We will contract with the Research Competitiveness program (http://www.aaas.org/spp/rcp/) of the American Association for the Advancement of Science (AAAS) to provide independent evaluation and assessment of this project (see Milutinovich commitment letter). AAAS is experienced as an independent evaluator of this type of project having provided similar services to over two dozen NSF EPSCoR and NIH INBRE programs. AAAS, through the AAAS Research Competitiveness Service, will recruit and lead a diverse panel to provide assessment and guidance to the work supported by this award.

The panel will be comprised of four experts, including three recruited by AAAS for their expertise in the areas of the program initiatives, and one AAAS senior staff member. The panel will include expertise in using cyberinfrastructure to enable water resources research and modeling, and promote STEM education and collaboration. In addition, the panel will have an understanding of EPSCoR and experience in managing large, collaborative programs.

The AAAS panel will use the information provided through the internal evaluation process, along with stakeholder interviews during a site visit, in order to: analyze the information available; interpret it in the national context; assess progress toward program goals; and most importantly, make recommendations on guiding the program as a consequence. This leads to recommendations on scientific and technical directions; on management activities; on supporting infrastructure and policies; and on the evaluation process itself, if any or all need to be modified for the program to have the best chance of success.

There will be three visits over the course of the three year award. The result of each two- to three-day site visit will be a comprehensive report that will summarize the findings of the AAAS panel. While the report is intended to be used as a tool for managing the program, it will be written in a way that it can be communicated to all stakeholders: the participants, both scientists and management; the participating universities and colleges; the state committee and other interested parties within the state government and private sector; and the NSF.

External Advisory Committee: In addition to the AAAS panel, which will provide formal independent formative and summative assessments of the project, we will convene an external advisory committee (EAC) to provide direct peer review and expert guidance. The EAC will meet annually to monitor progress towards project goals. The EAC will (1) review progress toward achieving outcomes of the project; (2) make constructive suggestions for improving and or changing the direction of the work underway; and (3) advise the CI-WATER Steering Committee. Internationally and nationally recognized experts with an appropriate balance of expertise, skills and diversity have been selected (Table 4) to serve as EAC members. Michener has expertise in cyberinfrastructure facilities and data systems (components 1 and 2) as well as EPSCoR experience. Minsker has experience in HPC and cybercollaboratories (components 2 and 3). Britton has experience in science communication (component 4) and Wilhelmi has experience in data models and interfacing terrestrial and atmospheric modeling (components 2 and 3). All have agreed to serve in the capacity of EAC members for this project.

Table 4. External Advisory Committee Members

Dr. Bill Michener	University of New Mexico, PI of NSF DataOne DataNet project
Dr. Barbara Minsker	University of Illinois, National Center for Supercomputing Applications
Mr. Ryan Britton	Managing Partner of EarthSky and Principal Investigator of the EarthSky en Español project
Dr. Olga Wilhelmi	Head of NCAR's Geographic Information Systems (GIS) program

4.8 Sustainability Plan

The diverse portfolio of CI-WATER activities will be sustained by leveraging existing campus and statewide organizations for adoption and long-term support, stimulating development of new research proposals, and forming new partnerships. Furthermore, the adjacency of Utah and Wyoming and their shared demographic, climatic, hydrologic, and related environmental challenges provide an enhanced context for collaboration beyond the project term. The completion of two new, large data centers with research computing capabilities in both states – NWSC in Cheyenne with the supporting UWYO capabilities in Laramie and the Utah off-campus facility in Salt Lake City with optically linked campus capabilities at BYU, USU, and UU – also provides an excellent opportunity for robust, ongoing CI collaboration – especially for best practice sharing and resource optimization.

The UWYO is committed to the long-term sustainability of the proposed high-performance computing equipment. UWYO will provide appropriate space, power, connectivity, and additional staff for operation of the equipment. UWYO will operate the hardware and software beyond the end of the project, if necessary, through a combination of user fees, institutional support, and other mechanisms.

Similarly, through its Center for High Performance Computing (CHPC), the UU will commit to providing space, power, operational and end user support for the CI-WATER DATA STORE system over the course of its useful lifetime beyond the project term. CHPC provides large-scale computer and networking resources (i.e., hardware, software and expertise) to the UU computational research community and their collaborators. CHPC is led by Prof. Julio Facelli of Biomedical Informatics and Chemistry and has a staffing level of 28 FTEs and a base budget of over $2.5 million in addition to one-time capital infusions from both internal and external funding sources.

We have engaged five senior, CI-cognizant IT leaders (McDonald, Hawley, Corbató, Aylward, and Stewart) at the four collaborating universities and UEN to work with the CI Facilities project team. Beyond coordinating the necessary campus and regional network and other IT capabilities in support of this project, they will participate in the planning efforts for supporting new research directions and CI enhancements beyond the term and scope of the proposed work. Our objective is to establish long-term linkages between the CI-WATER faculty researchers and the corresponding campus IT leaders that will spur the development of new CI capabilities, both within and between the two states.

As a partner for both network connectivity and broader engagement in Utah, the state-funded UEN brings considerable expertise, resources, and over 30 years of experience working over a large state with geographically remote populations. UEN manages statewide telecommunications, internet, and educational systems to deliver applications and interactive services to Utah K-12 and higher education. UEN also recently was awarded a $13.6 million broadband development grant by the National Telecommunications and Information Administration to expand high-speed connectivity to additional community anchor institutions in Utah.

The sustainability of the software and related products developed for the new HPC-capable hydrologic model and the new class of hydrologic data modeling services will be achieved through (a) their ongoing use, enhancement, and further integration by CI-WATER scientists and their collaborators and (b) their broader distribution under open source licensing terms (e.g., the Berkeley Software Distribution license as currently used within CUAHASI HIS and potentially others as identified by the consortium).

To ensure the sustainability of our education and outreach efforts, we will leverage existing resources housed within UWYO's Water Resources Data System (WRDS) and its Wyoming State Climate Office (SCO) branch. WRDS has a 43-year history of working with Wyoming water issues, and it serves as the technical services branch of the Wyoming Water Development Office. WRDS and the SCO are key contributors to the state's water plan process, and they work closely with the Wyoming State Engineers Office on issues ranging from drought monitoring to the development of statewide hydroclimatic observing networks. As such, WRDS/SCO will connect CI-WATER activities and the two largest water management agencies in the state. WRDS/SCO maintains an extensive stakeholder database and is also tightly linked to the Wyoming Water Association, the largest water-related consortium in the state. In addition, the WRDS/SCO already has a full-time outreach coordinator and an extensive web and social media presence in the state. Finally, we will utilize UEN's own well-established web presence and other long-term public distribution channels to assure the long-term availability of the curriculum modules and associated media developed for this project.

A critical element of this project is the leveraging of external partnerships in all components to create a substantial basis for longer-term engagement. These partnerships are described in Letters of Commitment accompanying this proposal. This ensemble of partners includes for the CI component: the University Corporation for Atmospheric Research; for the modeling components: the Central Utah Water Conservancy District, the Division of Water Resources within the Utah Department of Natural Resources, the Colorado Basin River Forecast Center of the National Weather Service, and the Research Applications Laboratory at the National Center of Atmospheric Research; and for the engagement component, the Casper Science Zone, Wyoming PBS, and the Wyoming Science Posse.

4.9 Management and Coordination Plan

The organizational chart for the CI-WATER project is shown in Figure 4. The top-level oversight and administration will be provided by the Utah and Wyoming EPSCoR committees. The Utah EPSCoR Committee is composed of twelve members from the state legislature, state agencies, private industry, and universities. The four names listed in the Utah portion of the diagram represent the Executive Committee. Utah was recently granted EPSCoR status, and it is anticipated that a structured state EPSCoR office will be developed early in the timeline of this project. The Steering Committee is composed of the project PI and five other project members representing a mix of institutions and focus areas. Each member of the Steering Committee has extensive research and/or project management experience resulting in a diverse team with the skills and background necessary to manage a complex and multi-faceted project such as this. The Steering Committee will report to and coordinate with the Utah and Wyoming EPSCoR offices and will provide project leadership and administration. The researchers and staff involved with the project will be organized into four project teams, where each team is aligned with one of the four project components described in section 4.4. Each team includes members from each of the four participating institutions in a manner that will ensure cross-jurisdiction and cross-disciplinary scientific integration. The entire management structure will be subject to the multi-faceted evaluation and assessment program described in Section 4.7. Internal assessment will document indicators of project effectiveness and the independent evaluation program will help identify weaknesses and areas of concern that can be corrected by the Steering Committee.

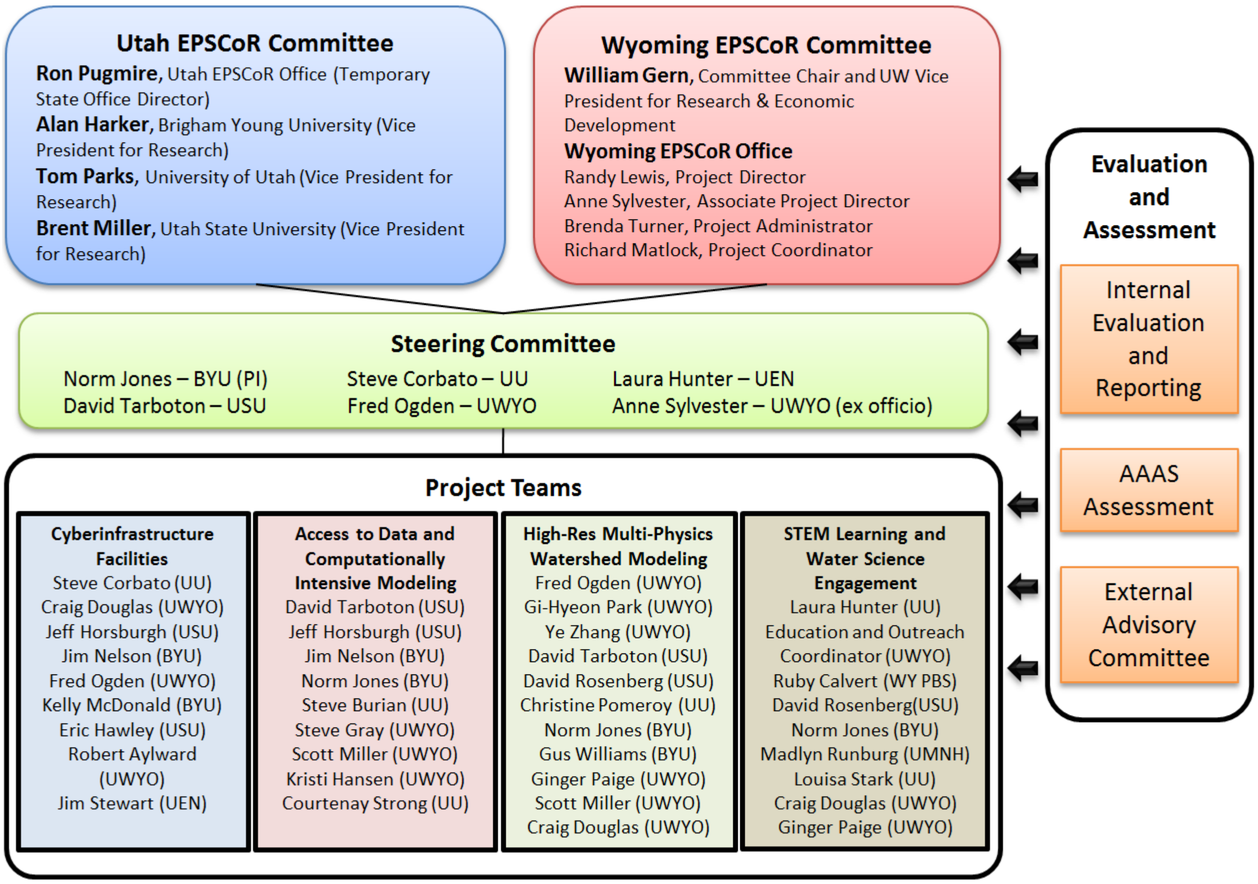

Figure 4. Comprehensive management and coordination plan.

Project Management: The CI-WATER project will be managed using the Agile (Scrum) method of project management [59]. The Agile method was originally developed for software development and is a widely-used and highly-efficient management approach. The main components and concepts of the Agile approach are as follows:

Teams. Each team in the project has a specific set of responsibilities for delivering products and services related to the project. The teams will be organized as shown in Figure 4.

Product Owners. The Product Owners are accountable for ensuring that the team activities are consistent with the overall project objectives. This role will be filled by the six members of the Steering Committee.

Product Backlog. The product owners (in consultation with the entire group) maintain a Product Backlog which is a prioritized list of products that need to be accomplished or developed by the group. Products can include software tools, plans, papers, hardware deployment, or other general tasks.

ScrumMaster. One person from each team will be designated as the ScrumMaster. This person acts as the team leader and is responsible for tracking the progress of the team through each of the development cycles.

Sprints. Each 30-day development cycle is called a Sprint. At the beginning of the Sprint, the teams will meet with the Product Owners and decide which products will be developed in the upcoming Sprint. This list of products is called the Sprint Backlog. This process takes into account the availability of the team members during the upcoming Sprint and an estimate of how much effort and resources will be required for each product. At the end of the Sprint, a Sprint Review Meeting is held where each team makes a presentation demonstrating the products developed during the Sprint. This meeting is typically followed by Spring Planning Meeting for the subsequent Sprint.

Scrums. Scrums are short meetings (conference calls, webinars) held frequently by teams during the Sprint. During the scrum, team members discuss progress on the Sprint Backlog and discuss problems and questions that have arisen that need to be resolved. ScrumMasters use feedback from the scrums to update a Sprint Burndown Chart illustrating the overall progress of the team toward achieving the monthly goal.

The Agile method will be advantageous for the project for several reasons. The monthly development cycles make it possible to track progress for long-term project goals and adjust priorities as needed. The more frequent scrum meetings allow for monitoring of short-term progress and provide a frequent level of exposure and accountability that promotes focus and productivity by team members. The Sprint Review Meetings make it possible for team members to demonstrate successes and progress and to get feedback from a larger pool of participants. The Review Meetings also give an opportunity for stake holders (Utah and Wyoming EPSCoR Committees) to monitor the progress of the project. The Sprint and Scrum meeting will provide ample opportunities for cross-jurisdiction and cross-disciplinary scientific integration.

Most Scrum meetings will be conducted via conference call. In some cases we will use webinars and/or video conferencing depending on the nature of the tasks being discussed. The monthly Sprint meetings will be conducted using webinars and/or video conferencing. In addition to the monthly meetings, we will hold annual review meetings at a central location where all team members will meet and review long-term progress and set goals for the coming year. These annual meetings will allow the state EPSCoR committees to review progress of the project and will provide an opportunity for review by the external evaluators discussed in Section 4.7.

Budget: The management and coordination plan will require resources that have been allocated in the project budget. The management costs are listed under the BYU budget and they will be administered by the project PI, Norm Jones. The monthly meetings will rely on low-cost web, audio, and video conferencing. We have allocated $6000 per year for the two-day annual meetings. This number includes facility costs, food, and document preparation. Travel costs for project members related to the annual meetings are itemized under each institution's budget. We have allocated $35,000 per year for the AAAS review based on quoted price from AAAS and $8800 for the four-person external advisory committee (travel plus $1000 stipend). With overhead, the management costs amount to $69,700, $57,200, and $57,200 for each of the three years.

Multi-Jurisdictional Strengths: Water scarcity in the western U.S., and worldwide, remains a significant scientific and societal challenge now and into the future. The CI-WATER Consortium offers solutions by using the power of CI to merge interdisciplinary research and education. Utah and Wyoming are uniquely situated to address water resource problems, and each jurisdiction brings unique expertise in modeling, CI expertise, resource management and STEM education. Together, these activities will accelerate our nation's ability to analyze, model, predict and offer water management solutions to an informed and CI-literate society.

References

1. Horsburgh, J.S., D.G. Tarboton, D.R. Maidment, and I. Zaslavsky, (2008), *A Relational Model for Environmental and Water Resources Data.* Water Resour. Res. **44**: p. W05406. doi:10.1029/2007WR006392.
2. Horsburgh, J.S., D.G. Tarboton, M. Piasecki, D.R. Maidment, I. Zaslavsky, D. Valentine, and T.Whitenack, (2009), *An integrated system for publishing environmental observations data.* Environmental Modelling & Software. **24**(8): p. 879-888. http://dx.doi.org/10.1016/j.envsoft.2009.01.002.
3. Tarboton, D.G., J.S. Horsburgh, D.R. Maidment, T. Whiteaker, I. Zaslavsky, M. Piasecki, J. Goodall, D. Valentine, and T. Whitenack. (2009), *Development of a Community Hydrologic Information System*. in *18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation*: Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, July 2009. 988-994. http://www.mssanz.org.au/modsim09/C4/tarboton_C4.pdf.
4. Horsburgh, J.S., D.G. Tarboton, K.A.T. Schreuders, D.R. Maidment, I. Zaslavsky, and D. Valentine. (2010), *Hydroserver: A Platform for Publishing Space-Time Hydrologic Datasets*. in *2010 AWRA Spring Specialty Conference Geographic Information Systems (GIS) and Water Resources VI*. Orlando Florida: American Water Resources Association, Middleburg, Virginia, TPS-10-1. http://www.awra.org/orlando2010/doc/abs/JefferyHorsburgh_7cb420e3_6602.pdf.
5. WATERS Network, (2008), *WATERS Test Bed Sites*. Accessed 2/18/2009. http://www.watersnet.org/wtbs/.
6. Zaslavsky, I., D. Valentine, and T. Whiteaker, (2007), *CUAHSI WaterML*, Open Geospatial Consortium Discussion Paper. http://portal.opengeospatial.org/files/?artifact_id=21743.
7. Ames, D.P., J. Horsburgh, J. Goodall, T. Whiteaker, D. Tarboton, and D. Maidment. (2009), *Introducing the Open Source CUAHSI Hydrologic Information System Desktop Application (HIS Desktop)*. in *18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation*: Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, July 2009. 4353-4359. http://www.mssanz.org.au/modsim09/J4/ames.pdf.
8. Horsburgh, J.S., D.G. Tarboton, D.R. Maidment, and I. Zaslavsky, (2011), *Components of an environmental observatory information system.* Computers & Geosciences. **37**(2): p. 207-218. http://dx.doi.org/10.1016/j.cageo.2010.07.003.
9. Horsburgh, J.S., A. Spackman Jones, D.K. Stevens, D.G. Tarboton, and N.O. Mesner, (2010), *A Sensor Network for High Frequency Estimation of Water Quality Constituent Fluxes Using Surrogates.* Environmental Modelling & Software. **25**(1031-1044). http://dx.doi.org/10.1016/j.envsoft.2009.10.012.
10. Jones, A.S., D.K. Stevens, J.S. Horsburgh, and N.O. Mesner, (2010), *Surrogate Measures for Providing High Frequency Estimates of Total Suspended Solids and Total Phosphorus Concentrations.* JAWRA Journal of the American Water Resources Association. http://dx.doi.org/10.1111/j.1752-1688.2010.00505.x.
11. Horsburgh, J.S., (2008), *Hydrologic Information Systems: Advancing Cyberinfrastructure for Environmental Observatories* in *Civil and Environmental Engineering*, Utah State University: Logan. p. 223.
12. National Research Council Committee on the Scientific Bases of Colorado River Basin Water Management, (2007), *Colorado River Basin Water Management: Evaluating and Adjusting to Hydroclimatic Variability*, National Academy Press: Washington, DC. http://books.nap.edu/catalog.php?record_id=11857.

13. A Plan for Wyoming Science and Technology, (2010). http://www.uwyo.edu/EPSCOR/_files/documents/WYScienceAndTechnologyPlan2010%20.pdf. Accessed 3/13/11.
14. Hey, T., S. Tansley, and K. Tolle, (2009), *The Fourth Paradigm, Data-Intensive Scientific Discovery*, Microsoft Research: Redmond, Washington. 283. http://research.microsoft.com/en-us/collaboration/fourthparadigm/.
15. Technical Working Group and Ad Hoc Committee for a Petascale Collaboratory for the Geosciences, (2005), *Establishing a Petascale Collaboratory for the Geosciences: Scientific Frontiers*, A Report to the Geosciences Community, UCAR/JOSS. p. 80.
16. Wilkins-Diehr, N., D. Gannon, G. Klimeck, S. Oster, and S. Pamidighantam, (2008), *TeraGrid Science Gateways and Their Impact on Science.* Computer. **41**(11): p. 32-41. http://dx.doi.org/10.1109/MC.2008.470
17. McLennan, M. and R. Kennell, (2010), *HUBzero: A Platform for Dissemination and Collaboration in Computational Science and Engineering.* Computing in Science & Engineering. **12**(2): p. 48-53. http://dx.doi.org/10.1109/MCSE.2010.41
18. Tarboton, D.G., D. Maidment, I. Zaslavsky, D.P. Ames, J. Goodall, and J.S. Horsburgh, (2010), *CUAHSI Hydrologic Information System 2010 Status Report*. http://his.cuahsi.org/documents/CUAHSIHIS2010StatusReport.pdf.
19. Harou, J.J., D. Pinte, A. Tilmant, D.E. Rosenberg, D.E. Rheinheimer, K. Hansen, P.M. Reed, A. Reynaud, J. Medellin-Azuara, M. Pulido-Velazquez, E. Matrosov, S. Padula, and T. Zhu. (2010), *An open-source model platform for water management that links models to a generic user-interface and data-manager*. in *2010 International Congress on Environmental Modelling and Software: 'Modelling for Environment's Sake', Fifth Biennial Meeting*. Ottawa, Canada: International Environmental Modelling and Software Society (iEMSs). http://www.hydroplatform.org/index.php/downloads/doc_download/9-iemss-2010-conference-proceedings-article.
20. Donatelli, M. and A.E. Rizzoli. (2008), *A Design for Framework-Independent Model Components of Biophysical Systems*. in *Proceedings of the iEMSs Fourth Biennial Meeting: International Congress on Environmental Modelling and Software (iEMSs 2008)*. International Environmental Modelling and Software Society, Barcelona, Catalonia. 727-734. http://www.iemss.org/iemss2008/.
21. Rizzoli, A.E., G. Leavesley, J.C. Ascough, R.M. Argent, I.N. Athanasiadis, V. Brilhante, F.H.A. Claeys, O. David, M. Donatelli, P. Gijsbers, D. Havlik, A. Kassahun, P. Krause, N.W.T. Quinn, H. Scholten, R.S. Sojda, and F. Villa, (2008), *Integrated modelling frameworks for environmental assessment and decision support*, in *Environmental modelling, software and decision support: state of the art and new perspectives*, A.J. Jakeman, A.A. Voinov, A.E. Rizzoli, and S.H. Chen, Editors, Elsevier Science: Amsterdam. p. 101-118. http://digitalcommons.unl.edu/usgsstaffpub/196/.
22. Bhatt, G., M. Kumar, and C. Duffy. (2008), *Bridging gap between geohydrologic data and Integrated Hydrologic Model: PIHMgis*. in *Proceedings of the iEMSs Fourth Biennial Meeting: International Congress on Environmental Modelling and Software (iEMSs 2008)*. International Environmental Modelling and Software Society, Barcelona, Catalonia. 743-750. http://www.iemss.org/iemss2008/.
23. Kumar, M., G. Bhatt, and C.J. Duffy, (2009), *An efficient domain decomposition framework for accurate representation of geodata in distributed hydrologic models.* International Journal of Geographical Information Science. **23**(12): p. 1569 - 1596. http://dx.doi.org/10.1080/13658810802344143
24. Gregersen, J.B., P.J.A. Gijsbers, and S.J.P. Westen, (2007), *OpenMI: Open Modelling Interface.* Journal of Hydroinformatics. **9**(3): p. 175-191. http://dx.doi.org/10.2166/hydro.2007.023.
25. Tarboton, D.G., K.A.T. Schreuders, D.W. Watson, and M.E. Baker. (2009), *Generalized terrain-based flow analysis of digital elevation models*. in *18th World IMACS Congress and MODSIM09*

International Congress on Modelling and Simulation: Modelling and Simulation Society of Australia and New Zealand and International Association for Mathematics and Computers in Simulation, July 2009. 2000-2006. http://www.mssanz.org.au/modsim09/F4/tarboton_F4.pdf.
26. Milly, P.C.D., K.A. Dunne, and A.V. Vecchia, (2005), *Global pattern of trends in streamflow and water availability in a changing climate.* Nature. **438**(17): p. 347-350. doi:10.1038/nature04312.
27. Christensen, N.S. and D.P. Lettenmaier, (2007), *A multimodel ensemble approach to assessment of climate change impacts on the hydrology and water resources of the Colorado River Basin.* Hydrol. Earth Syst. Sci. **11**(4): p. 1417-1434. http://www.hydrol-earth-syst-sci.net/11/1417/2007/.
28. Gleick, P.H., (2000), *Water: The Potential Consequences of Climate Variability and Change for the Water Resources of the United States, The Report of the Water Sector Assessment Team of the National Assessment of the Potential Consequences of Climate Variability and Change, For the U.S. Global Change Research Program*. Accessed 2/27/2011 http://www.gcrio.org/NationalAssessment/water/water.pdf.
29. Downer, C.W. and F.L. Ogden, (2004), *GSSHA: A model for simulating diverse streamflow generating processes.* J. Hydrol. Engrg. **9**(3): p. 161-174.
30. Qu, Y. and C.J. Duffy, (2007), *A semidiscrete finite volume formulation for multiprocess watershed simulation.* Water Resour. Res. **43**: p. W08419. doi:10.1029/2006WR005752.
31. Maxwell, R.M. and N.L. Miller, (2005), *Development of a Coupled Land Surface and Groundwater Model.* Journal of Hydrometeorology. **6**: p. 233-247.
32. Gao, H., Q. Tang, X. Shi, C. Zhu, T.J. Bohn, F. Su, J. Sheffield, M. Pan, D.P. Lettenmaier, and E.F. Wood, (2010), *Water Budget Record from Variable Infiltration Capacity (VIC) Model.* In Algorithm Theoretical Basis Document for Terrestrial Water Cycle Data Records (in review).
33. Zagona, E.A., T.J. Fulp, H.M. Goranflo, and R.M. Shane. (1998), *Riverware: A General River and Reservoir Modeling Environment*. in *Proceedings of the First Federal Interagency Hydrologic Modeling Conference*. Las Vegas, Nevada. 19-23 April. 5:113-120.
34. Bentz, B.J., J. Régnière, C.J. Fettig, E.M. Hansen, J.L. Hayes, J.A. Hicke, R.G. Kelsey, J.F. Negrón, and S.J. Seybold, (2010), *Climate Change and Bark Beetles of the Western United States and Canada: Direct and Indirect Effects.* BioScience. **60**(8): p. 602-613. http://www.bioone.org/doi/abs/10.1525/bio.2010.60.8.6.
35. Liu, J.S., (2008), *Monte Carlo Strategies in Scientific Computing*, Springer-Verlag: New York.
36. Vrugt, J.A., H.V. Gupta, W. Bouten, and S. Sorooshian, (2003), *A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters.* Water Resour. Res. **39**(8): p. 1201. http://dx.doi.org/10.1029/2002WR001642.
37. Evensen, G., (2009), *Data Assimilation: The Ensemble Kalman Filter*. 2nd ed, Springer-Verlag: Heidelberg.
38. Slater, A.G. and M.P. Clark, (2006), *Snow Data Assimilation via an Ensemble Kalman Filter.* Journal of Hydrometeorology. **7**(3): p. 478-493. http://journals.ametsoc.org/doi/abs/10.1175/JHM505.1.
39. Thompson, J.F., Z.U.A. Warsi, and C.W. Mastin, (1985), *Numerical Grid Generation: Foundations and Applications*, North-Holland: Amsterdam.
40. Thompson, J.F., B. Soni, and N.P. Weatherrill, (1998), *Handbook of Grid Generation*, CRC Press: Boca Raton.
41. Downer, C.W., F.L. Ogden, J.M. Niedzialek, and S. Liu, (2006), *Gridded Surface/Subsurface Hydrologic Analysis (GSSHA) Model: A Model for Simulating Diverse Streamflow Producing Processes*, in *Watershed Models*, V.P. Singh and D. Frevert, Editors, CRC Press. p. 131-159.
42. Efendiev, Y. and T. Hou, (2009), *Multiscale Finite Element Methods: Theory and Applications*, Springer-Verlag: New York.
43. Talbot, C.A. and F.L. Ogden, (2008), *A method for computing infiltration and redistribution in a discretized moisture content domain.* Water Resour. Res. **44**(8): p. W08453. http://dx.doi.org/10.1029/2008WR006815.

44. Ji, Z., (1998), *General hydrodynamic model for sewer/channel network systems.* J. Hydraulic Eng. **124**: p. 307-315.
45. Bales, R.C., N.P. Molotch, T.H. Painter, M.D. Dettinger, R. Rice, and J. Dozier, (2006), *Mountain hydrology of the western United States.* Water Resour. Res. **42**(8): p. W08432. http://dx.doi.org/10.1029/2005WR004387.
46. Liston, G.E. and K. Elder, (2006), *A Distributed Snow-Evolution Modeling System (SnowModel).* Journal of Hydrometeorology. **7**(6): p. 1259-1276. http://journals.ametsoc.org/doi/abs/10.1175/JHM548.1.
47. Liston, G.E., R.B. Haehnel, M. Sturm, C.A. Hiemstra, S. Berezovskaya, and R.D. Tabler, (2007), *Simulating complex snow distributions in windy environments using SnowTran-3D.* Journal of Glaciology. **53**: p. 241-256.
48. Storck, P., D.P. Lettenmaier, and S.M. Bolton, (2002), *Measurement of snow interception and canopy effects on snow accumulation and melt in a mountainous maritime climate, Oregon, United States.* Water Resour. Res. **38**(11): p. 1223. http://dx.doi.org/10.1029/2002WR001281.
49. Lehning, M., J. Doorschot, N. Raderschall, and P. Bartlet, (2000), *Combining snow drift and SNOWPACK models to estimate snow loading in avalanche slopes*, in *Snow Engineering*, E. Hjorth-Hansen, I. Holand, S. Laset, and H. Norem, Editors, Balkema: Rotterdam. p. 113-121.
50. Kumar, S.V., C.D. Peters-Lidard, Y. Tian, P.R. Houser, J. Geiger, S. Olden, L. Lighty, J.L. Eastman, B. Doty, P. Dirmeyer, J. Adams, K. Mitchell, E.F. Wood, and J. Sheffield, (2006), *Land information system: An interoperable framework for high resolution land surface modeling.* Environmental Modelling & Software. **21**(10): p. 1402-1415. http://dx.doi.org/10.1016/j.envsoft.2005.07.004.
51. Vrugt, J.A. and B.A. Robinson, (2007), *Treatment of uncertainty using ensemble methods: Comparison of sequential data assimilation and Bayesian model averaging.* Water Resour. Res. **43**(1): p. W01411. http://dx.doi.org/10.1029/2005WR004838.
52. Carriero, N. and D. Gelernter, (1990), *How to Write Parallel Programs: A First Course*, The MIT Press: Cambridge, MA.
53. Linda, *Scientific Computing Associates*. Accessed 3/8/2011: New Haven, CT. http://www.lindaspaces.com.
54. Charm++, *Parallel Computing Laboratory, UIUC*. Accessed 3/8/2011: Champaign, IL. http://charm.cs.uiuc.edu.
55. Committee on Prospering in the Global Economy of the 21st Century: An Agenda for American Science and Technology, (2007), *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*, National Academy of Sciences, National Academy of Engineering, Institute of Medicine http://www.nap.edu/catalog.php?record_id=11463.
56. Utah Society for Environmental Education, (2009), *Utah Environmental Literacy Plan*, brochure, USEE, Utah.
57. Public Broadcasting Service, (2010), *Roper Public Opinion Poll Results*. Accessed 3/13/11. http://www.pbs.org/roperpoll2010/PBS_Roper_brochure_2.18.10.pdf.
58. U S Census Bureau. Accessed 3/13/11. http://quickfacts.census.gov/qfd/states/56000.html.
59. Schwaber, K., (2004), *Agile Project Management with Scrum*, Microsoft Press: Redmond, WA. 192.