Improving the Discovery, Access, and Use of Water Data for Basin-Scale River Management

Phase 2 Workshop Report

October 2021

Dan Reicher[1], Jack Seagrist[1], Rahul Dhakal[1], Kyle Larson[2], Brian Bellgraph[2], Bo Saulsbury[2], Jerry Tagestad[2], Corrine Deciampa[2], Erin McCann[2], Debjani Singh[3], Tom Ruggles[3], Cuihong Song[4], Weiwei Mo[4], David Hart[5], and Betsy Barber[5]

[1] Stanford Woods Institute for the Environment, Stanford, CA

[2] Pacific Northwest National Laboratory, Richland, WA

[3] Oak Ridge National Laboratory, Oak Ridge, TN

[4] University of New Hampshire, Durham, NH

[5] University of Maine, Orono, ME

Note: Author order by institution

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor Battelle Memorial Institute, nor any of their employees, makes **any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights**. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or Battelle Memorial Institute. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Table of Contents

Executive Summary

The project assessed the value of accessible and usable water data for advancing river basin management in the context of five river basins across the U.S. This work was conducted by three universities (Stanford, University of New Hampshire, University of Maine) and two DOE National Laboratories (Pacific Northwest and Oak Ridge) with the support of the U.S. Department of Energy's (DOE) Water Power Technologies Office (WPTO). The project was part of the U.S.–China Clean Energy Research Center for Water-Energy Technologies (CERC-WET) initiative of the Department of Energy, which is a collaboration with the following institutions: University of California – Berkeley, University of California – Irvine, Lawrence Berkeley National Laboratory, University of California – Davis, University of California – Merced, University of California – Los Angeles, and the Stockholm Environment Institute.

The goal of the project was to provide WPTO guidance on how it may facilitate improving the discovery, sharing, and use of water data. Extensive literature research and stakeholder interviews were conducted in five river basins across the United States, including the Eel and Russian river basins in California, Wisconsin River basin in Wisconsin, and the Connecticut and Penobscot river basins in the New England region (Figure 1). The information collected was used to inform the development of data stories about each basin (see *Workshop Structure and Themes*) that illustrate the variety and accessibility of water-related data and its importance to basin-scale river management. A fifth data story about the HydroSource data platform, stewarded by Oak Ridge National Laboratory, was also written. HydroSource was developed to help inform hydropower-related decision making and research in the U.S.
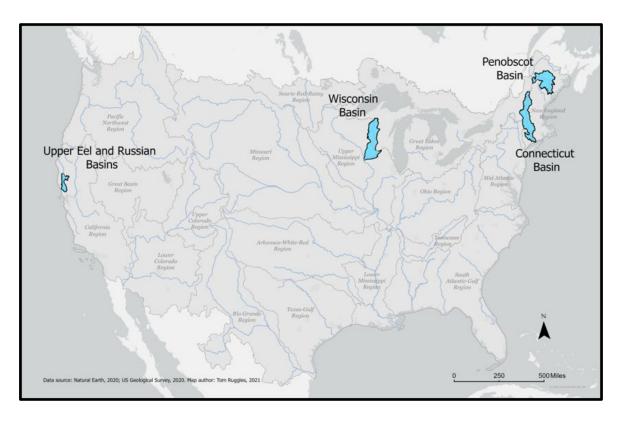
Figure 1 - Locations of Case Study U.S. River Basins.

The project also held a virtual workshop on May 7th, 2021 where the team shared key findings and collected input from approximately 70 water management experts and stakeholders from federal and state agencies, non-government organizations, academia, and industry from across the U.S.

The project team began the workshop with discussion of the opportunities and challenges of basin-scale river management and three key themes about water data that emerged from development of the data stories. These three themes include:

- **Variety** – Information about water quantity and quality are essential, but most water-related issues require a wider variety of data to make informed decisions about complex issues.
- **Accessibility** – The accessibility of water data has greatly improved but significant technical and socioeconomic challenges remain. Better access to information helps streamline decision-making and enable more equitable participation in river management.
- **Modernization** – Efforts to improve access to water data, and ultimately river management, must be done in ways that are consistent with modern information systems and ways in which people access data.

The second major component of the workshop was breakout discussions with participants about challenges and benefits of improving access to water data and its role in facilitating basin-scale river management. Participants were also asked to comment on actions that DOE and the broader water community can take to facilitate basin-scale river management and improve access to water data. Below is a summary of the key findings for each question revealed during the breakout discussions and panel. More detailed analyses of the findings are provided later in the report.

1. *What challenges do you see in improving access to water data and its role in facilitating basin-scale river management?*

   Four primary challenges were identified: security and trust concerns; funding issues; equitable access to data; and centralizing data from disparate sources.

   The first challenge area, security and trust concern, refers to the lack of trust present among stakeholders at the local and state levels and limited access to proprietary water data, preventing collaboration and equitable involvement in the decision-making process. Disparate levels of funding across basins were also identified as a challenge area that limits data sharing and inhibits better decision-making processes. Another challenge area that was identified was equitable access to data, which refers to incorporating under-served communities which lack technical and financial resources in water data improvement efforts. Finally, centralizing data from disparate sources, refers to the challenges that exist with aggregating the large quantity of water data supplied from a variety of different sources in basins across the U.S.

2. *What benefits do you see in improving access to water data to address the challenges in basin-scale river management?*

   Three opportunity areas were identified for the benefits associated with improving access to water data: aggregation of data for easier access; reduced stakeholder conflicts through interest optimization; and better system/infrastructure management.

   The first opportunity area, aggregation of data for easier access, describes the benefits that aggregation efforts for water data could bring to the decision-making process. Reducing stakeholder conflicts through interest optimization refers to the ability to reach optimal project decisions, which maximize benefits for all parties, due to improved access to water data. The third opportunity area identified was improved system/infrastructure management. This refers to the improved ability to manage assets with state-of-the-art technology and more accessible, usable water data.

3. *What are the most important actions that DOE and the broader water community can take to facilitate basin-scale planning and improving access to water data?*

Four action areas were identified: incentives; partnerships; science and technology; and funding and technical support.

The first action area, incentives, refers to programs that provide incentives to the larger water management community that allow them to adopt or prioritize basin-scale river management approaches. The second action area, partnerships, refers to the need for creating and strengthening partnerships across the diverse sets of stakeholders, at all levels, to improve basin-scale river management. Science and technology were identified as the third action area. This refers to the need for using the latest advances in science and information technology across basins to create more effective data platforms for better river management. Funding and technical support was identified as the fourth action area and refers to the need to increase the financial resources being provided to water data. It was even suggested that it be considered in federal infrastructure investments.

Workshop Structure, Themes and Data Stories

As part of a series of investigations with funding from WPTO, a 3-hour virtual workshop was conducted on May 7, 2021, to: 1) publicize a collection of "data stories" that illustrate the variety and importance of data relevant to river management; 2) identify key challenges, opportunities, and actions to help inform future DOE and other public-private initiatives related to basin-scale management and accessibility of water data; and 3) assess community views regarding discoverability, accessibility, and usability of water data.

The workshop was facilitated by the Stanford Woods Institute for the Environment and moderated by Dan Reicher (Stanford). A total of 70 people representing federal and state agencies, academia, nongovernmental organizations (NGOs), the hydropower industry, and engineering consulting companies, participated in the workshop. The workshop also included a panel discussion featuring Jaime Pinkham (Principal Deputy Assistant Secretary of the Army for Civil Works), Alejandro Moreno (Deputy Assistant Secretary for Renewable Power in the DOE Office of Energy Efficiency and Renewable Energy), and Tom Kiernan (President and CEO of American Rivers). The workshop agenda was structured as follows:

Table 1 - Workshop Agenda

Workshop Agenda	
12:00 – 12:45pm	Introduction
12:45 – 1:45pm	Breakout discussion
1:45 – 2:00pm	Break
2:00 – 2:45pm	Panel discussion
2:45 – 3:00pm	Synthesis
(All times are in Eastern Daylight Time)	

"Water data" is defined as data and information about water quantity, quality, and other water-dependent factors such as aquatic ecosystem health, socioeconomics, recreation, and cultural practices. Improving the discoverability, accessibility, and usability of water data is important to addressing a variety of river management challenges, including basin-scale solutions and coordination, diverse stakeholder interests, increased uncertainty in the natural and human systems, intensified environmental, economic, and social tradeoffs, and trust and relationship building.

In preparation for the workshop, a series of data stories illustrating the variety and accessibility of water-related data and its importance to basin-scale river management were developed. The data stories were focused on real-world management issues in four select U.S. river basins where basin-scale approaches have been successful or are in use, including Penobscot River (Maine), Connecticut River (New England), Eel and Russian Rivers (California), and Wisconsin River (Wisconsin). A fifth data story that describes WPTO's platform for U.S. hydropower data, HydroSource, was also produced.

Table 2 - Web Links to Data Stories

Story Short Name	Permanent URL
Upper Eel and Russian River Basins	https://arcg.is/1TCS0H0
Penobscot River Basin	https://arcg.is/54iGz
Connecticut River Basin	https://arcg.is/1HqTyi0
Wisconsin River Basin	https://arcg.is/1yevDi
HydroSource	https://arcg.is/1q4CGb

Three common themes emerged from the development of data stories that were used to help focus workshop dialogue. These themes include: Data Variety; Data Accessibility; and Data Modernization. The following describes each theme in more detail.

Data Variety. Many water-related issues require data about a variety of topics (e.g., ecology, hydrography, water quality, climate, recreation/aesthetics, and socioeconomics) to facilitate informed decision-making. Water data are disparate in nature, meaning they are usually collected and distributed by different entities. In addition, water data are often collected through different methods in various formats, metadata standards, and spatiotemporal scales, making it challenging to use and

discover the data. These factors can lead to concerns about the credibility or usability of data. Inclusion of a wide variety of water data can help us understand diverse stakeholder interests and resources at stake and tackle complex river management issues more effectively. In addition, developing data models or centralized data infrastructure may facilitate the process of harmonizing data formats and standards and the process of data modernization.

Data Accessibility. Data may become hidden away in portals, come in obscure formats, or come with inaccurate or insufficient descriptions. This can constrain effective communication with decision-makers and stakeholders. Better access to existing information can help streamline decision-making and enable more equitable participation in river management. This is challenging, however, because there are often poor incentives for data sharing and some data are too sensitive to share openly or without significant use restrictions. Data sharing is enhanced by productive relationships based on trust and common ground, which often take a long time to build and can remain fragile. Duplicate data collection and dissemination efforts sometimes occur due to a lack of coordination.

While the challenges are vast, so are the opportunities. For example, improving data access can provide economic benefits in the long-term by reducing time and resources spent on data aggregation and facilitating better informed outcomes. Improved access to data also can help level the playing field among stakeholders, leading to more equitable outcomes. Additionally, efforts to improve access to data can foster efforts to be more transparent and reduce conflict among stakeholders. Efforts to improve interorganizational coordination on data collection and dissemination, stakeholder engagement, and consensus building are also needed. The latter two may be facilitated via applying social sciences methods such as negotiation simulation, scenario analysis, and participatory modeling.

Data Modernization. The Digital Age has transformed how we discover, share, and use data. However, digitalization does not guarantee it to be accessible or usable. With the modernization of water data comes new considerations, such as the Five V's of Big Data (Volume, Velocity, Variety, Veracity, and Value). Modernizing water data can provide more effective means to assess how the value can be improved, however, it remains a significant challenge. The variety of water data creates additional challenges for creating interoperability of data between systems and software. This becomes even more challenging as data structures and standards are updated or improved. The ongoing drive to modernize water data makes this an opportune time to create or improve centralized data systems. This can take shape in the form of new Information Technology (IT) platforms for sharing data, or platforms for connected distributed data systems. Opportunities to boost funding and resources aimed at modernizing water data

systems also should be considered. Finally, seeking opportunities to collaborate with the tech industry may help drive modernization at a faster pace, especially for larger initiatives seeking to tie together many disparate sources of water data.

Breakout Discussion

The breakout session was attended by approximately 70 participants who were divided into six breakout groups. Chatham house rules were followed during the workshop to encourage open dialogue; thus, participant's answers are anonymous. Participants were asked to respond to two questions regarding challenges and benefits to improving access to water data and its role in facilitating basin-scale river management. Participants were also asked to comment on a third question regarding actions that can be taken to address the challenges, which is summarized in the following *Workshop Synthesis* section.

Question 1: What __challenges__ do you see in improving access to water data and its role in facilitating basin-scale river management?

There were several commonalities in the challenges discussed among the breakout groups, including: security and trust concerns; funding issues; equitable access to data; and centralizing data from disparate sources. Each of these common areas are discussed in more detail below.

- **Security and trust concerns:** Participants suggested that to improve access to proprietary or otherwise sensitive water data, novel approaches are needed to ensure it is secure while also more accessible and helpful to stakeholders. There are strong concerns about lack of trust among stakeholders, which can inhibit collaboration and equitable involvement in decision-making processes. Some participants also felt that more effort is needed to strengthen trust among stakeholders at local and state levels than at the federal level.

- **Funding issues:** Participants felt that disparity in funding across basins limits data sharing, and often there is no funding available to bring multiple stakeholders together for working towards a better decision-making process.

- **Equitable access to data:** Some participants indicated that efforts to improve access to water data should also consider how to improve access for under-served communities, non-English speaking communities, or stakeholder groups who lack technical and financial resources and are commonly under-represented in decision-making processes.

- **Centralizing disparate sources of data:** Participants expressed that certain categories of water data have become much more accessible from a variety of sources; however, navigating through the abundance of data sources demands a lot of time and skills to get actionable insights. Developing more centralized data systems that can aid in the discovery and distillation process were discussed as a possible solution to this problem. One example that was discussed was the Internet of Water project, which seeks to develop information architecture that will help better connect water data sources to end users. More centralized data systems could also lessen the divide between the science and non-science

communities. Participants also expressed a need to train people to be more data literate and proficient with information technology.

*Question 2: What **benefits** do you see from improving access to water data to address the challenges in basin-scale river management? What are some real-world examples of the benefits?*

There were several commonalities in the benefits discussed among the breakout groups, including: aggregation of data for easier access; reduce stakeholder conflicts; and better system/infrastructure management. Each of these common areas are discussed in more detail below.

- **Aggregation of data:** Some participants suggested that, in certain cases, access to data should be improved by aggregating disparate sources into a centralized location. California's Open Water Data Act AB1755 was cited as an example where the state government has taken a lead role in aggregating its water data. Currently, projects like the Internet of Water and data platforms like ORNL's HydroSource, US Army Corps's National Inventory of Dams, USGS's National Water Dashboard and Streamstats, and America's Watershed Initiatives are good examples of the benefits of large-scale data aggregation initiatives. Projects and platforms like these could also improve coherence of multiple datasets by helping users gain access to information that improves their understanding of the complex coupling between human and environmental systems.

- **Reduce stakeholder conflicts**: River management requires multi-stakeholder collaboration; thus, better access to data could improve efficiency of decision-making processes by leveling the field of information which stakeholders have access to. This could help reduce conflicts, especially between states (and within states' agencies).

- **Better system/infrastructure management:** A paradigm shift is occurring where water managers and stakeholders are seeking ways to manage rivers as integrated systems rather than a collection of independent projects. With better access to water data, these parties can more effectively explore complicated tradeoffs associated with such systems-based management decisions. Participants also noted that improved access to data is needed to acquire the volume of data required for more advanced analytical methods (e.g., machine learning, Artificial Intelligence) that may help to answer complex, system-scale questions.

Workshop Synthesis

Participants were asked during breakout discussions what actions DOE and the broader water community could take to facilitate basin-scale planning and improve access to water data. Their suggestions were analyzed for commonalities and summarized at the conclusion of the workshop. Suggested actions were grouped into four general categories: Incentives, Partnerships, Science and Technology, and Funding. Each of these four action areas is described in greater detail below.

Incentives

Workshop participants generally agreed there are benefits to basin-scale approaches to river management and improving access to water-related data, although some suggested that additional actions are needed to further incentivize the water management community to pursue these objectives. Efforts to incentivize the community should be conducted in both top-down (government) and bottom-up (grassroots) manners. For example, some participants indicated that initiatives aimed at improving access to water-related information are consistent with the President Biden's Administration's emphasis on environmental justice and social equity and should be tied to those policies to provide greater top-down influence. To strengthen these ties, future initiatives should seek to identify underserved communities and disadvantaged stakeholder groups within local, regional, and national water management forums and improve the information needs of those groups.

Bottom-up approaches to incentivizing the water management community to adopt basin-scale management principles and improving access to water-related information were frequently cited by workshop participants. In general, these approaches involve studying and publicizing real-world examples of how adaptive management and information sharing have improved local discourse about water management. Participants also suggested there be broader examination of the environmental and economic tradeoffs of dam removal, rehabilitation, and retrofitting, and the data used to inform and evaluate those decisions.

Partnerships

Many participants expressed the need to build partnerships among stakeholder groups, industries, non-government organizations (NGOs), and state and federal agencies, that are based, in part, on making data more accessible and facilitating basin-scale river management. Such partnerships should also be rooted in addressing challenges such as improving inter-organizational trust and coordination, co-leveraging financial and institutional resources, and improving cross-domain transfer of knowledge. Workshop participants also cited the importance of partnerships to resolving some of the more complex challenges facing water management, most notably better understanding the range of uncertainties posed by climate change on the nation's water resources.

Identifying and promoting successful examples of unlikely partnerships, such as those achieved in the Penobscot River Restoration Project in Maine or the recent joint

[statement](#) of collaboration by leading members of the U.S. hydropower industry and environmental conservation community, was discussed as an action that DOE and the larger water management community could continue. The two-year process through which the joint statement of collaboration was reached, known as an Uncommon Dialogue, led by Stanford Woods Institute of the Environment, was cited as a potential model for partnership building among other stakeholder groups in the water management community.

Science and Technology

Workshop participants indicated that science and technology are critical aspects of improving the accessibility of water data and more effective river management. Suggested actions regarding technology commonly centered around making data more discoverable. This includes a range of actions such as developing better ontologies (i.e., formal definitions of the categories, properties and relations between datasets or information systems) to improve standardization and develop a common vocabulary, developing scalable data platforms aimed at connecting users to distributed data systems, and better leveraging the power of cloud computing and Artificial Intelligence. Additional technological actions that were suggested focused on improving data availability (i.e., filling categorical, spatial, and temporal data gaps). For example, developing new technologies or approaches to measure water quantity and quality where traditional methods are not feasible.

Some participants pointed to the needs and practices of the scientific community as a key driver for improving access to water data and river management. The scientific community is one of the largest consumers and producers of water-related data. The community also has become increasingly focused on interdisciplinary approaches and Big Data to solving complex water management problems. These factors have helped to motivate efforts to improve access and use of water-related data both within and among different scientific fields. A notable example is the [Consortium of Universities for the Advancement of Hydrologic Science, Inc](#). (CUAHSI), which is working to create a shared scientific infrastructure (data, models, and education) aimed at understanding the interactions between water, earth, ecosystems, and society.

Funding and Technical Support

Workshop participants indicated that more financial and technical support is needed to expedite and expand efforts to improve access to water data, particularly at state and local levels where these resources are more limited. Some participants even argued that certain categories of water data are crucial to managing the nation's water resources, and therefore should be considered part of our national infrastructure and related investments. Key areas where financial resources are needed include reducing the cost burden of modernizing legacy information systems and data warehouses, facilitating collaboration with the technology sector, participating in cost-sharing activities among stakeholders, and improving coordination of data collection and dissemination by federal and state agencies.