

California Data Collaborative in Brief

A new **California Data Collaborative** effort is currently underway involving water agencies throughout California to collect and share data among utilities to improve efficiencies, refine demand management strategies and promote long-term sustainable solutions across California's natural resources. This initiative aims to utilize "Big Data," an approach that has proven successful in the private and other public sectors, to answer the call for improvements in California's water data infrastructure from leading water and technology experts. The first phase of this pilot project focuses on integrating customer-level water usage data across participating utilities to develop more robust measurements of water usage behavior across California and illustrate how this effort can scale statewide.

The critical role of customer-level usage data

California currently collects a variety of reported utility level water usage metrics through the State Water Resources Control Board conservation program (monthly averages since 2013), Department of Water Resources Urban Water Management Plan program (annual averages last updated in 2010), and the Clean Drinking Water program. Ultimately, however, water usage decisions are made at the customer level, and such granular data is necessary to effectively manage water demand across California. Customer level usage data enables analysts to dig deeper into water usage trends by customer class, geographic area, hydrologic features and other unique local characteristics. In addition, customer level usage data across utilities supports more comprehensive research into what conservation actions work in what contexts to effectively manage demand.

The need for inter-utility collaboration in understanding water usage behavior

Consider the recently published study evaluating the impact of turf rebates in Moulton Niguel Water District. This study supplements the existing literature by examining water usage behavior as a function of demographic characteristics and water usage brackets, and by investigating conservation program participation behavior. But with data from only one utility, we are unable to answer: what would have happened if Moulton had implemented the turf rebates differently during that time? Yet in South Orange County, other districts have historically offered different rebate amounts, creating a reference point that analysts can use to understand the drivers of turf rebate program participation and scope next steps to achieve turf market transformation. Other subsets of California's 411 other major urban water retailers offer different conservation programs, pricing structures, and marketing approaches to a variety of customer populations. Putting all of California's customer level usage data together with key contextual information in a centralized, secure cloud platform can accelerate the understanding of the unique characteristics of communities throughout California and lead to an effective, customized approach to local demand management that have statewide impacts.

Integrating the entire lifecycle of CA's water usage data

California water retailers employ a wide variety of data management and customer classification practices, but every urban metered water connection in California collects three key data points: 1) the amount of water used 2) the address where that water was used and 3) the time period over which that water was used. Those geographic and time identifiers enable the integration of contextual demographic, weather, evapotranspiration and other data sources such as irrigable area derived from aerial or satellite imagery. That data can be integrated with utility operational data (e.g. SCADA, GIS asset maps) to support system efficiencies like leak detection and improving operations through demand forecasting. Centralizing that data in cloud infrastructure enables the water community to better leverage modern data science tools



regularly employed by private technology companies and increasingly in other industries such as energy, agriculture, retail, healthcare, and financial services.

Pioneering new data infrastructure so California can adapt to whatever the future holds

Using computer programming scripts to integrate water usage data with important contextual information provides two key advantages over current manual approaches: better information and less time/cost to the agency. Computers can repeatedly run routine tasks and regularly update the data over secure web connections allowing for more current analysis that can aid in water management decisions. Water managers need timely, comprehensive measurements to accelerate how we achieve water efficiency and adapt to California's challenging water supply realities.

Contact Information

Patrick Atwater | patrickatwater@gmail.com
Project Manager, California Data Collaborative Phase One Pilot

Project Supporters and Collaborators

Andrew J. Marx, Ph.D, Assistant Professor of the Center for Information Science and Technology
Claremont Graduate University

Roger Bales, Distinguished Professor of Engineering, Director
UC Water Security & Sustainability Research Initiative
Sierra Nevada Research Institute, UC Merced

Neil Fromer, Executive Director for the Resnick Sustainability Institute
Applied Physics & Materials Science, California Institute of Technology

Galib Rustamov, PhD. The Center for Behavioral and Economic Studies
Division of Politics and Economics, Claremont Graduate University

Richard Atwater, Executive Director
Southern California Water Committee

Greg Webber, Executive Director
California Urban Water Conservation Council

Wesley Hartman, Professor of Marketing
Graduate School of Business, Stanford University

Ken Baerenklau, Associate Professor of Environmental Economics & Policy
School of Public Policy, UC Riverside

Kurt Schwabe, Professor of Environmental Economics & Policy
School of Public Policy, UC Riverside