

Advancing Hydrological Modeling and Water Resources Management with Machine Learning

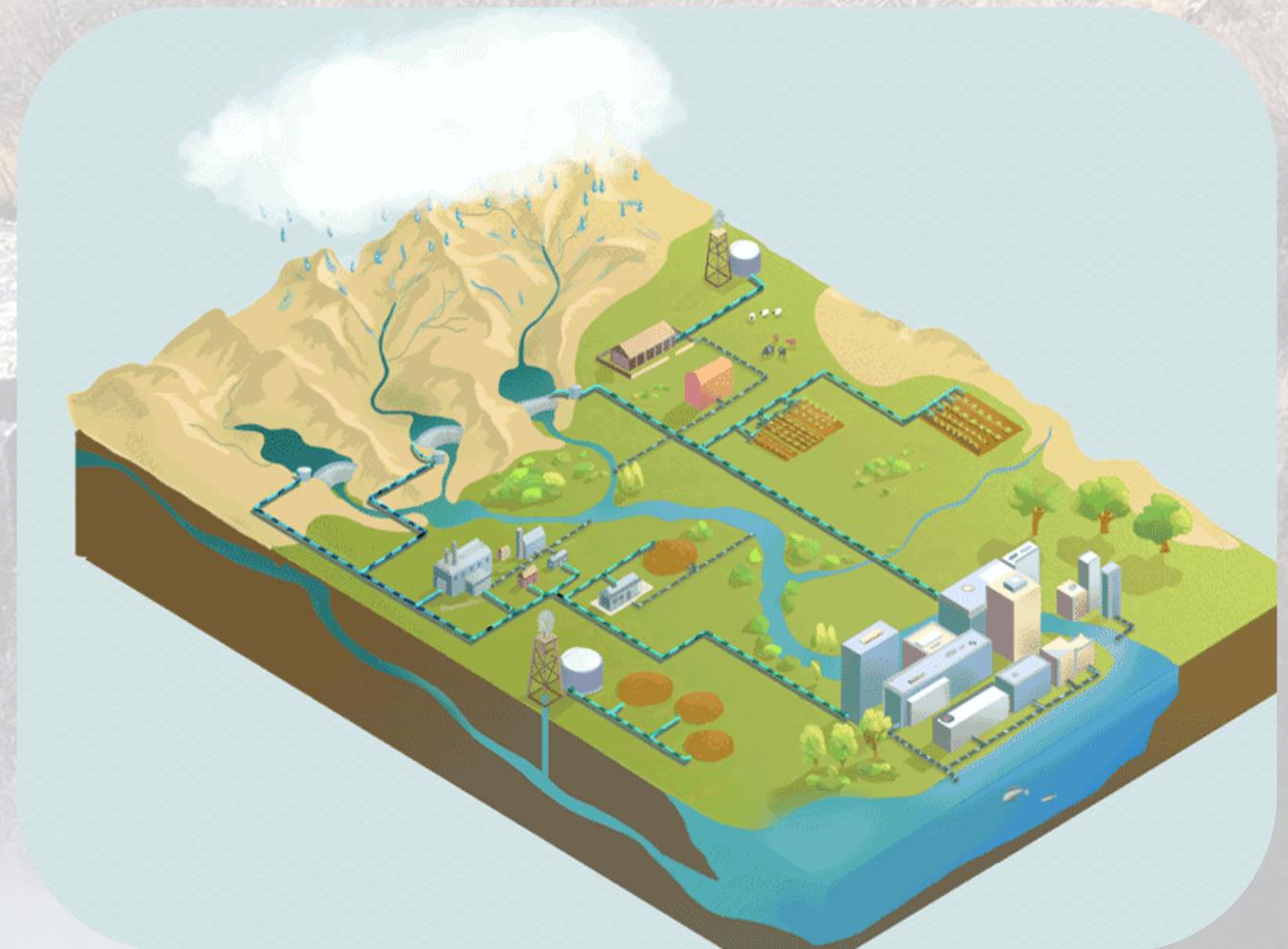
Dr. Ryan Johnson
Research Scientist
The University of Alabama



whitelightning450/Presentation_Materials



College of
Engineering
Civil, Construction and Environmental Engineering



Presentation Outline

- Introduction
- Research Background
- Previous Accomplishments
- Active Areas of Research
- Research Program at the University
- Teaching Goals, Philosophy, Ideas



whitelightning450/Presentation_Materials

2

Dr. Ryan Johnson

Academic Background

- Environmental Sciences
- Earth System Science
- Civil and Environmental Engineering
- Water Resources

“The richness I achieve comes from Nature, the source of my inspiration” – Claude Monet



Water Quality

Climate Change



Urban Water Systems



Water Resources

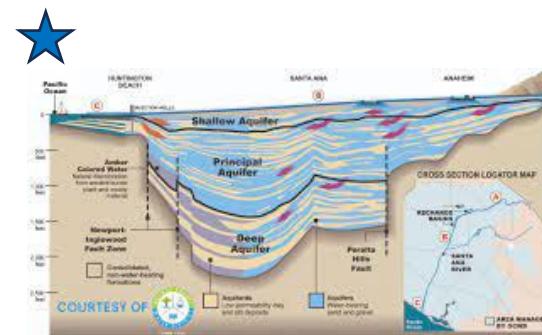
Background: Drivers of Water Research



Environmental Degradation
and Social Inequality



Water Quality

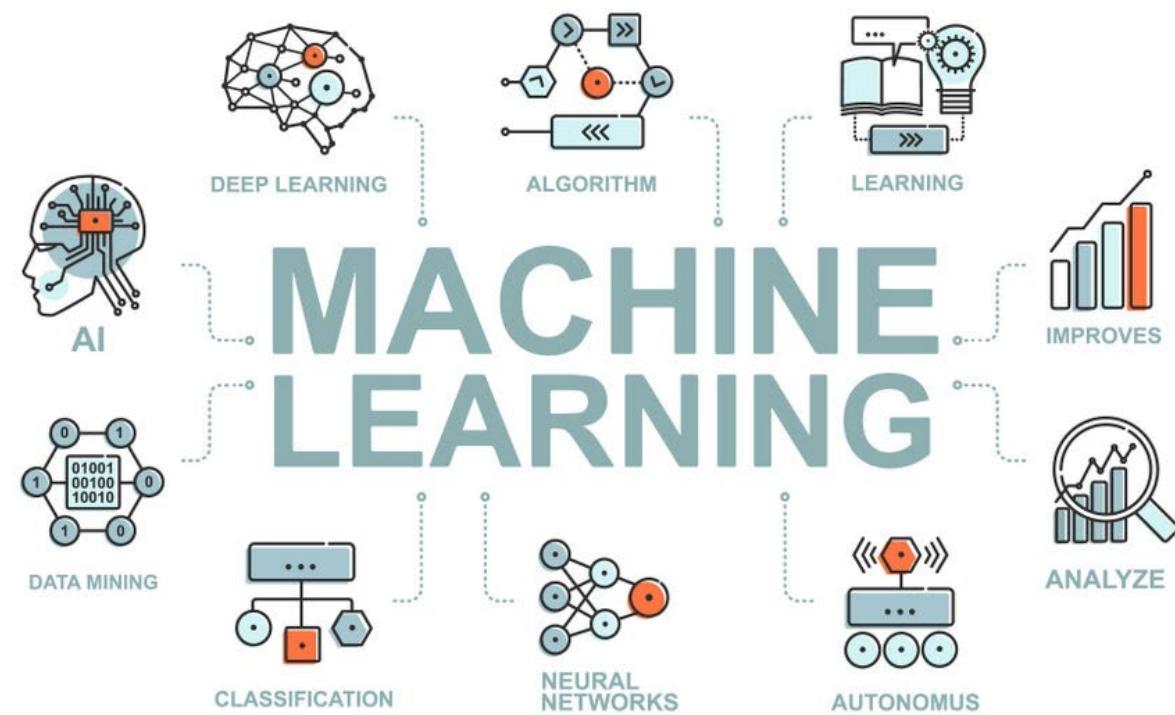


Enhancing Water Resources



Environmental Management

My Water Research



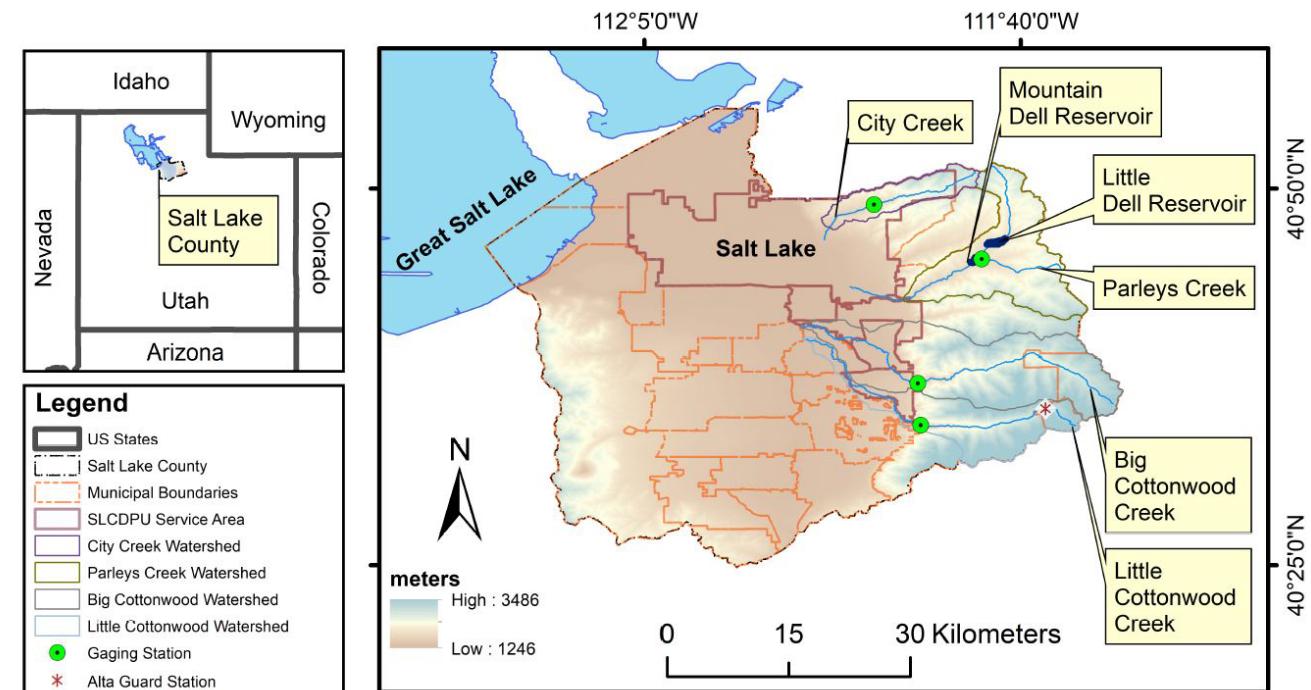
Previous Research



Urban Water System Modeling

Urban Water System Modeling

- ML to model water demand
- Climate influences on water system performance
- ML for water system modeling



Objective: Advance the science in water system planning in a dynamic and arid climate.



ML to estimate water demand

Research Question: Can machine learning and feature engineering produce measurable improvements in water demand estimates compared to traditional methods?

Objective: Identify indicators/drivers of season-to-season and year-to-year variability in water demand, and integrate them in machine learning workflows to improve prediction skill.

Received: 29 September 2022 | Accepted: 16 November 2023

DOI: 10.1111/1752-1688.13186

RESEARCH ARTICLE



Data-driven modeling to enhance municipal water demand estimates in response to dynamic climate conditions

Ryan C. Johnson¹  | Steven J. Burian² | Carlos A. Oroza³ | Carly Hansen⁴ |
Emily Baur³ | Danyal Aziz²  | Daniyal Hassan⁵  | Tracie Kirkham⁶ | Jessie Stewart⁶ |
Laura Briefer⁶

ML to estimate water demand

Methods:

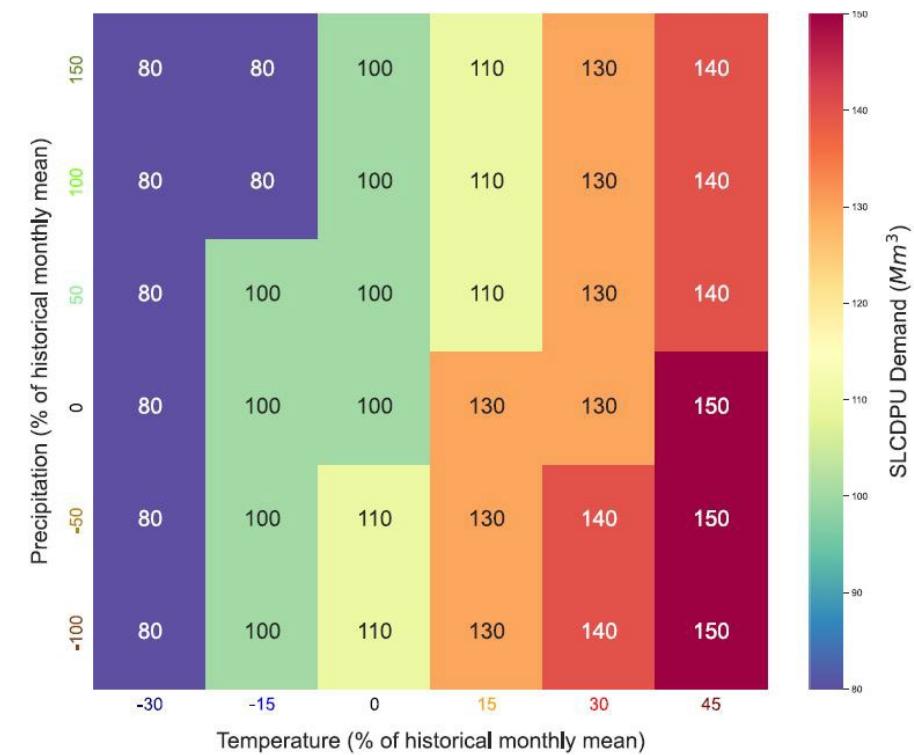
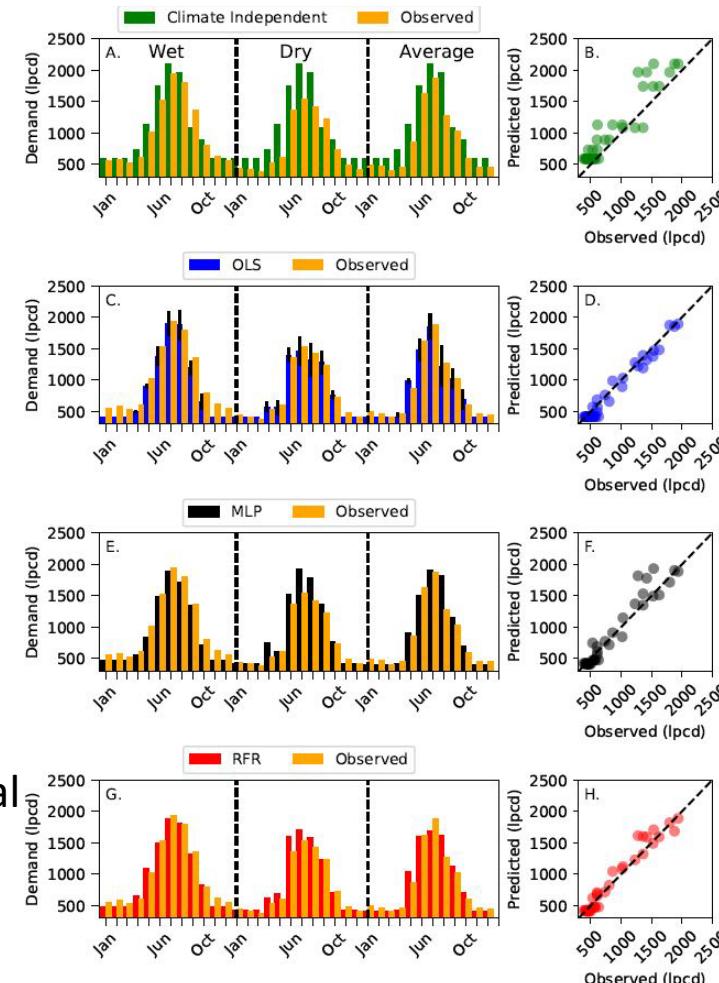
- Developed multiple ML algorithms
- Monthly precipitation and temperature inputs
- Population density for change over time
- Dry, average, and wet conditions

Results:

- 35% reduction in error during drought conditions
- All ML models display greater skill than conventional methods

Impact:

- 1st demonstration of ML for seasonal water demand estimation
- ML can improve seasonal water demand estimates
- Created operational tools for SLCDPU



Machine Learning for Water System Modeling

Research Question: Many utilities do not have the resources or personnel to develop a representative water system model to support planning and management activities. Can machine learning accurately model water system interactions and performance across different supply and demand scenarios?

Objective: Use simulated water system data to train machine learning models to predict daily to seasonal groundwater, reservoir levels, and imported water. Compare model skill to simulated targets.



© 2023 The Authors



Journal of Hydroinformatics Vol 25 No 5, 1584 doi: 10.2166/hydro.2023.170

Data-driven modeling of municipal water system responses to hydroclimate extremes

Ryan Johnson *, Steven John Burian , Carlos Anthony Oroza , James Halgren^a, Trevor Irons , Danyal Aziz , Daniyal Hassan , Jiada Li , Carly Hansen , Tracie Kirkham^g, Jesse Stewart^g and Laura Brieferg

^aAlabama Water Institute, University of Alabama, Tuscaloosa, Alabama, USA

^bCivil, Construction and Environmental Engineering, University of Alabama, Tuscaloosa, Alabama, USA

^cCivil and Environmental Engineering, University of Utah, Salt Lake City, Utah, USA

^dMontana Technical University, Butte, Montana, USA

^eCivil and Environmental Engineering, Colorado State University, Fort Collins, Colorado, USA

^fOak Ridge National Laboratory, Oak Ridge, Tennessee, USA

^gSalt Lake City Department of Public Utilities, Salt Lake City, Utah, USA

*Corresponding author. E-mail: rjohnson18@ua.edu

RJ, 0000-0003-2195-739X; SJB, 0000-0002-5278-3188; CAO, 0000-0001-5522-7665; TI, 0000-0003-3547-9341; DA, 0000-0002-0734-6894;
DH, 0000-0001-8812-7230; JL, 0000-0002-7238-5101; CH, 0000-0002-1288-2632

Machine Learning for Water System Modeling

Methods:

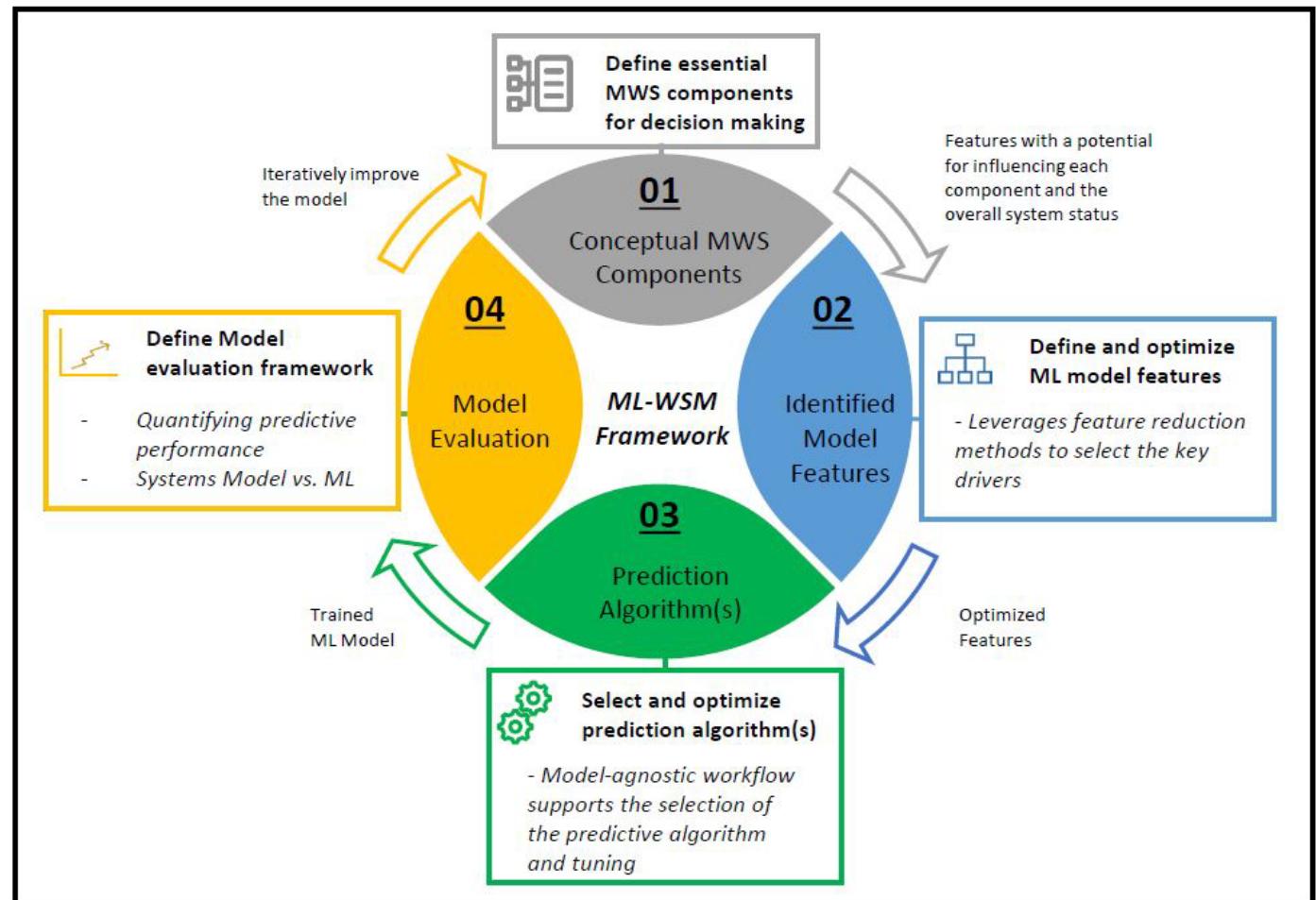
- Model key features of water system
- Incorporate serial correlation
- Model agnostic workflow
- System agnostic
- Dry, wet, average scenarios

Results:

- High model skill during critical conditions
- Representative vulnerability metrics

Impact:

- Demonstration of ML algorithms for modeling water system interactions
- Expedited development
- Applicable to smaller water systems and/or complementary to others



Research Impact

Advanced Water Resources Planning and Management



Department of
ATMOSPHERIC SCIENCES
MINES AND EARTH SCIENCES | THE UNIVERSITY OF UTAH



THE UNIVERSITY OF UTAH
COLLEGE OF SOCIAL AND BEHAVIORAL SCIENCES
Department of Geography



Prelim estimates of streamflow yield

January

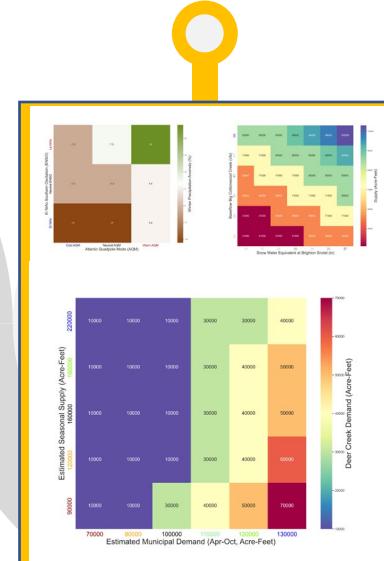


October

Seasonal water system projections

Winter precipitation estimates

March

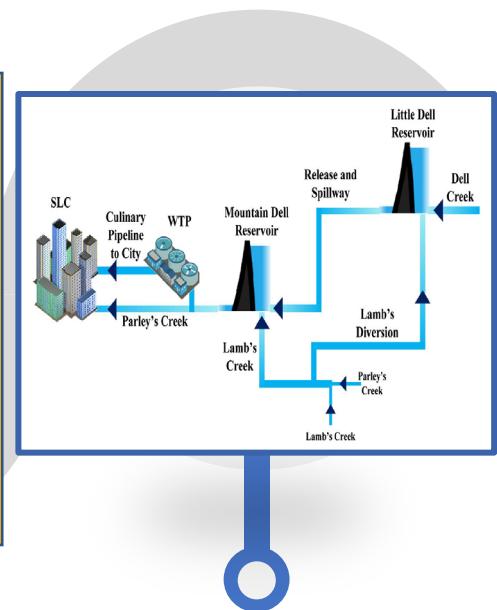


February

April-June

Assess system status
Revise/Initiate operational decisions

Summer precip & temp outlook
Refined estimates of surface water yield
Refined estimates of system vulnerabilities
Begin formulating management decisions



Other: Previous research

Water Resources Research

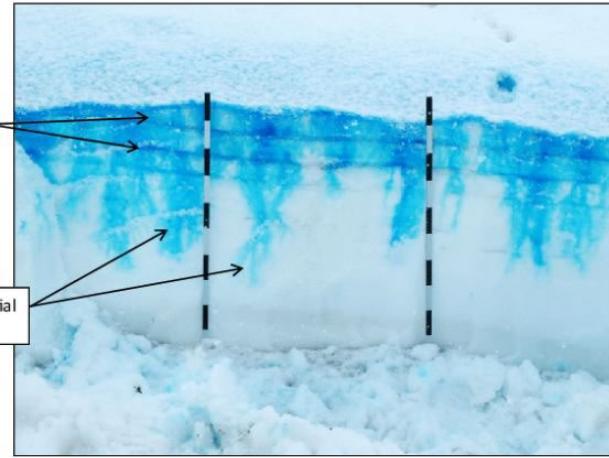
RESEARCH ARTICLE
10.1029/2019WR024828

Special Section:
Big Data & Machine Learning in
Water Sciences: Recent Progress
and Their Use in Advancing
Science

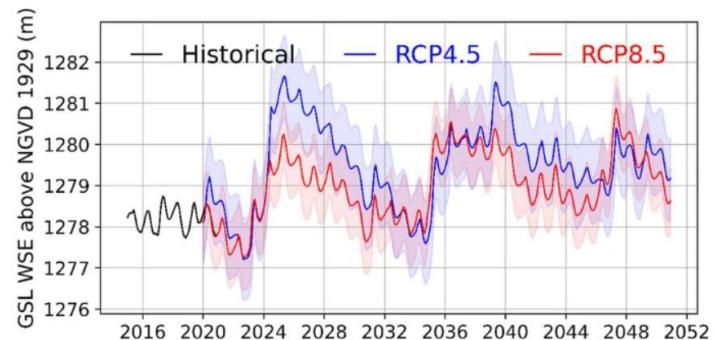
Insights Into Preferential Flow Snowpack Runoff Using Random Forest

Francesco Avanzi¹, Ryan Curtis Johnson², Carlos A. Oroza², Hiroyuki Hirashima³,
Tessa Maurer¹, and Satoru Yamaguchi³

¹Department of Civil and Environmental Engineering, University of California, Berkeley, CA, USA, ²Department of Civil and Environmental Engineering, University of Utah, Salt Lake City, UT, USA, ³Snow and Ice Research Center, National Research Institute for Earth Science and Disaster Resilience, Nagaoka, Japan



Lazzaro et al., 2015



Water Resources Management (2023) 37:2697–2720
<https://doi.org/10.1007/s11269-022-03376-x>



The Great Salt Lake Water Level is Becoming Less Resilient to Climate Change

Daniyal Hassan¹ · Steven J. Burian^{2,3} · Ryan C. Johnson³ · Sangmin Shin⁴ ·
Michael E. Barber¹

Received: 10 August 2022 / Accepted: 1 November 2022 / Published online: 21 December 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022



Current Research Activities

Cooperative Institute for Research to Operations in Hydrology (CIROH)

- Advancing CONUS-Scale Operational Snow Modeling Capabilities
 - Snow Water Equivalent Machine Learning (SWEML)
- Community Streamflow Evaluation System (CSES)
- Enhancing Supply Forecasting for Systems Management
 - National Water Model Season-to-Season Water Supply Forecasting
- Advancing Snow Observation Systems to Improve Operational Streamflow Prediction Capabilities
 - Low-Cost Low-Power Snow Observing Systems



Advancing CONUS-Scale Operational Snow Modeling Capabilities



PI: Ryan Johnson, PHD – University of Alabama

Co-PI: Andy Wood , PHD – Colorado School of Mines

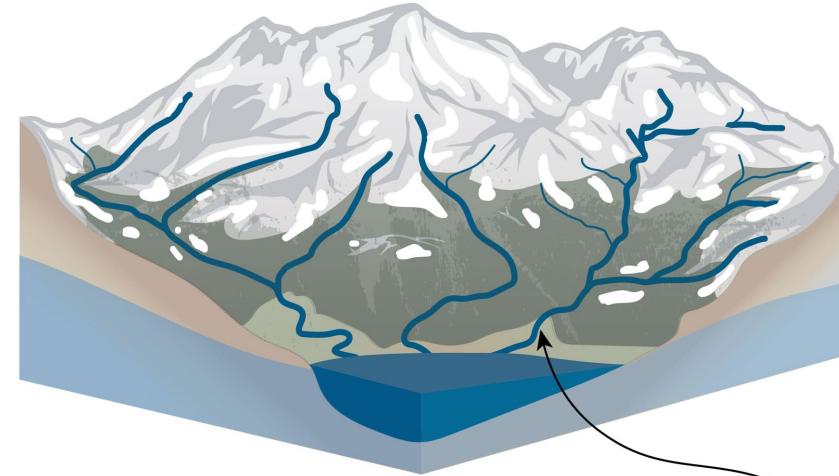
Co-PI: Katherine Hale , PHD - University of Vermont

Co-PI: McKenzie Skiles , PHD - University of Utah



Snow Water Equivalent Machine Learning (SWEML)

Spatially continuous near-real-time SWE estimates



Soil Saturation


SWE

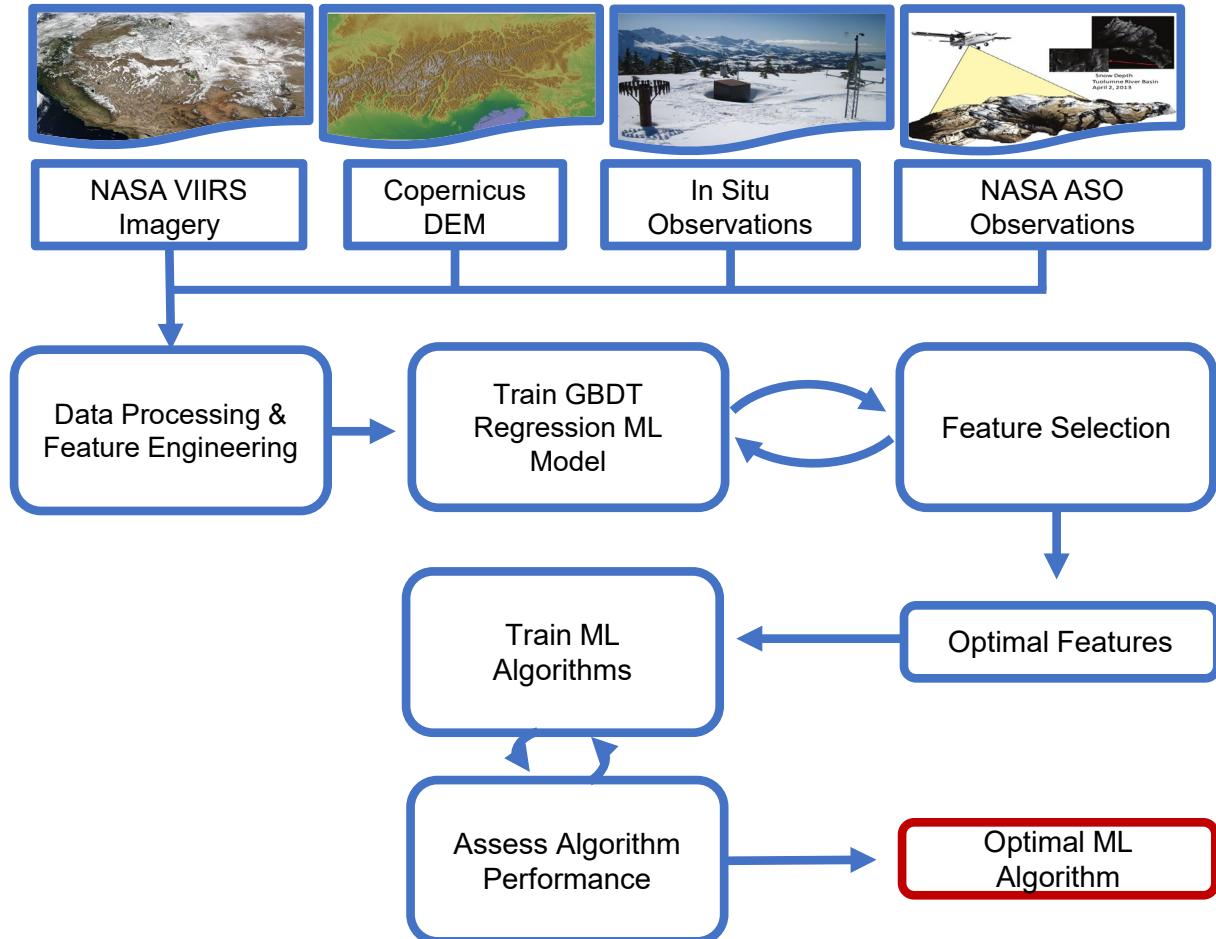
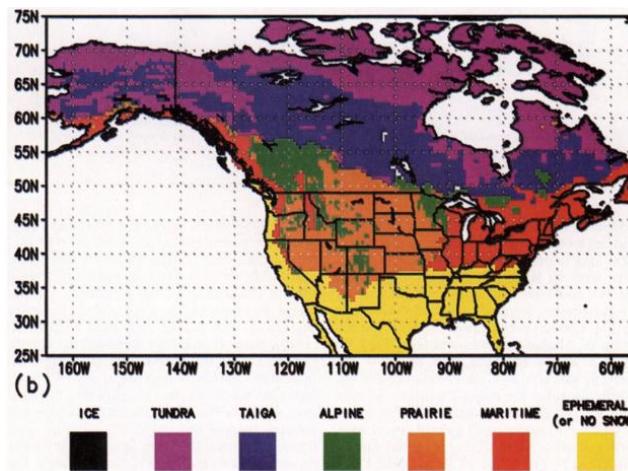

Melt Timing


Flow

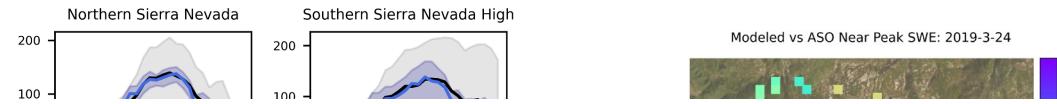
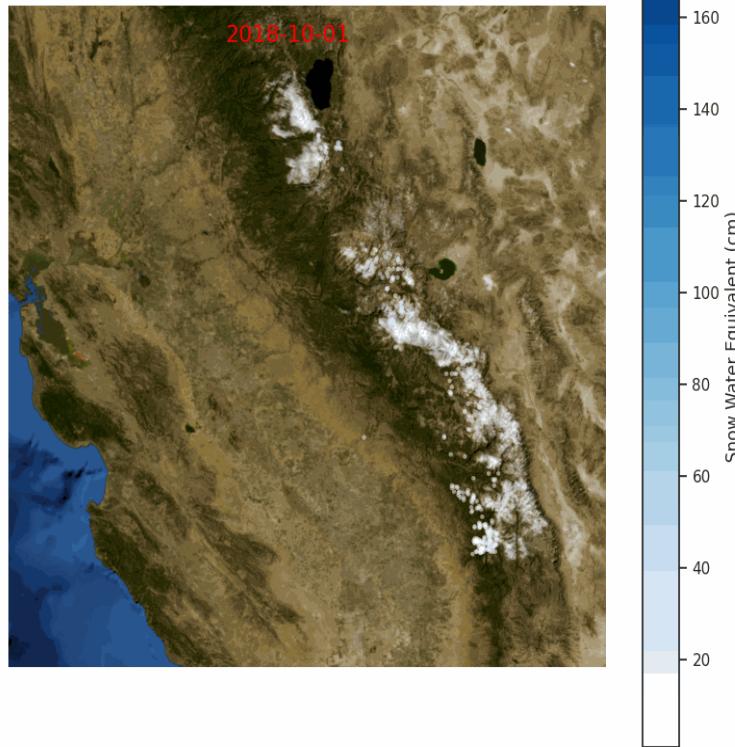

Snow Water Equivalent Machine Learning (SWEML)

Model Development

- Airborne Snow Observatory and snow course observations
- 1km spatial resolution
- Regionalization via Sturm's classification
- Multiple ML algorithms explored
- Examined 3 feature selection methods



Snow Water Equivalent Machine Learning (SWEML)



Quantifying Regional Variability of Machine-Learning-Based Snow Water Equivalent Estimates Across The Western United States

Dane Liljestrand^a, Ryan Johnson^b, S. McKenzie Skiles^c, Steven Burian^d and Josh Christensen^e

^aDepartment of Civil and Environmental Engineering, University of Utah, Salt Lake City, 84112, UT, USA

^bAlabama Water Institute, University of Alabama, Cyber Hall 1046, Tuscaloosa, 35487, AL, USA

^cDepartment of Geography, University of Utah, Salt Lake City, 84112, UT, USA

^dDepartment of Civil and Construction Engineering, University of Alabama, 3043 H.M. Comer 245 7th Avenue, Tuscaloosa, 35487, AL, USA

^eComputer Science Department, Brigham Young University, Provo, 84602, UT, USA

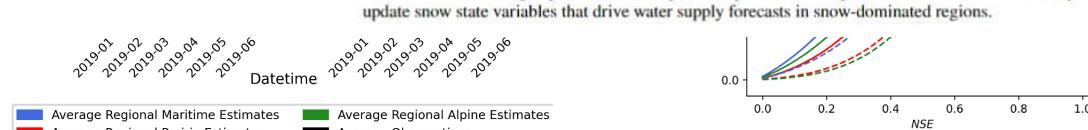
ARTICLE INFO

Keywords:

Snow Water Equivalent
Machine Learning
Recursive Feature Engineering
Artificial Neural Network
Gradient Boosting Decision Trees

ABSTRACT

Seasonal snow-derived water is a critical component of the water supply in the mountains and downstream regions, and the accurate characterization of available water in the form of snow-water-equivalent (SWE), peak SWE, and snowmelt onset are essential inputs for water management efforts. Arising from recent advancements in artificial intelligence (AI) and machine learning (ML), we introduce a large-scale ML SWE model leveraging publicly available data sources and open-source software. The model demonstrates the application of a limited feature space in a relatively simple ML architecture without the need for process-based formulations to effectively estimate spatially continuous SWE at a daily temporal resolution. Beginning with in situ SWE measurements (i.e., SNOTEL), lidar-derived terrain features, and temporal variables, we employ localized feature engineering and optimization via gradient-boosting decision trees to identify regionally unique drivers of snowpack dynamics and use the optimal features to train regionally independent artificial neural networks to estimate regional SWE at a 1-km spatial resolution. The model results yield respectable skill in reconstructed 1-km gridded SWE magnitudes in a hindcast simulation of the 2019 water year that is independent of the training and testing data. Comparing model estimates to over 6200 observations, the model demonstrates a weighted RMSE of 15.4-cm, Kling-Gupta Efficiency metric of 0.86, and a percent bias of 0.71% across 23 snow-influenced regions in the western U.S. The model simulation produces peak SWE estimates within 10-cm for twenty of the twenty-three regions, demonstrating capability in effectively capturing regional snow accumulation processes. The demonstration of low-error ML workflows capable of providing near-real-time, spatially continuous SWE estimates at a high spatial resolution provides proof-of-concept and a foundation to effectively update snow state variables that drive water supply forecasts in snow-dominated regions.





Hydroinformatics: Community Streamflow Evaluation System: CSES



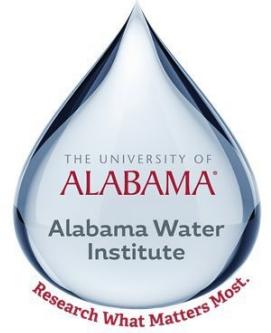
PI: Shahabul Alam, PHD
CO-PI: Ryan Johnson, PHD

Support:

- Nathan Swain, PHD
- Giovanni Romero



AQUAVEO™



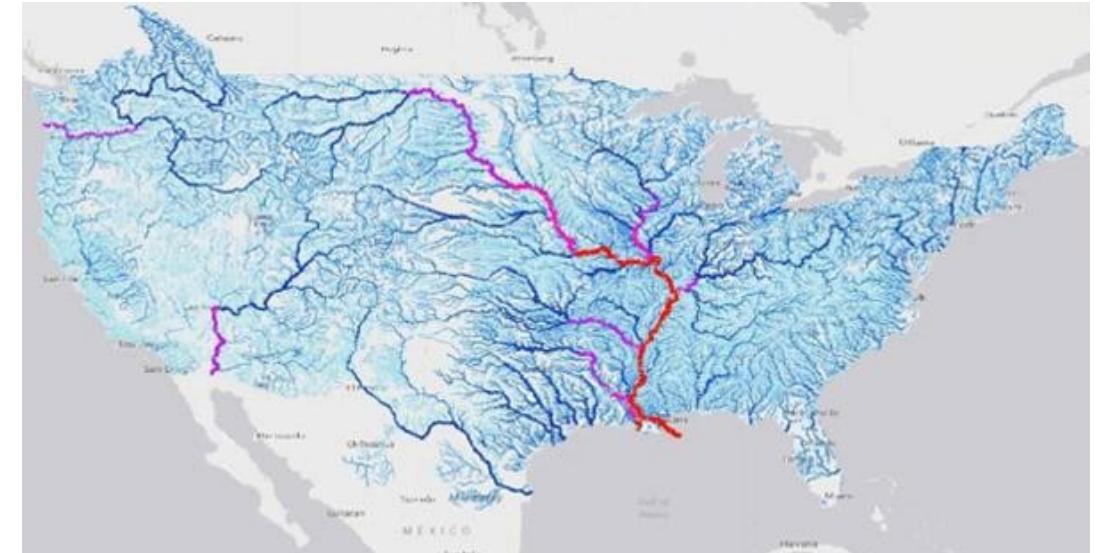


Community Streamflow Evaluation System: CSES



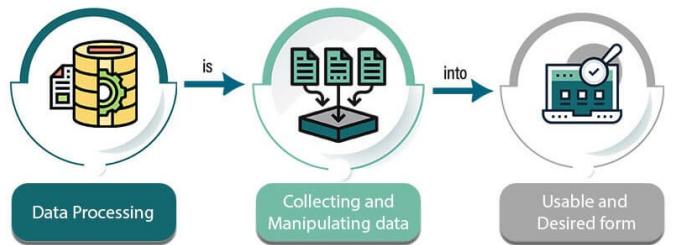
Objectives:

- Timely evaluations of National Water Model
- Model agnostic structure
- Simple data model
- Publicly accessible
- General to advanced users
- Open-Source
- CONUS scale





Community Streamflow Evaluation System: CSES



EDUCBA



A	B
1 Datetime	USGS_flow
2 10/1/2007	0.06236111
3 10/2/2007	0.060937498
4 10/3/2007	0.06
5 10/4/2007	0.06791667
6 10/5/2007	0.106458336
7 10/6/2007	0.10329165
8 10/7/2007	0.085520834
9 10/8/2007	0.080104165
10 10/9/2007	0.088854164
11 10/10/2007	0.12875
12 10/11/2007	0.0803125
13 10/12/2007	0.0584375
14 10/13/2007	0.044166666
15 10/14/2007	0.044791665
16 10/15/2007	0.060729165
17 10/16/2007	0.063333333
18 10/17/2007	0.069270834
19 10/18/2007	0.07729167
20 10/19/2007	0.34375
21 10/20/2007	0.2875
22 10/21/2007	0.13729167
23 10/22/2007	0.12916666
24 10/23/2007	0.29770833
25 10/24/2007	0.54300004
26 10/25/2007	5.883421
27 10/26/2007	9.178437
28 10/27/2007	5.9686456
29 10/28/2007	3.759375
30 10/29/2007	2.7270834
31 10/30/2007	2.318125

Community Streamflow Evaluation System (CSES)

Community Streamflow Evaluation System (CSES) is a novel system to evaluate hydrological model performance using a standardized NHDPlus data model. CSES evaluates modeled streamflow to a repository of over 5,000 in situ USGS monitoring sites, with interactive visualizations supporting an in-depth analysis.

The system includes:

- README
- Code of conduct
- MIT license
- Security

Metrics:

$$NSE = \frac{(\bar{Q}_{obs} - \bar{Q}_{sim})^2}{(\bar{Q}_{obs} - \bar{Q}_{mean})^2}$$
$$KGE = 1 - \sqrt{\frac{(\bar{Q}_{obs} - \bar{Q}_{sim})^2 + (\frac{1}{\bar{Q}_{obs}} - \frac{1}{\bar{Q}_{sim}})^2 + (r - 1)^2}{3}}$$
$$R = \frac{\sum_{i=1}^n (Q_{obs,i} - \bar{Q}_{obs})(Q_{sim,i} - \bar{Q}_{sim})}{\sqrt{\sum_{i=1}^n (Q_{obs,i} - \bar{Q}_{obs})^2} \sqrt{\sum_{i=1}^n (Q_{sim,i} - \bar{Q}_{sim})^2}}$$
$$Bias(\%) = \frac{\sum_{i=1}^n (Q_{obs,i} - Q_{sim,i})}{\sum_{i=1}^n Q_{obs,i}} \times 100$$
$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (R_{Si} - R_{Gi})^2}$$
$$MAE = \frac{1}{n} \sum_{i=1}^n |(R_{Si} - R_{Gi})|$$

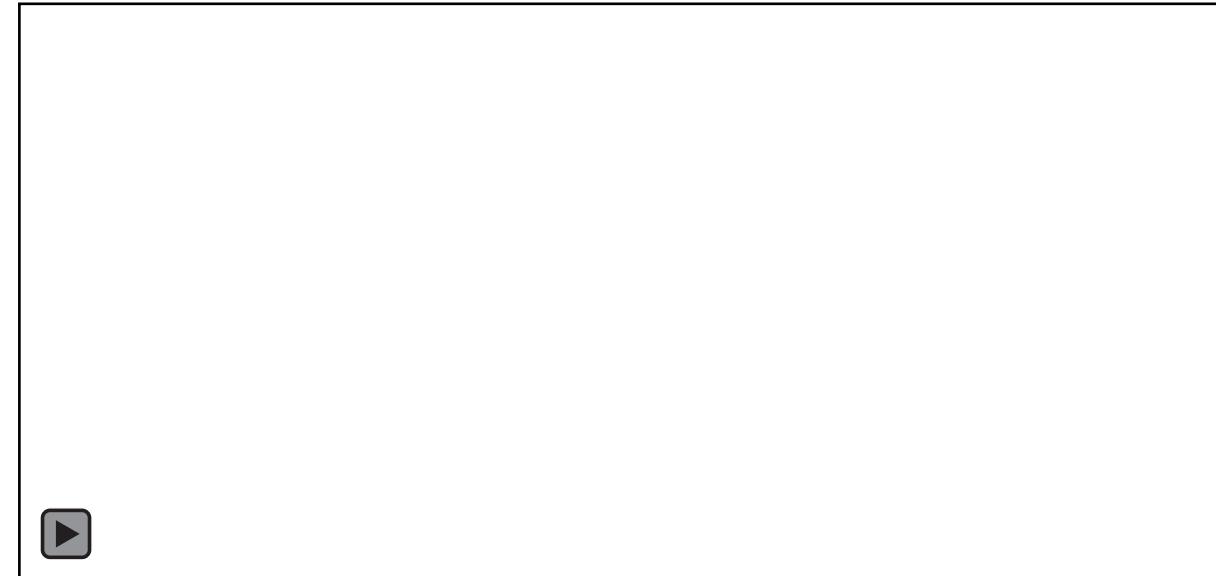
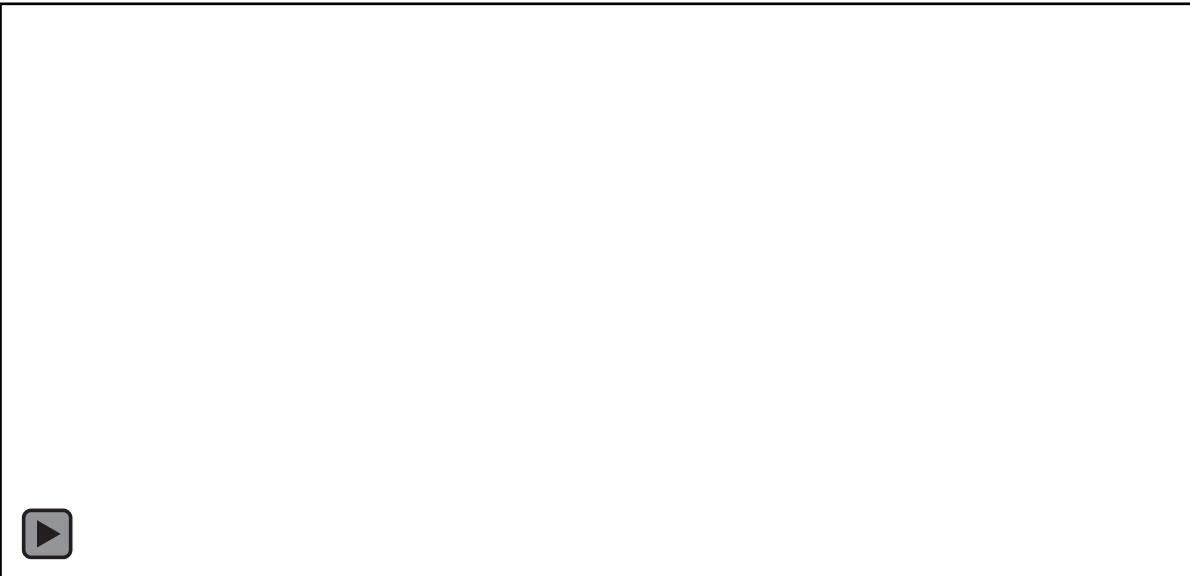



Community Streamflow Evaluation System: CSES



Tethys web-based platform

Jupyter-based platform





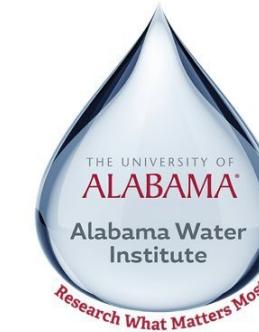
Enhancing Water Supply Forecasting for Systems Management



PI: Steven Burian, PHD
CO-PI: Ryan Johnson, PHD

Researchers:

- Shahabul Alam, PHD
- Savalan Neisary, PHD Student

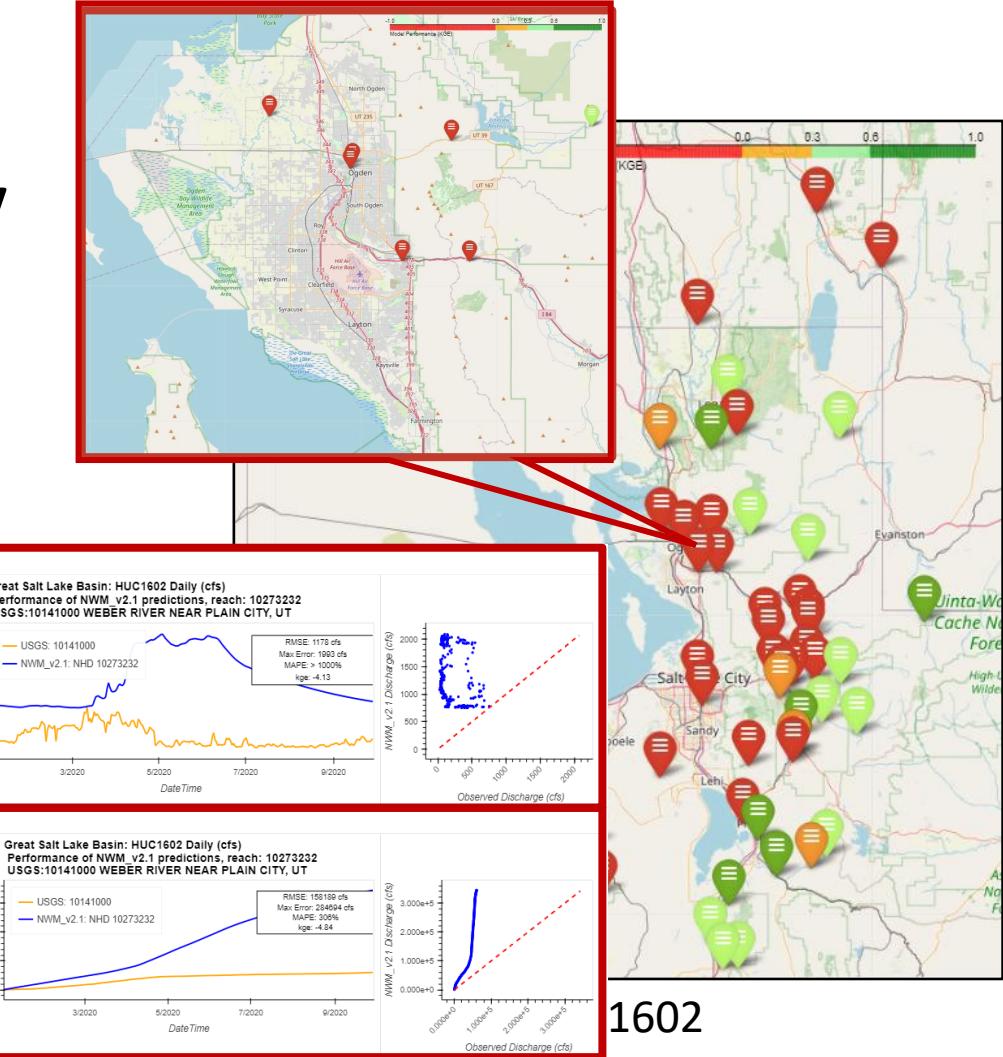


National Water Model Season-to-Season Water Supply Forecasting

Improved Season-to-Season water supply forecasting needed in Western US.

- Ecosystem Management
- Power Generation (Hydro)
- Environmental Management
- Water Supply
- Recreation

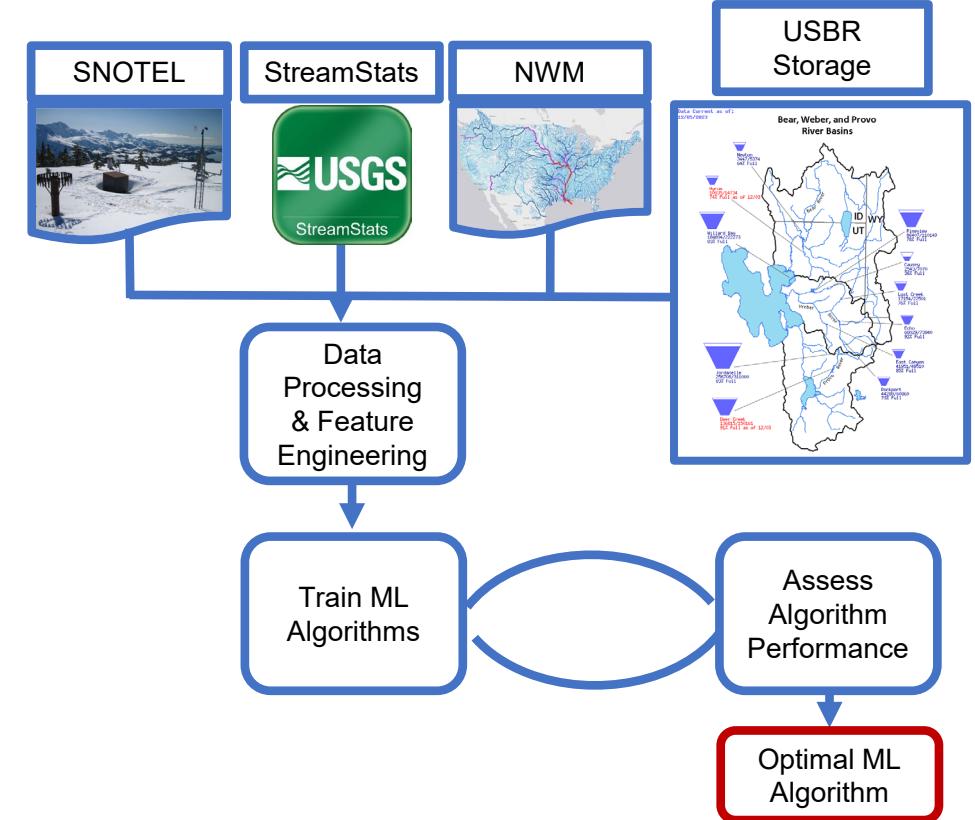
Leverage National Water Model



National Water Model Season-to-Season Water Supply Forecasting

Season-to-Season water supply estimates for up to 2.7 million stream segments.

- Post-processing NWM streamflow
- Catchment frozen water storage
- Catchment characteristics
- Catchment reservoir storage
- Machine Learning



National Water Model Season-to-Season Water Supply Forecasting

USGSid	NHDPlusid	NWM RMSE (Mm ³)	NWM-ML RMSE (Mm ³)	NWM Pbias (%)	NWM-ML Pbias (%)	NWM kge	NWM-ML kge
10126000	4605050	437	133	-62	26	-0.73	0.62
10130500	10092262	78	17	-132	19	-0.77	0.59
10134500	10277268	46	7	-158	20	-1.18	0.70
10136500	10274616	248	34	-164	16	-1.50	0.63
10137500	10274270	18	16	38	33	0.56	0.58
10141000	10273232	457	19	-461	-20	-7.83	0.75
10155000	10373622	72	77	42	44	0.46	0.40
10164500	10329013	40	1	-237	0	-2.38	0.98
10171000	10390290	1339	29	-2329	-52	-30.38	0.38

National Water Model for Water Resources Management:
Post-Processing with Machine Learning in Controlled Basins.

Savalan Neisary^a, Ryan C. Johnson^b, Md Shahabul Alam^b and Steven J. Burian^{a,b}

^aDepartment of Civil and Construction Engineering, University of Alabama, 3043 H.M. Comer 245 7th Avenue, Tuscaloosa, 35487, AL, USA

^bAlabama Water Institute, University of Alabama, Cyber Hall 1046, Tuscaloosa, 35487, AL, USA

ARTICLE INFO

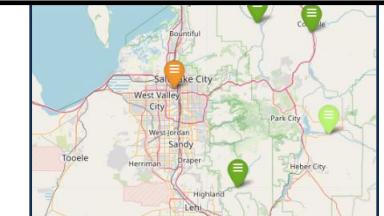
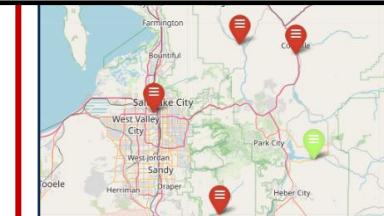
Keywords:
 Machine Learning
 Decision Trees
 Water Supply Forecasting
 Water Resources Management
 Great Salt Lake

ABSTRACT

Accurate streamflow prediction is critical for water resources management in the western United States (US), especially in balancing the beneficial uses of multiyear reservoir carryover, ecological interests, irrigation, and recreational use. With Great Salt Lake (GSL) in Utah, US facing decreasing lake levels and potential environmental disaster, we investigate the utility of the National Water Model (NWM) as a water resources management tool. Using the NWM retrospective v2.1 simulations and comparing daily and seasonal flow volumes to 48 USGS gauge stations, the NWM demonstrated decreasing model performance at locations closer to the GSL where there is extensive water resources infrastructure. Given the reaches closest to the GSL are critical for lake-level management, we explored XGBoost, Random Forest, Long Short-Term Memory, and Convolutional Neural Network Machine Learning (ML) algorithms to post-process the outputs of the NWM to account for the impacts of large-scale water resources infrastructure. Using upstream reservoir storage percentage of full capacity, the seasonality index, Snow Telemetry (SNOTEL) snow water equivalent (SWE) catchment observations, and NWM outputs as inputs into the ML pipeline, we observed significant improvement in the estimated daily flow (cms) and seasonal accumulated flow volumes (mm³). The overall Kling-Gupta Efficiency (KGE), Nash-Sutcliffe (NSE), Percent Bias (PBIAS), and Root Mean Square Error (RMSE) for the GSL basin increased by x, y, and z, respectively. The results of the study demonstrate that water resources infrastructure may be the single greatest challenge in hydrological modeling, being the dominant regional hydrological process, and novel ML methods can account for the impacts of the infrastructure on regional streamflow without explicitly adding infrastructure rulesets.

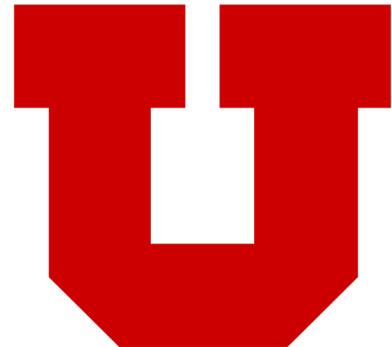
NWM-ML Post Processing:

- Accounts for reservoir operations w/o explicit p
- Works in controlled and natural basins
- ML “learns” hydrological processes
- Monthly to Annual flow estimates
- Leverages NWM-Retrospective products (flow + climate)





Advancing Snow Observation Systems to Improve Operational Streamflow Prediction Capabilities



PI: McKenzie Skiles – University of Utah

Co-PI: Christian Skalka –University of Vermont

Co-PI: Jeff Horsburgh – Utah State University

Co-PI: Ryan Johnson – University of Alabama

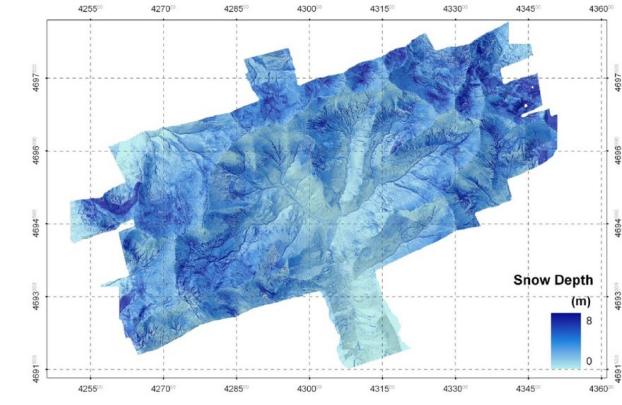


Advancing Snow Observation Systems to Improve Operational Streamflow Prediction Capabilities

Problem: *In situ* snow monitoring infrastructure has proven valuable but exhibits limitations for hydrologic models in because of the spatial heterogeneity of snow distribution.

Research Activities

- Develop and adapt snow survey optimization ML algorithm to identify the optimal sensor deployment locations
- Develop a low-cost low-power snow sensing system network (w/communications) based on open-source firmware (e.g. Arduino IDE)
- Create catchment snow-off and snow-on datasets using UAV (Structure-for-motion & LiDAR)

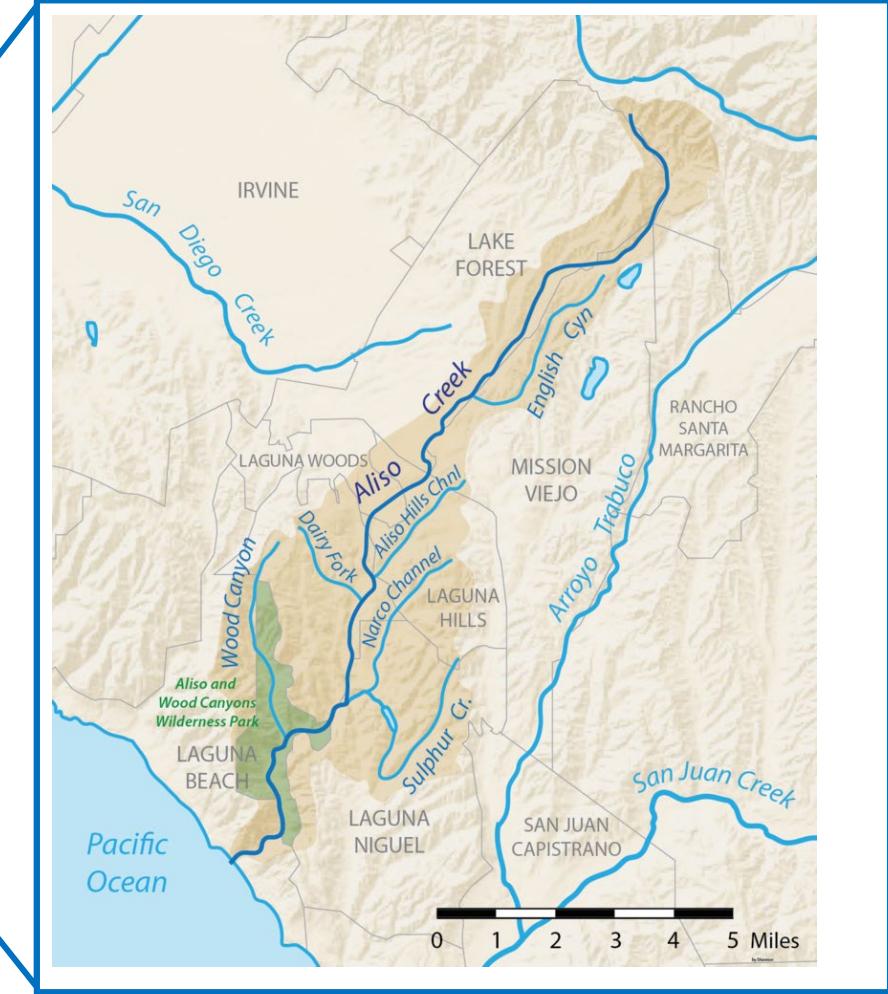


Questions



Aliso Creek Water Quality

- Water quality (nitrates) increased in the urban water system, likely from landscaping runoff.
- Water quality improved as the stream traveled through natural riparian habitat (Aliso Wilderness Park).
- Water treatment facility effluent reduced water quality.
- Golf course reduced water quality.
- Beach closures due to poor water quality.

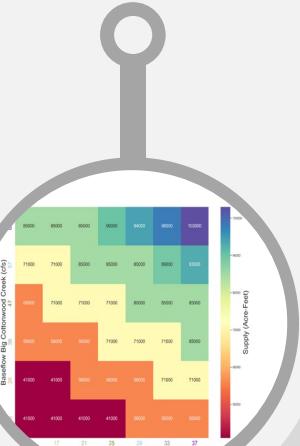
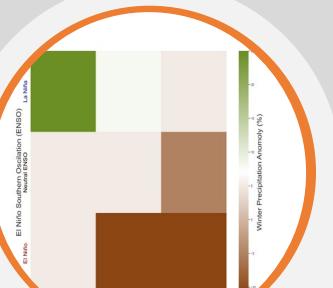


Operational Decision Timeline

Hydrology

Prelim estimates of streamflow yield

January



October

Climate

Winter precipitation estimates

Climate

Summer precip & temp outlook

Hydrology

Refined estimates of surface water yield

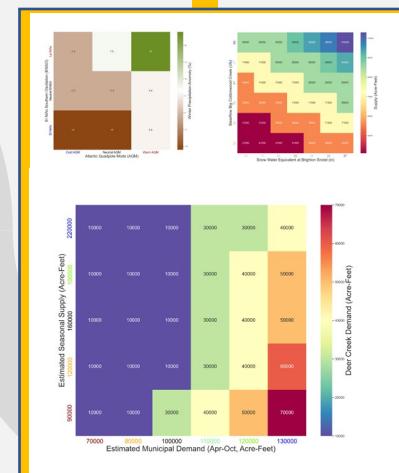
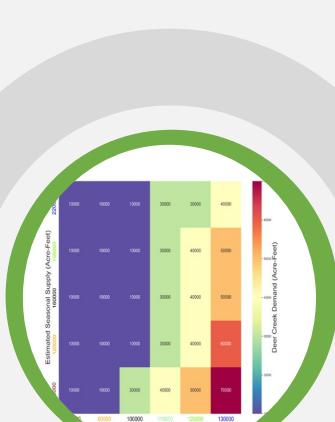
Engineering

Refined estimates of system vulnerabilities

SLC

Begin formulating management decisions

March



February

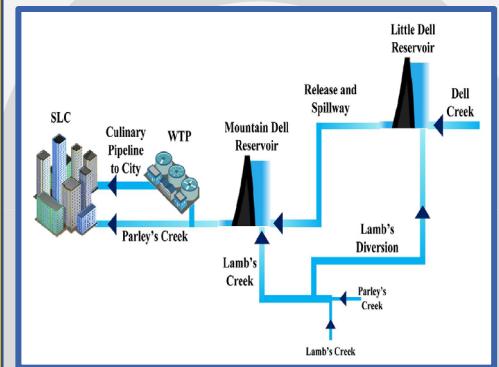
Engineering

Seasonal water system projections

April-June

All

Assess system status
Revise/Initiate operational decisions



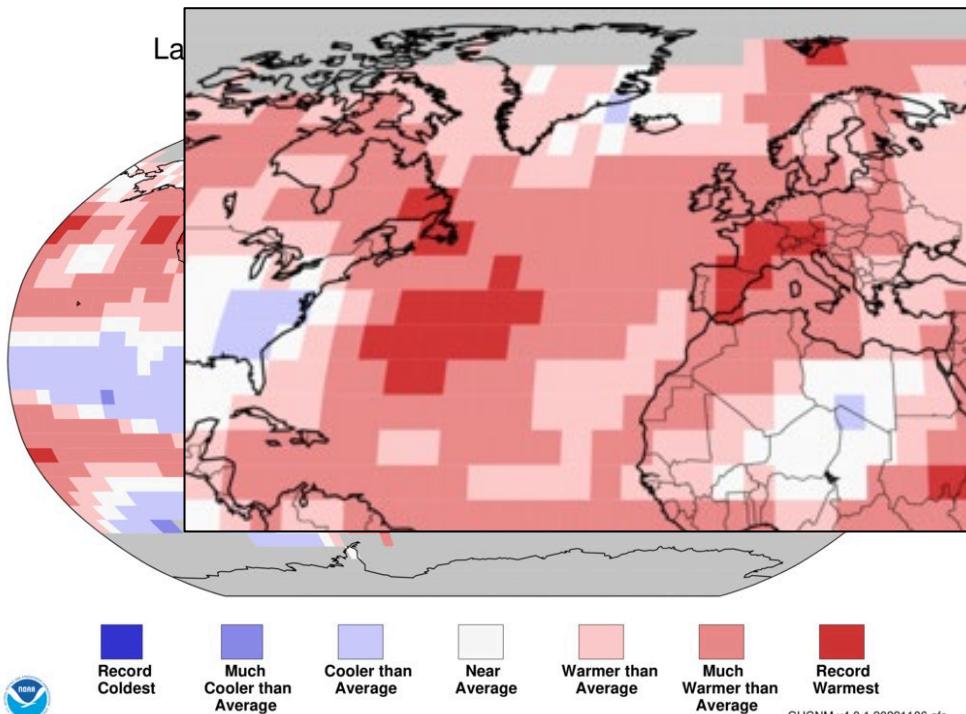
Precipitation Estimates (October)

CURRENT STATUS:

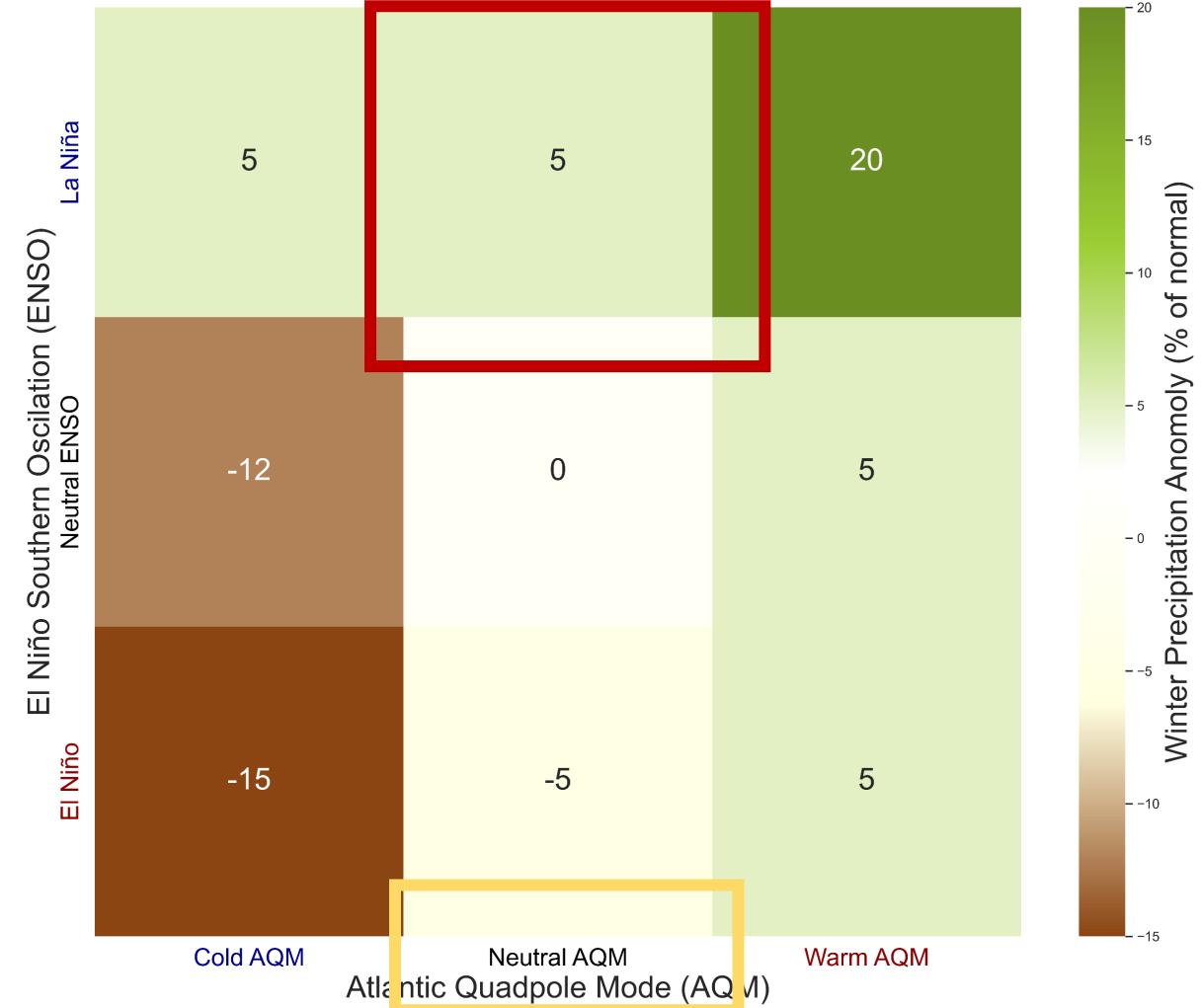
LA NIÑA ADVISORY

OCTOBER 13, 2022

Forecasts indicate a 75 percent chance that La Niña—the cool phase of the ENSO climate pattern—will persist across the tropical Pacific for the third winter in a row. The map at left shows the broad swath of cooler-than-average water (blue) across the Pacific at the equator, one of the hallmarks of La Niña. Scroll down to learn about La Niña's influence on global and U.S. seasonal climate.

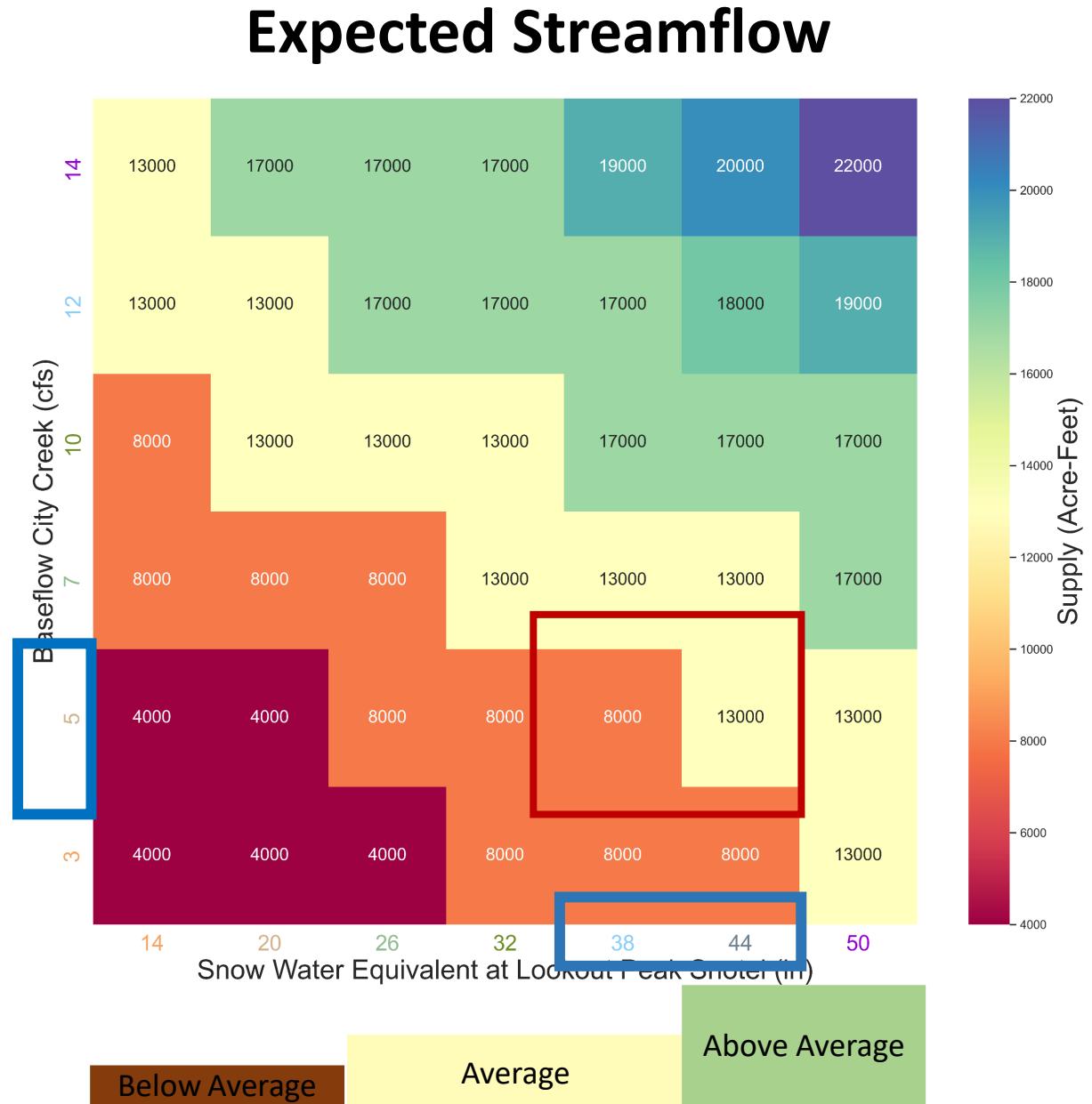


Precipitation Anomaly



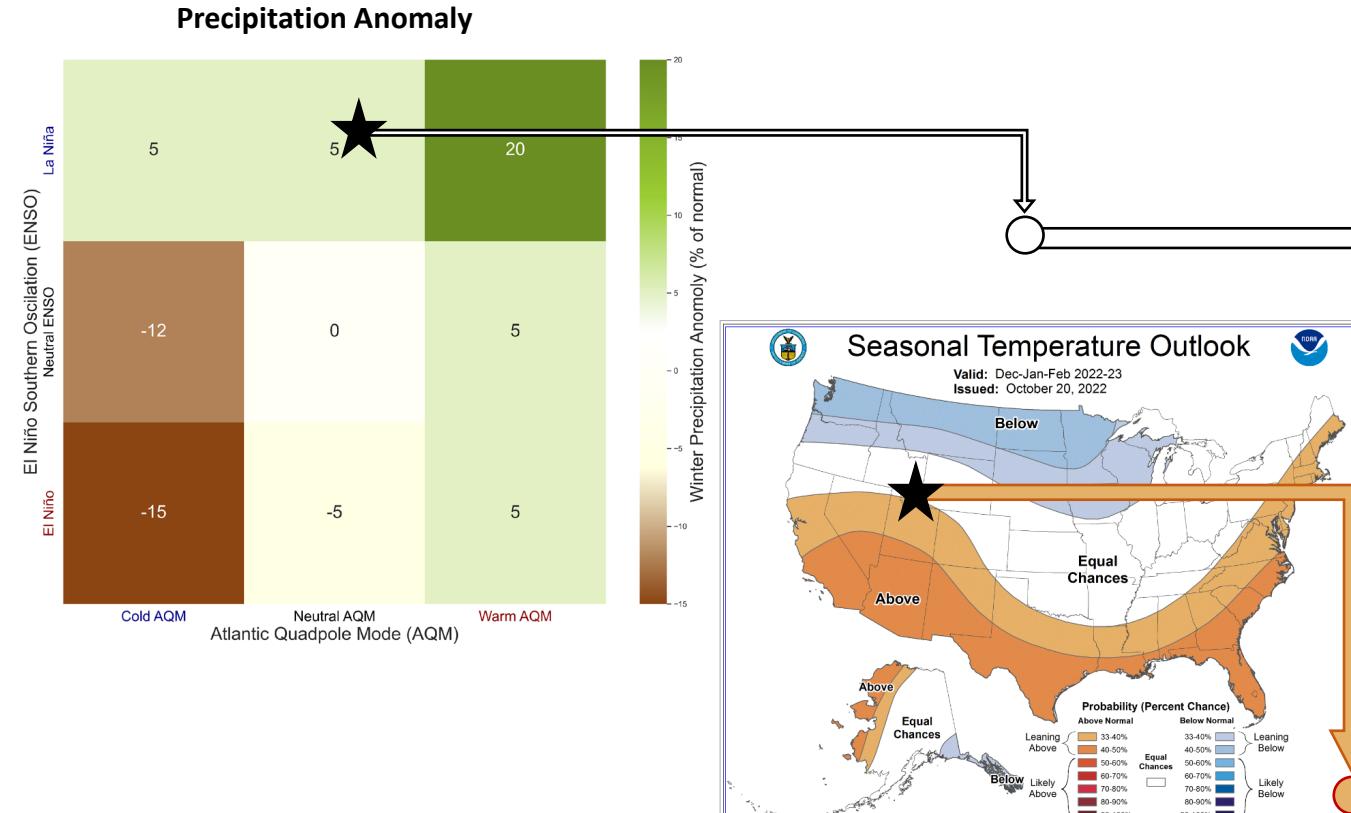
Surface Supply Yield (February)

- Use precipitation prediction to determine total max SWE.
 - Average annual winter baseflow.
 - Reduce uncertainty in supply estimate.

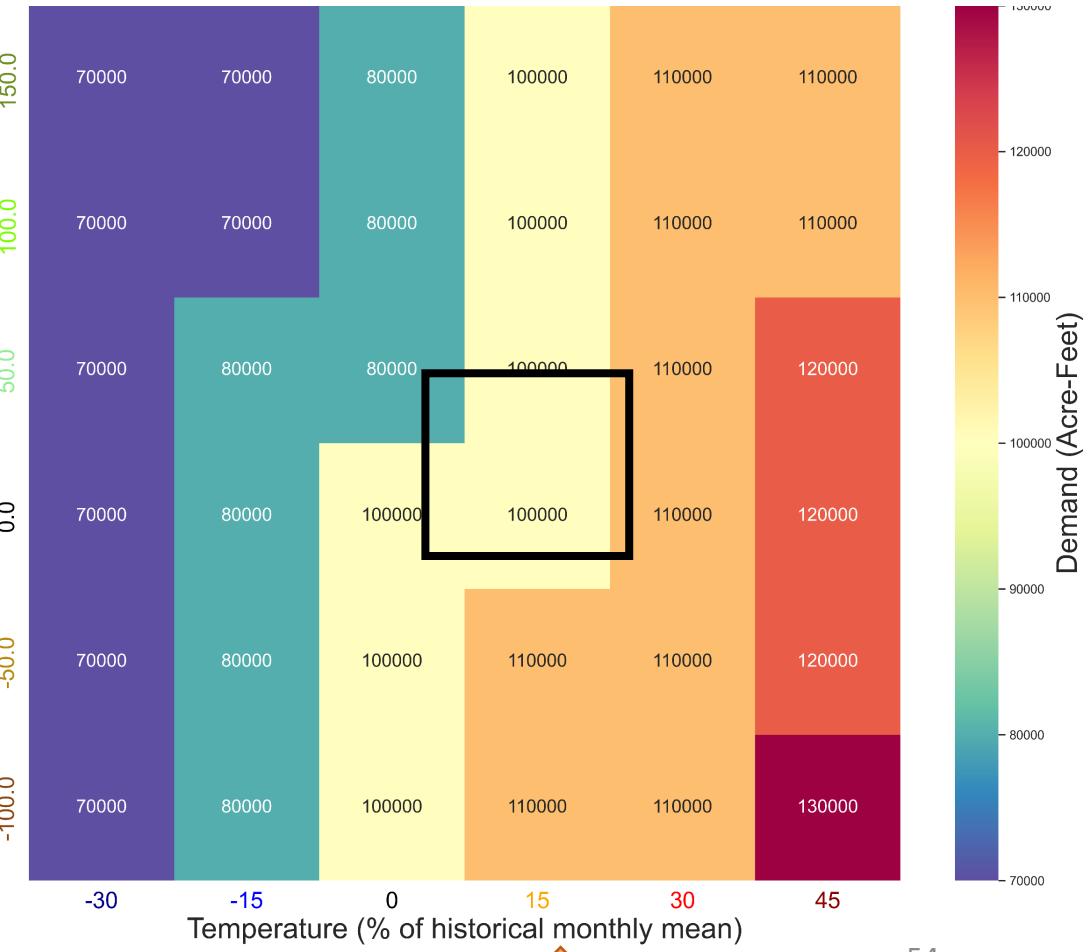


Municipal Demand Estimates (February - April)

Use climate prediction (e.g., spring NOAA outlook and/or Climate Matrix) to determine expected monthly temperature and precipitation

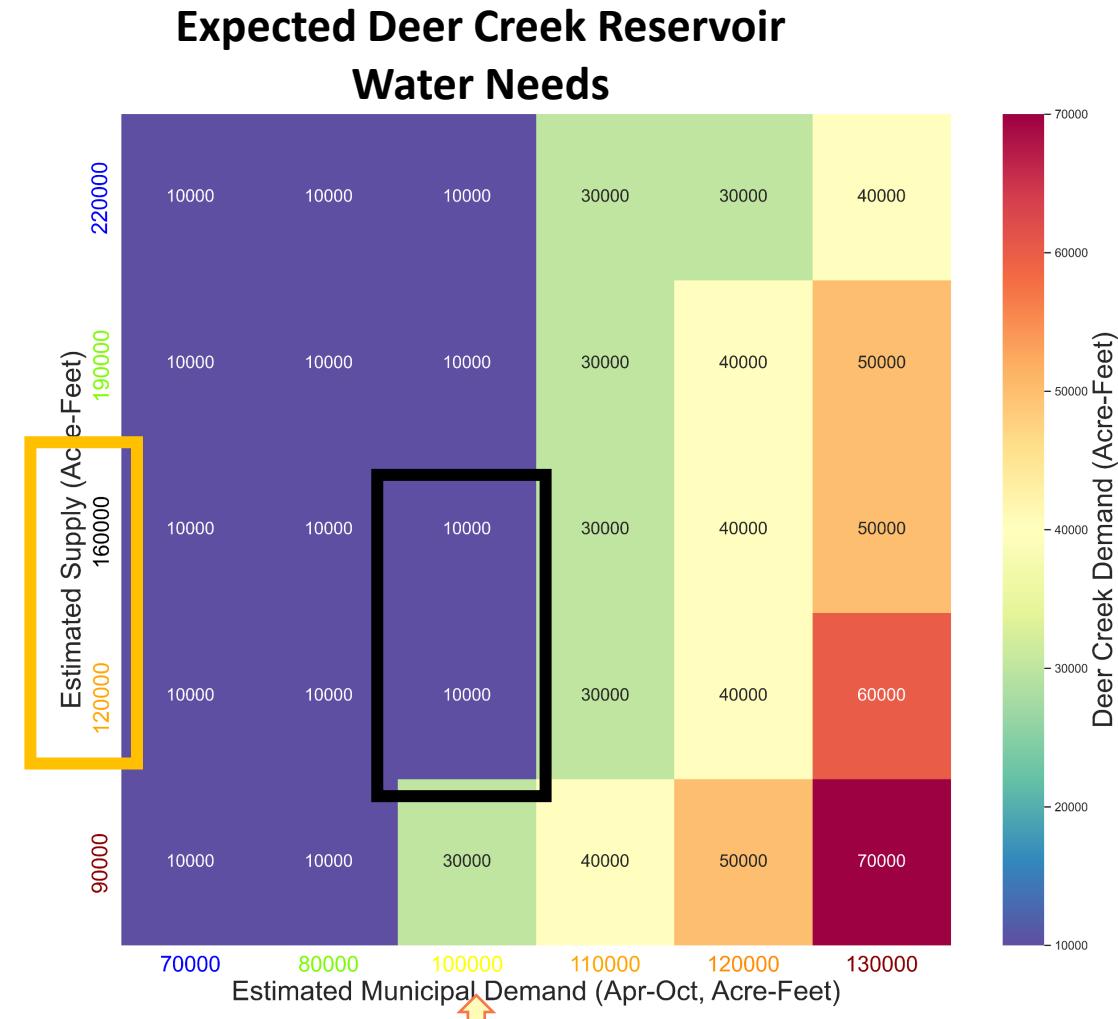
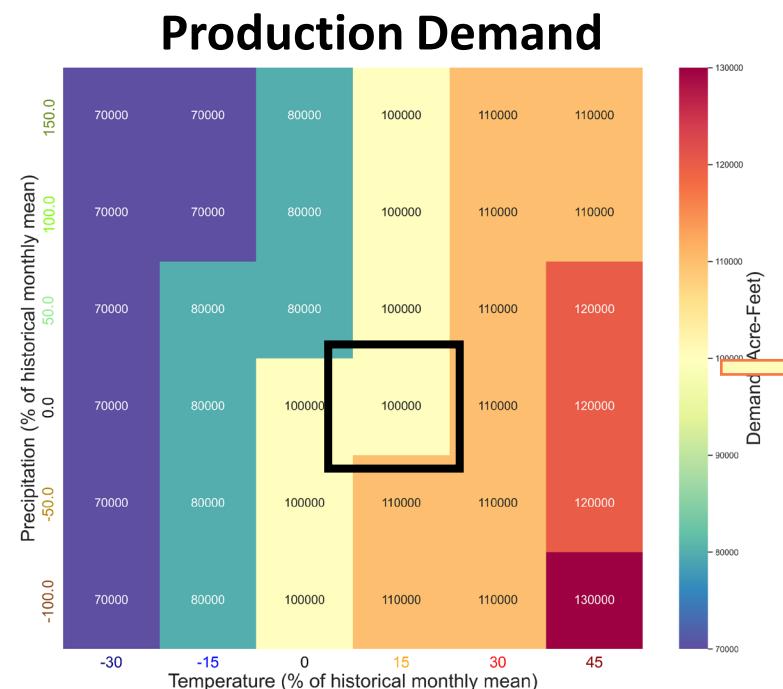


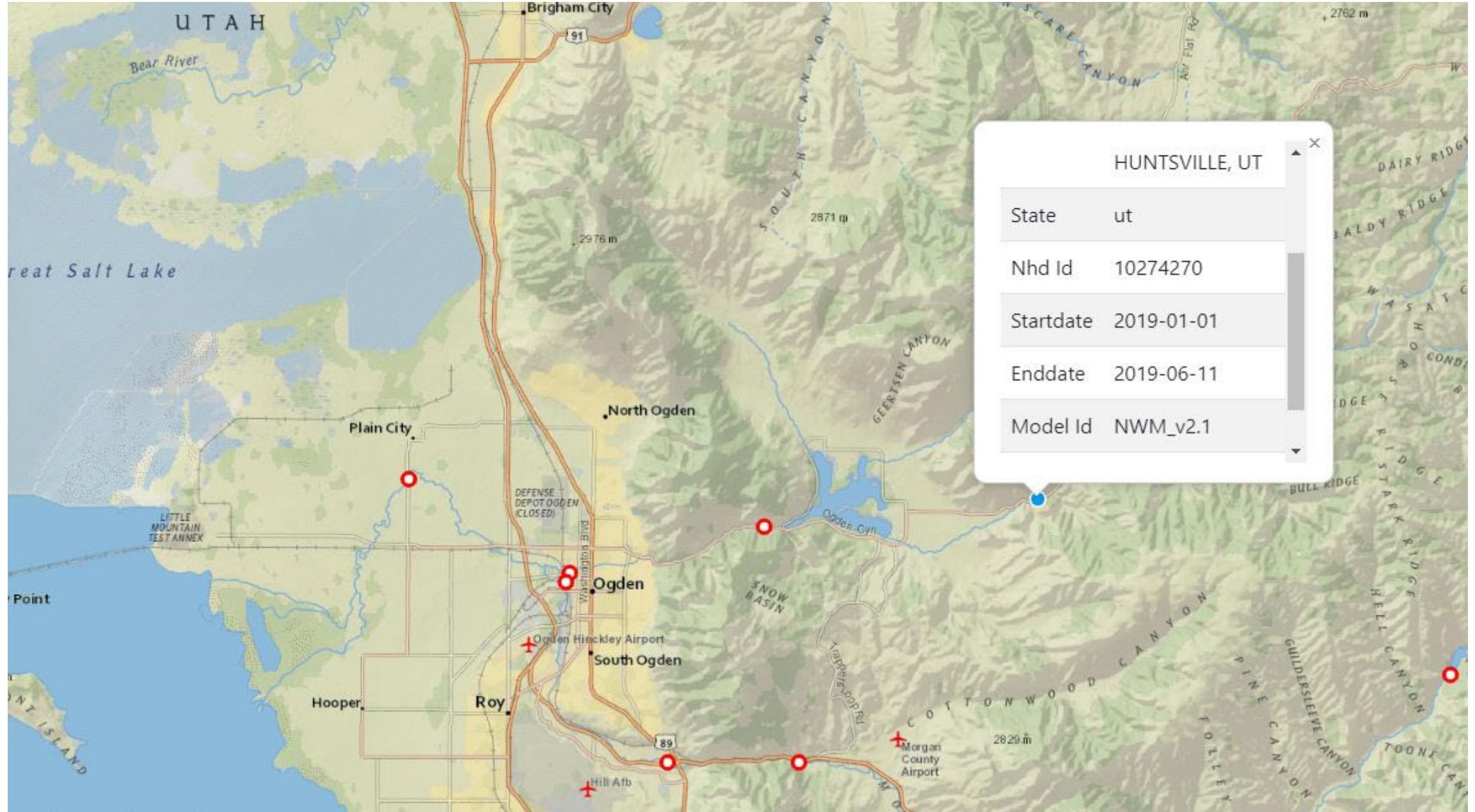
Total Production Demand

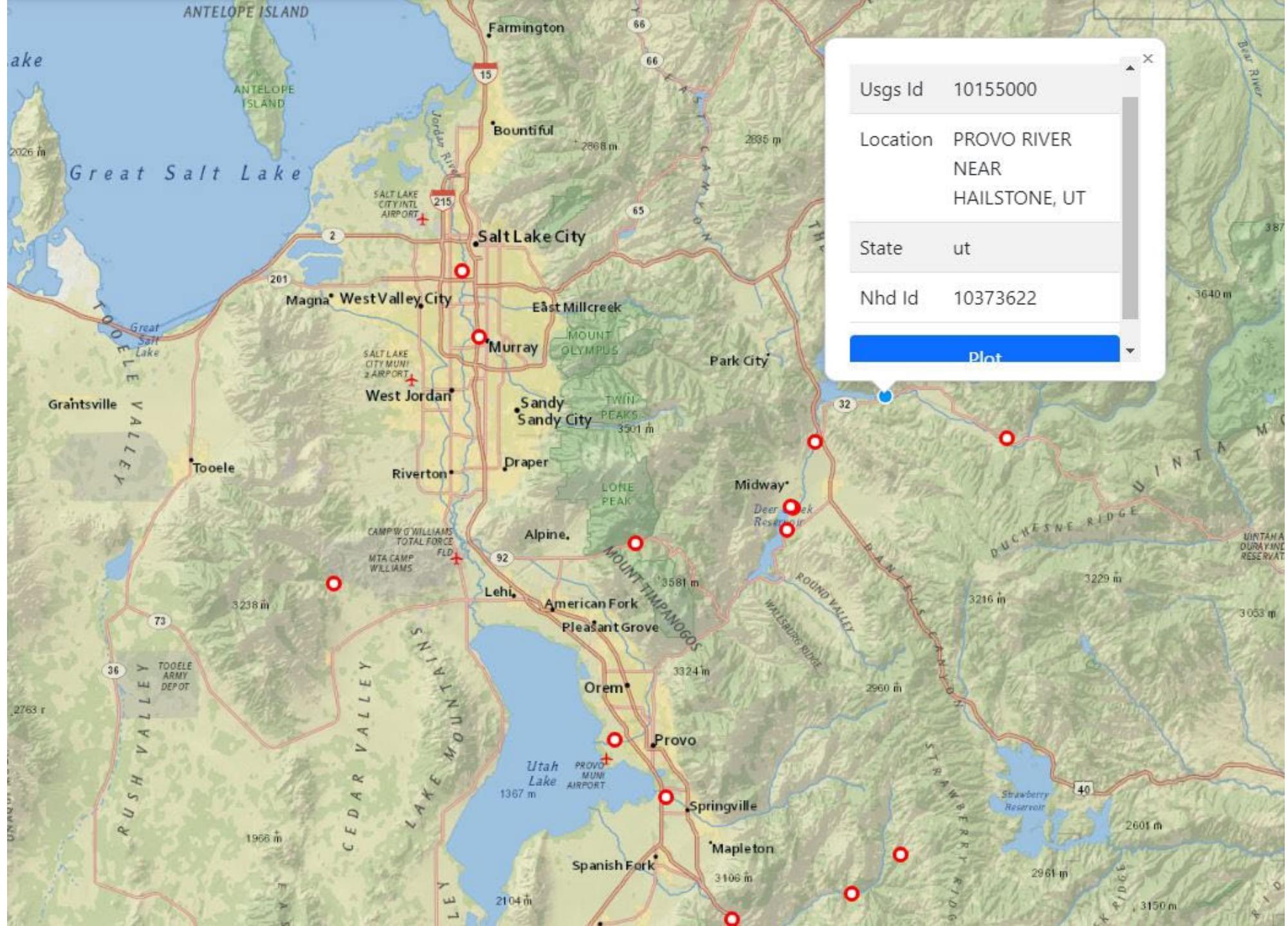


Water System Performance (February - April)

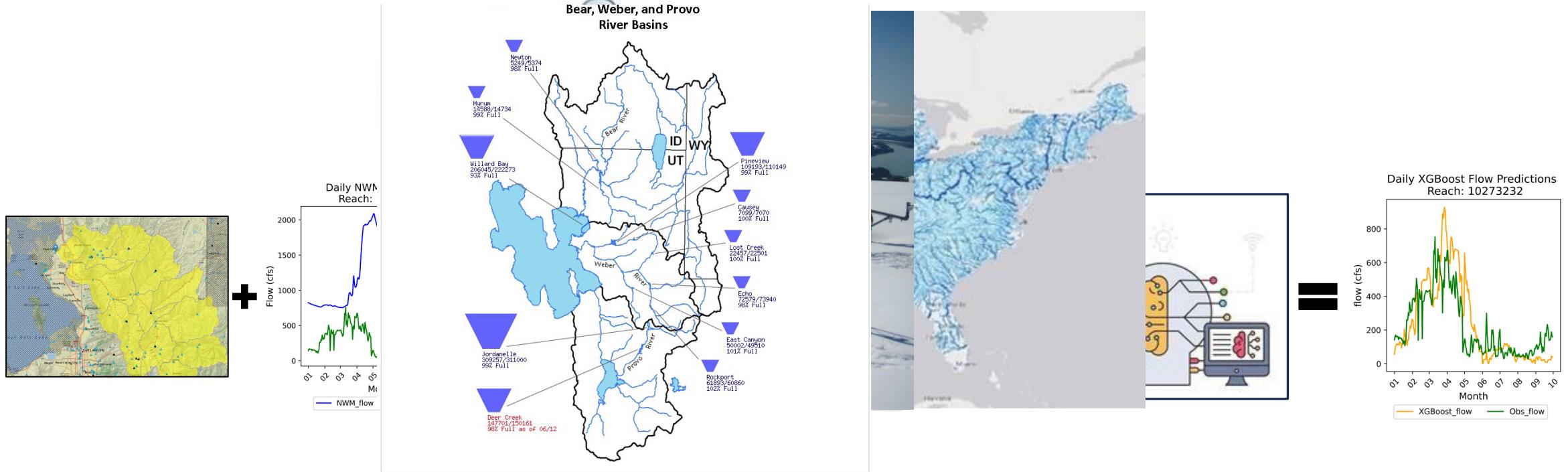
- Using supply tools, estimate 120,000 to 160,000 acre-feet.
- From NOAA climate outlook, demand should range between 80,000 to 110,000 acre-feet (mean 100,000 acre-feet).
- Estimated Deer Creek needs: 10,000 to 40,000 acre-feet.







National Water Model Season-to-Season Water Supply Forecasting





Advancing Research in Geosciences using AI/ML

Summary: Connections with snow and streamflow AI/ML methods to advance the understanding of geosciences using AI/ML methods.

- Advance core geosciences program goals and use AI/ML methods for addressing scientific problems. Build the capacity of AI/ML to explore and/or emulate physically based models.

Core geoscience goal: Advance the representation of snow on the earth's surface to support water supply forecasting and explore the impacts of a changing climate.

Broader impacts: developing AI/ML training datasets, python-based tools (evaluation), open source software, supporting workforce development. Supporting the greater hydro-meteorological practioning and research community.



U.S. National
Science
Foundation

Environmental Engineering

Summary: Investigation of a no-snow future within the western US, beginning with the Great Salt Lake using SNOTEL to investigate precipitation phase during the snow season (October to July).

Research questions: how has precipitation phase changed in the past half century and what can we expect in a changing climate in montane, high elevation watersheds. Most studies relate a no snow future to the impacts of temperature on snow melt, investigation peak SWE and the duration of snow. However, snowmelt is primarily driven by the energy from SW radiation.

Applications benefits: connect newfound knowledge related to precipitation phase (or at least the temperature when precip is occurring) to naturalized and managed water systems. How much snow can we expect to occur in the future and relate it to the quantity of melt to fill reservoirs/GW recharge,

Environmental Sustainability: GSL sustainability, no snow future, climate change/scenarios/impact of scenarios on water resources management.



NIDIS: future drought risk across the West and in the context of a changing climate

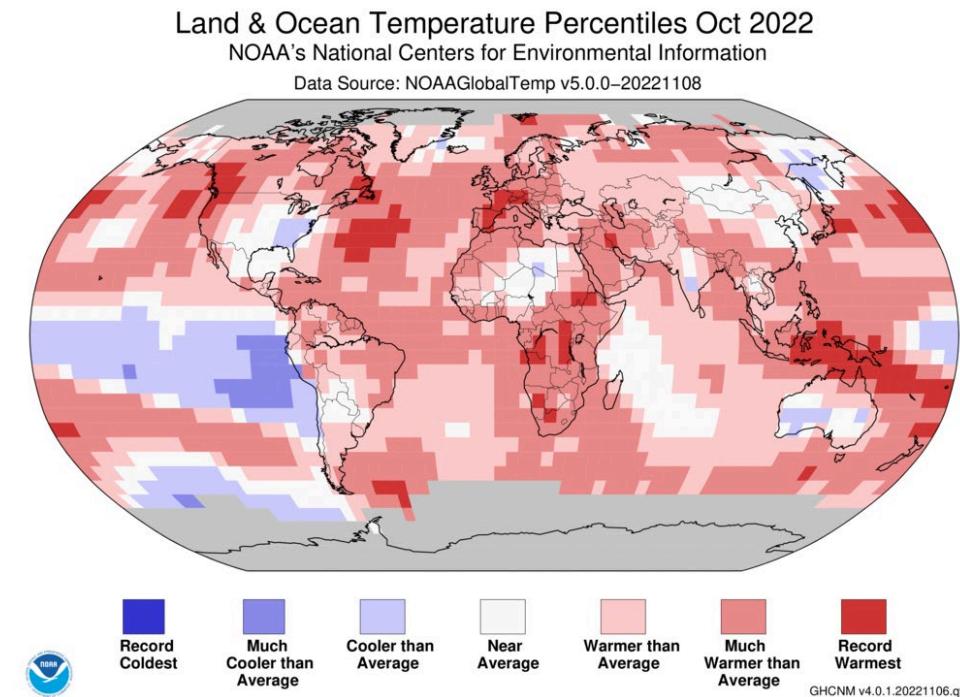
- Drought vulnerability assessments
- Developing drought plans and communication plans
- Identifying primary drought impacts
- Optimal drought indicators and/or triggers and improving monitoring
- Developing drought dashboards with relevant tools and information and demonstrating the application of drought data to enhance decision-making





CVP – Advancing the Understanding of multi-year to decadal climate variability and predictability for US climate predictions

- Data analysis and investigation focused on mechanisms that govern the variability of the coupled climate systems and its predictability on the multi-year to decadal timescale with long-term observation data and or model data.
- Investigation of the relationship between the Atlantic meridional overturning circulation (AMOC) and impacts on the cryosphere and hydroclimate in the Western US



Positive Action Principles

1. Positive actions for your mind
2. Responsible self-management
3. Continuous Improvement