Isabelle Anne Goebel, Rain Zou, Claire Meaney, Maram Ahmed

May 1, 2025

Discovery Symposium Award

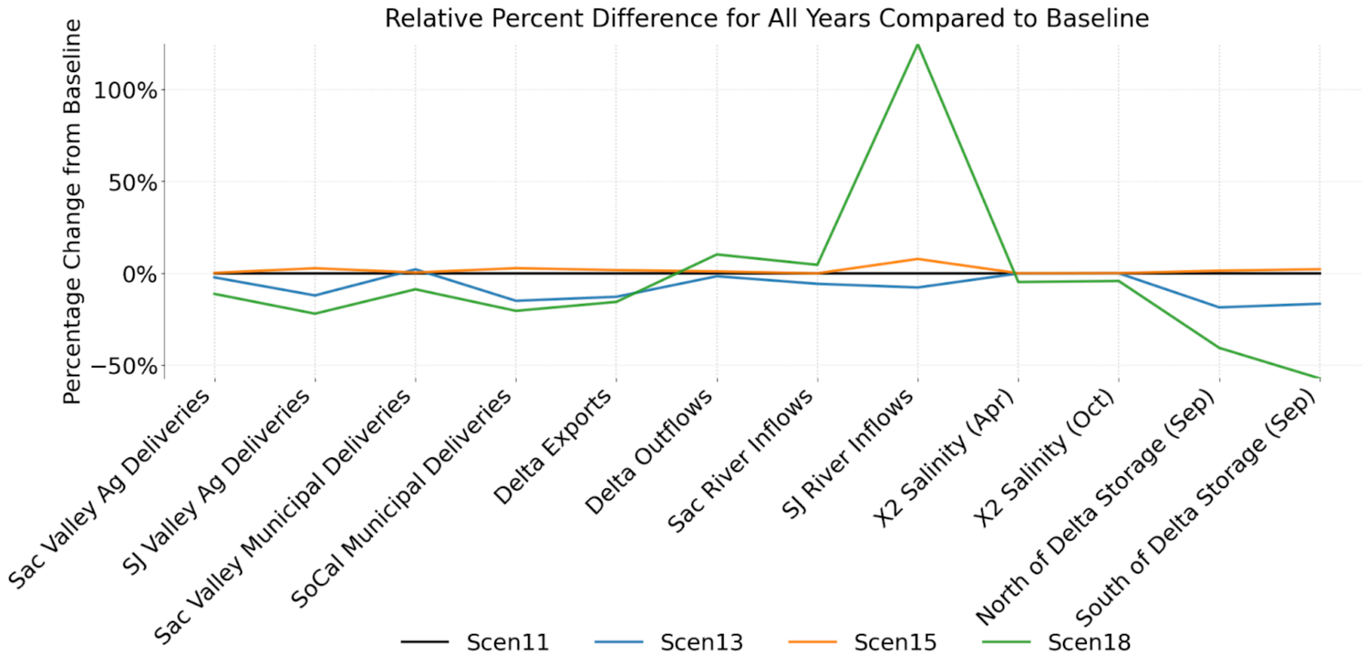
The COEQWAL team is providing tools to tackle a critical problem in California: how to manage water resources in a way that is equitable for people, food production, and the environment, in a way that is resilient to threats such as those brought on by climate change. The goal of our project is to show how water is allocated and where it gets distributed, with the potential to support more informed water policy decisions in the future. For this goal, we developed a data science-based workflow to process complex hydrological data. In particular, we developed a library of methods to evaluate the costs and benefits of alternative management strategies, along with visualization tools, including a powerful interactive dashboard shown below that greatly simplifies the task of data analysis and makes it easier to understand. We focused on turning the data into stories that can be shared with a wide range of stakeholders and are applying for the data visualization award for our work utilizing data science to address critical societal, environmental, and economic problems.

We applied these methods to evaluate a library of different scenarios. For our poster, we illustrate this by plotting three of these different scenarios: climate change (13), groundwater pumping restrictions (15), and minimum flow requirements (18) against a baseline scenario (11). The baseline scenario serves as a reference point that allows us to easily compare the changes. It reflects current operations under the historical climate and hydrology in combination with the occasional relaxation of salinity and flow regulations following the Temporary Urgency Change Petition (TUCP) process during extreme drought conditions. We focused on three main types of variables in our plots: storage, deliveries, and flows. Storage variables measure water levels in reservoirs, delivery variables represent water deliveries through the State Water Project (SWP) and Central Valley Project (CVP) to different contractor groups, and flow variables track the movement of water through rivers and canals. Using these variables, we can see where water moves under different scenarios and water year types.

Isabelle Anne Goebel, Rain Zou, Claire Meaney, Maram Ahmed

May 1, 2025

Parallel Line Plot

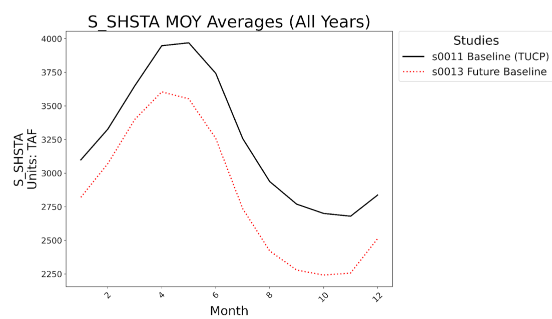


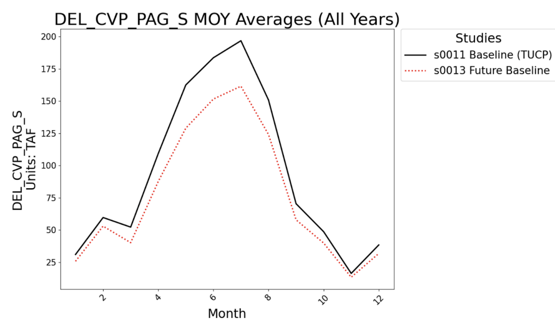
One of the primary visualizations we used in our analysis is the parallel line plot shown above, which represents the relative percent difference from the baseline scenario (scen11) across a range of variables. This specific plot was chosen because it concisely displays multiple scenarios and variable types (such as deliveries, flows, and storage) into one interpretable graph. By having continuous lines for each scenario, viewers are able to easily track patterns and quickly identify increases and decreases for related variables. Previously, we considered creating individual exceedance probability plots for each variable and scenario, but found it difficult to compare multiple scenarios simultaneously or identify consistent patterns across variables. However, the single parallel line plot clearly illustrates overarching patterns across a diverse set of variables.

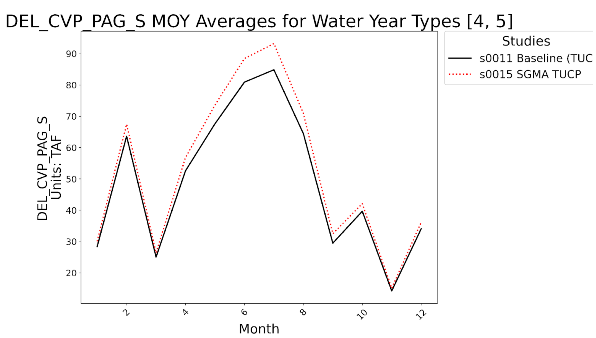
To improve readability and accommodate longer variable names, we rotated the x-axis labels and assigned each scenario a unique color. Percent difference was selected as the y-axis metric in order to allow changes across different types of variables, such as deliveries, flows, and storage, to be interpreted on a consistent scale. The horizontal reference line at 0% (where the baseline scenario 11 lies) makes it easy to see where scenarios deviate from the baseline.

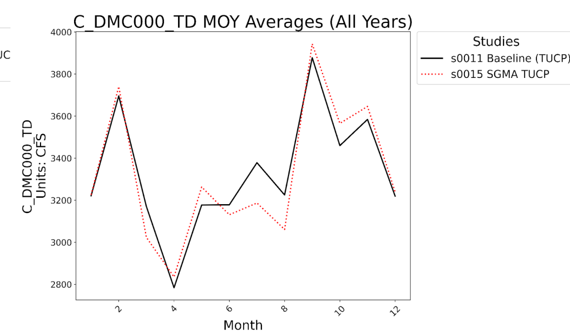
This visualization highlights the advantages and disadvantages of each water management scenario. For example, in scenario 18, which introduces new minimum environmental flow requirements based on ecologically informed seasonal flow targets to benefit ecosystems and fish populations. We observe a significant increase in San Joaquin River Inflows (nearly 150%) and a greater than 50% decrease in South of Delta Storage in September. This reflects the cost of meeting environmental goals in terms of reduced available water in storage. In contrast, scenario 15 minimally differs from the baseline scenario, likely due to the fact that scenarios 11 and 15 are both built on historical hydrology operations with TUCPs, but scenario 15 applies SGMA groundwater pumping restrictions. Visualizing this comparison alongside others helps illustrate how incorporating specific policy changes translates into measurable outcomes, making analysis more approachable for both technical and non-technical audiences.

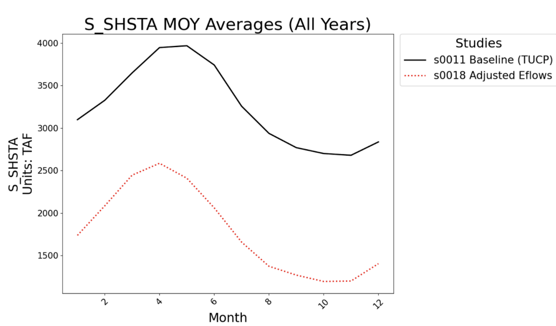
Month-of-Year Average Plots

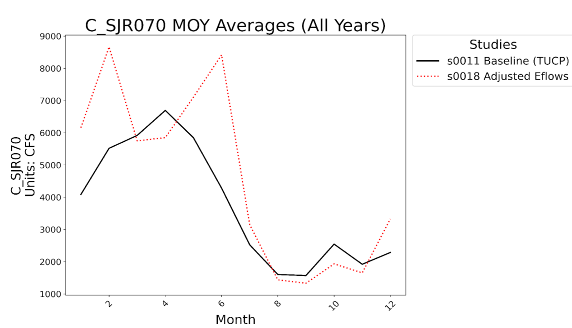












Isabelle Anne Goebel, Rain Zou, Claire Meaney, Maram Ahmed

The other primary visualizations on our poster are the line charts above. These six selected plots showcase the seasonal variations between our baseline scenario, s0011, and the scenario. We chose to show the month-of-year average plots instead of our other types of line plots — timeseries, annual totals, and annual exceedances — because they are the easiest to interpret. Out of the 17 variables we plotted, we chose to show the ones above because they were the most significantly affected in their respective scenarios. We chose them by analyzing the parallel line plots and the month-of-year averages for all the variables. For example, in scenario 18, reservoir storage, such as Shasta (bottom-left), is much lower than the baseline throughout the year. Using these month-of-year line charts, it is easy to see the differences between the scenarios and seasons, such as summer water storage being low due to reduced precipitation and increased demands.

We chose to keep the baseline scenario the same style, a solid black line, so it was easy to differentiate when looking through different scenarios. It is the same for the parallel line plot. We added a short description of what each scenario is in the legend to remember what each plot represents. The title starts with the variable, then is followed by the type of plot, and what years are being averaged. This semester, we added a functionality to our plotting library that allows us to select water year types, numbers 1 through 5, which represent the amount of precipitation for that year. 1 being an extremely wet year and 5 being extremely dry. Numbers 4 and 5 represent a drought, which can be seen in the middle-left plot. This helps see how certain variables will be affected in wet versus dry years, which can have a large impact on the communities that rely on the water provided to them by the State Water Project or Central Valley Project.

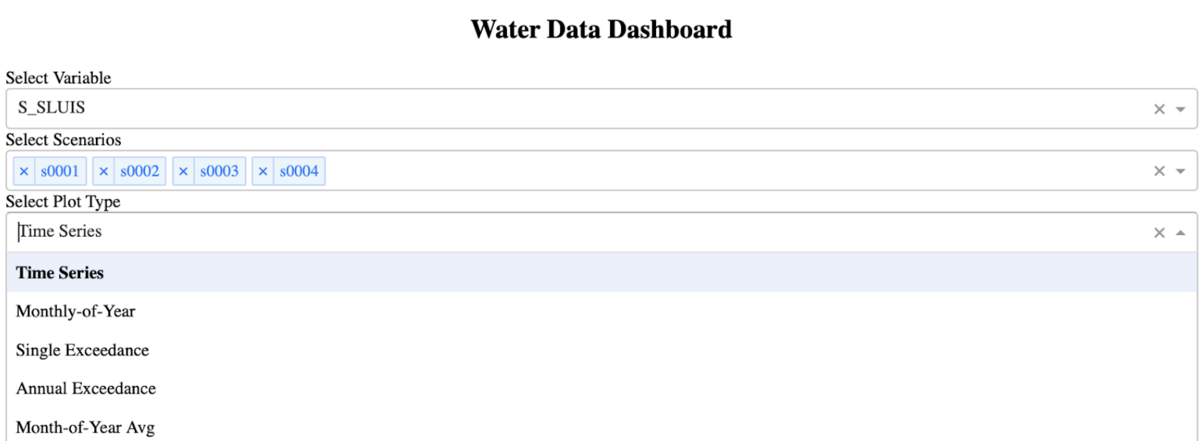
After seeing the general trends in our plots, we asked ourselves why this is the case and why it is important. In our climate change scenario, we can see that reservoir storage is generally lower than the baseline, especially in spring and winter months. We can also see that agricultural Central Valley Project contractors south of the Delta are most significantly affected by climate change, as their deliveries are reduced the most. This happens because of reduced precipitation and snowpack, as well as the delivery priorities to different contractors. This is important because over 10% of the United States' food supply comes from California's agriculture[1]. In our groundwater pumping restrictions scenario, we can actually see the opposite: agricultural deliveries increase, especially in dry years. This may be a good thing because groundwater

[1] https://www.fsa.usda.gov/Internet/FSA_File/10cafacts_v3.pdf

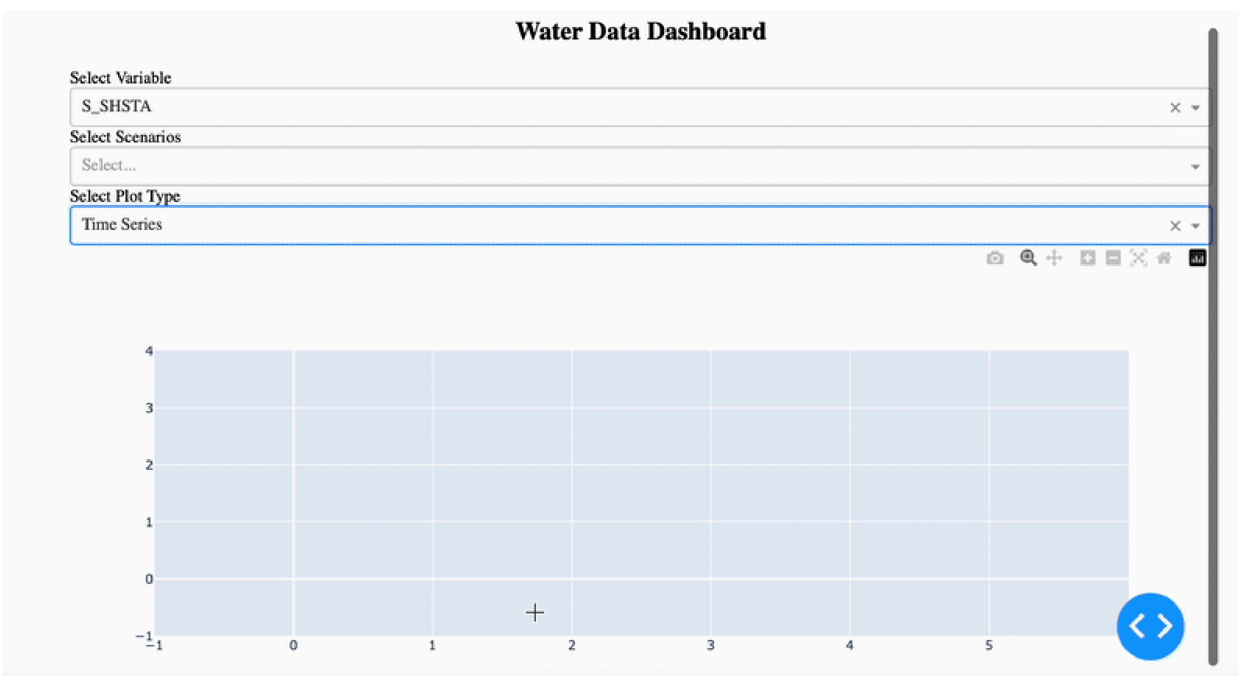
pumping causes other issues such as land subsidence, well dryups, and groundwater contamination, however, many rely on groundwater to meet their needs. In our minimum flow requirement scenario, we can see that reservoir storage and water deliveries drop substantially. This is because the water that would be getting stored in reservoirs is being allocated for the environmental flows, which are important for protecting fish and wildlife that rely on the water for annual migrations. Minimum flows are also important for preventing saltwater intrusion from the San Francisco Bay. These plots help tell a story of how different scenarios will affect the future of California's water resources, and it is important to keep in mind who is being affected in each.

Interactive Dashboard

Another visualization tool that makes this analysis easier to understand includes an interactive dashboard. Below is our interactive water data dashboard that allows users to choose from three main options: variables (e.g., reservoir storage, water delivery, flow), scenarios (based on different combinations of climate conditions and policy assumptions), and plot types (time series, monthly distribution, exceedance, and monthly average plots). The interactive dashboard offers an intuitive way to explore the data and quickly navigate to the areas of interest, making it a valuable tool for both specialists and the general public.

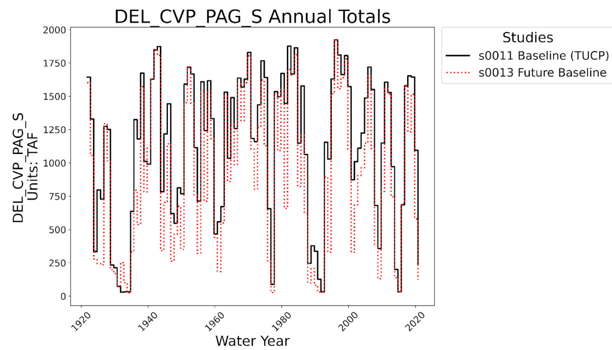


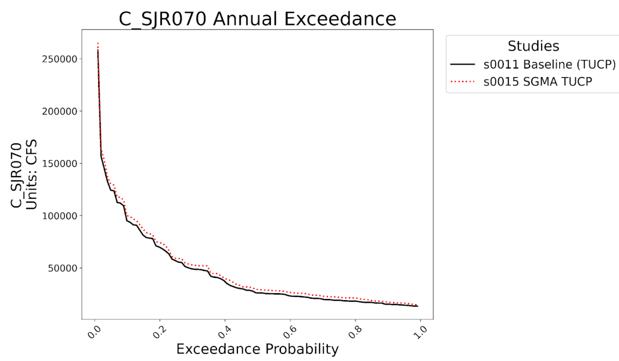
Below is a brief demonstration of how the dashboard can be used. While only one variable and one plot type can be selected at a time, multiple scenarios can be chosen simultaneously for direct comparison of how water resources change under different conditions. Each scenario is represented by a different color, with a legend displayed on the right. The time series plot displayed changes in water availability for the selected scenario from October 1921 to September 2021. Users can freely zoom in and out and interact with the plot to view the most accurate values. The monthly-of-year plot presents the same time series data, but allows users to select specific years and take a closer look at them. Drought years are marked in the selection menu for clarity. The single exceedance plot shows the probability that a value will be equaled or exceeded. A higher exceedance probability indicates that the variable hits this value more frequently. The annual exceedance plot shows how often the annual totals reach or exceed chosen thresholds, for example 90th percentile. An option is provided to users to evaluate exceedance patterns from a specific month of the year, such as April or September. The exceedance plots help users to understand how often the water availability falls within normal ranges or becomes too high or too low. The month-of-year average plot shows the average monthly water availability data across all years, which is useful for analyzing seasonal variabilities. A filter labeled 'Water Year Type' is provided for users to select data to include in the average calculation by year type, ranging from type 1 (wettest) to type 5 (driest).

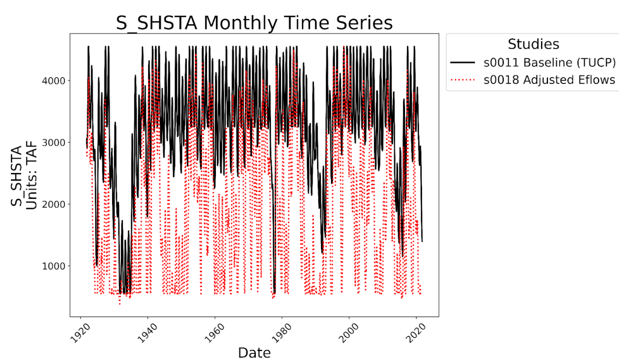


One potential improvement to make the dashboard more user-friendly and clear is adding brief explanations for each variable, scenario, and the data processing methods used in each plot type. Making the menu dropdown, label, legend, and axis description for the dashboard more intuitive is also an improvement that can be made in the future. Additional features, such as allowing the users to switch between units such as TAF and m^3 and saving the data and plots to downloadable files, can also be added.

Additional Line Plot Examples







Above are some examples of the other types of line plots we have. As we can see in the annual totals plot, the agricultural contractors get fewer annual deliveries in scenario 13. In scenario 15, the San Joaquin River is more likely to have higher flows, as we can see its line is above the baseline. Lastly, we can see the monthly time series for scenario 18, which is significantly lower than the baseline. Our tools allow us to easily switch between time series and annual total or exceedance plots, which help visualize and explain the trends in our variables and scenarios.