Final Write-Up

Lillian Makhoul

5/30/2025

1 Data

The data for this project was obtained from a series of sources. First, for the data on water flow, we used the simulated data gathered by the Bureau of Reclamation (April 2024, 24-month study). This data breaks us down into our three scenarios:

1. The actual water flow is often beyond what we predicted (wet season)
2. The actual water flow is often below what we predicted (dry season)
3. The actual water flow is typical to what we predicted (average season)

The data on demands are gathered from the Colorado River Compact, litigation for water rights of indigenous groups near the upper basin, or news reports on how Colorado allocated its water from the river to the municipal, industrial, and agricultural sectors. This is all summarized in the `demands data.csv` file. See `notes.tex` for tables.

2 Assumptions

Due to lack of access to data on Lower Basin water flows, we will only focus on the Upper Basin and the users of it. This also makes most sense to do since Colorado sources its water from the Upper Basin.

The users (i) in the notebook are labeled as following:

0. Colorado (Municipal)
1. Colorado (Industrial)
2. Colorado (Agriculture)
3. Upper Basin Tribes
4. Utah
5. Wyoming
6. New Mexico
7. Arizona
8. Colorado (as a whole, for constraints only)

Note that we are only optimizing water allocation for users 0-2. The remainder of users is meant for the constraints only.

The periods (j) in the notebook are labeled as following:

0. April 2023
1. May 2023

2. June 2023
3. July 2023
4. August 2023
5. September 2023
6. October 2023
7. November 2023
8. December 2023
9. January 2024
10. February 2024
11. March 2024
    End of WY
12. April 2024
13. May 2024
14. June 2024
15. July 2024
16. August 2024
17. September 2024
18. October 2024
19. November 2024
20. December 2024
21. January 2025
22. February 2025
23. March 2025
    End of WY
24. April 2025
25. May 2025
26. June 2025
27. July 2025
28. August 2025
29. September 2025
30. October 2025
31. November 2025
32. December 2025
33. January 2026
34. February 2026
35. March 2026
    End of WY

We will consider three scenarios (k):

- "Probable maximum": wet hydrologic condition
- "Probable minimum": dry hydrologic condition
- "Most probable": median hydrologic condition

3 Final Model

The parameters are defined as the following:

- NB_{ij} – net benefit when per unit of water is allocated to user i in period j
- Q_j unregulated water inflow in period j, given by data in Bureau of Reclamation
- C_{ij} – loss when per unit of water is **not** allocated to user i in period j
- $T_{ij,\text{cap}}$ – amount of maximum allowable allocation for user i
- p_{jk} – probability that scenario k occurs for period j
- R_i – legal requirements for water allocation to states/tribes. This can either be a proportion (for states) or a fixed value (for tribes).

Our variables will be

- T_{ij} – Fixed amount of water allocation target promised to user i in period j
- D_{i,Q_j} – amount of water shortage to user i when seasonal flow is Q_j
- $\epsilon_{(j-1),Q_{j-1}}$ is the surplus water inflow in period j according to Q_{j-1}

The final model we put together was

$$\text{Maximize } \sum_{i=0}^{2}\sum_{j=0}^{35} NB_{ij}T_{ij} - \sum_{i=0}^{2}\sum_{j=0}^{35}\sum_{k\in\{\text{min, most, max}\}} p_{jk}C_{ij}D_{ijk} \tag{1}$$

$$\text{s.t. } \sum_{i=3}^{8}(T_{ij}-D_{ijk}) \le Q_{jk}+\epsilon_{(j-1),k}, \qquad j=0,\dots,35,\ k=\text{min, most, max}, \tag{2}$$

$$\epsilon_{(j-1),k} = \left[Q_{(j-1)k} - \sum_{i=3}^{8}\left(T_{i,(j-1)}-D_{i,(j-1),k}\right)\right] + \epsilon_{(j-2),k}, \qquad j=0,\dots,35,\ k=\text{min, most, max}, \tag{3}$$

$$\sum_{j=Y-12}^{Y-1} T_{ij}-D_{ijk} \ge R_i \sum_{i^*=4}^{8}\sum_{j=Y-12}^{Y-1} T_{i^*j}-D_{i^*jk}, \qquad Y=12,24,36,\ i=4,\dots,8,\ k=\text{min, most, max}, \tag{4}$$

$$\sum_{j=Y-12}^{Y-1} T_{3j}-D_{3jk} \ge R_3, \qquad Y=12,24,36,\ k=\text{min, most, max} \tag{5}$$

$$T_{8j}-D_{8jk} = \sum_{i=0}^{2} T_{ij}-D_{ijk}, \qquad j=0,\dots,35,\ k=\text{min, most, max} \tag{6}$$

$$T_{ij\text{cap}} \ge T_{ij} \ge D_{ijk} \ge 0, \qquad i=0,\dots,8,\ j=0,\dots,35,\ k=\text{min, most, max}. \tag{7}$$

The objective function (1) is maximizing the expect "profit" of our model. Note that the expenditures and revenue are represented by the net benefit/cost parameters, which are random parameters.

The constraint (2) is ensuring that for each period (and scenario of that period) we do not allocate more water than what we received that period, as well as what we had left over from the previous.

The constraint (3) initializes the surplus variable.

The constraint (4) is a constraint we decided to add on from the original formulation. This constraint ensures we are meeting legal requirements of how water must be allocated to other states. Note that the objective function only applies to Colorado users. We consider other users for the matter of meeting legal requirements in this objective function. Furthermore, note that this proportion is between the states. We are not considering tribal water rights when meeting this requirement.

The constraint (5) is a constraint we also added to meet legal requirements of allocating water to upper basin tribes. This requires a separate constraint since tribes are promised a fixed amount of water, rather than a proportion of.

(6) establishes a relationship between the water the entirety of Colorado receives, and the total amount of water each individual user within Colorado receives.

The constraint (7) is leading to a problem I discuss later. We are considering $T_{ij\text{cap}} = \infty$, $\forall i, \forall j$. This constraint is ensuring that we do not deficit any user more than what we promised them that period (as in, we cannot take away from their supply). Note that this upper bound's name was changed from the paper to not confuse it with the "max" scenario.

Finally, the solver used for this model was Gurobipy, due to the size of the problem.

4 Problems with the model

The model described gives a bounded objective value and is feasible. Problem - We're missing a constraint. Currently, we have a constraint that says, hey, each of these users need this much water from the river every year. This leads to a feasible solutions that allocates all the water to the user for the year in the month with highest benefit.

Furthermore, note that the agricultural sector is not receiving any water at all.

In terms of the non-Colorado users, I'm not too concerned about this, seeing that they are not included in the objective function. The purpose of including those users was to account for the fact that the Upper Basin serves multiple states, not just Colorado.

I was hoping that this net benefit / cost parameter would prevent this from happening. There are a couple of things I'm considering to resolve this:

1. **Adjust the hyperparameters for the net benefit / cost prior distributions**
   - In fact, maybe adjust the distributions we're using. I'm still trying to understand why the original paper chose to use a uniform distribution for the municipal parameters, but the normal distribution for industrial / agriculture. I'm wondering if we should make them all normal, but make the standard deviation for the municipal user so that the distribution is much more concentrated around the mean. I feel like there should be little room for error for that user.
   - I almost want to use more extreme parameters as well. In context of the scale of our units - Acre-feet - a decimal difference in the mean of the distributions is significant already, though.
2. **Find data on monthly water demands**
   - This is tricky, though. It was difficult enough as it is to find annual data on water allocation for specifically Colorado. I broke down the NB and C parameters by month already. Higher demand months have higher NB and higher C, lower demand months have lower NB and lower C. But I suppose writing that out explains to me why the solver chose to allocate all the water to the month with the highest NB.
3. **Make up data?**
   - As mentioned before, the point of this project is to get a functioning model. The data can be substituted in wherever necessary. I'm wondering if I should incorporate in monthly demands

into the "demand data.csv" file. I'm not so sure where to start with that. Maybe base it off proportions from the annual demand? I still feel like that new data and therefore new constraint would make the NB and C parameters redundant. At that point I feel like we're just maximizing some constant.

4. **Choose a $T_{ij\mathbf{cap}}$**
   - If we cap how many units of water can be allocated each period, this might be able to prevent all the water being allocated at once each year.

5 Github

In the Github repository[1], you will find a folder of all the code used to clean the water flow data and implement the model. You will also find a folder with all the data used for this project. In the main branch, there is also a `README` file that will explain the purpose of each individual file in the repository.

6 Appendix of Results

Any variable not included was simply zero. See python output on Github for full solution.

```
Optimal 'profit': 88079888.19335274
Final allocation (ac-ft):
  T[(0, 32)] = 1022952895.4412675
  T[(1, 10)] = 211228318.62000006
  D[(1, 10, 'most')] = 211228318.62000006
  D[(1, 10, 'min')] = 211228318.62000006
  T[(1, 15)] = 149236192.52999994
  D[(1, 15, 'min')] = 149236192.52999994
  T[(3, 0)] = 1152281.0
  D[(3, 0, 'most')] = 1152281.0
  T[(3, 7)] = 1152281.0
  D[(3, 7, 'max')] = 1152281.0
  D[(3, 7, 'min')] = 1152281.0
  T[(3, 13)] = 1152281.0
  D[(3, 13, 'max')] = 1152281.0
  T[(3, 16)] = 1152281.0
  D[(3, 16, 'most')] = 1152281.0
  D[(3, 16, 'min')] = 1152281.0
  T[(3, 25)] = 1152281.0
  T[(4, 7)] = 93879252.72000004
  D[(4, 7, 'most')] = 93879252.72000001
  D[(4, 7, 'min')] = 93879252.72000004
  T[(4, 15)] = 66327196.67999998
  D[(4, 15, 'min')] = 66327196.67999998
  D[(4, 17, 'max')] = 3.802900513013204e-09
  D[(4, 25, 'most')] = -4.718701044718425e-08
  T[(4, 31)] = 9.153570447649275e-08
  D[(4, 31, 'min')] = 9.153570447649275e-08
  T[(4, 32)] = 454645731.30723006
  D[(5, 2, 'most')] = -7.96798202726576e-09
  T[(5, 11)] = 57143892.96000003
  D[(5, 11, 'most')] = 57143892.96000003
  D[(5, 11, 'min')] = 57143892.96000003
```

[1] https://github.com/makhoullillian153/Optimally-Allocating-Water-Resources-from-the-Colorado-River/

```
T[(5, 21)] = 40373076.23999999
D[(5, 21, 'min')] = 40373076.23999999
D[(5, 22, 'most')] = 7.450580596923828e-09
T[(5, 31)] = 7.472176482712013e-08
D[(5, 31, 'max')] = 3.70796336684116e-08
D[(5, 31, 'min')] = 7.472176482712013e-08
T[(5, 35)] = 276740879.92614
T[(6, 9)] = 45919199.70000002
D[(6, 9, 'most')] = 45919199.70000002
D[(6, 9, 'min')] = 45919199.70000002
T[(6, 16)] = 32442650.54999999
D[(6, 16, 'min')] = 32442650.54999999
D[(6, 19, 'max')] = 3.743005673522534e-09
D[(6, 26, 'most')] = 1.7654636631841244e-08
D[(6, 32, 'min')] = 5.960464477539063e-08
T[(6, 35)] = 222381064.22636256
T[(7, 1)] = 3.079090348551845e-10
D[(7, 4, 'min')] = 3.079090348551845e-10
T[(7, 7)] = 2877336.000000001
D[(7, 7, 'most')] = 2877336.000000001
D[(7, 7, 'min')] = 2877336.000000001
T[(7, 16)] = 2032883.9999999993
D[(7, 16, 'min')] = 2032883.9999999993
D[(7, 17, 'max')] = -2.315707817052339e-10
T[(7, 27)] = 4.486474591096834e-09
D[(7, 27, 'min')] = 4.486474591096834e-09
D[(7, 29, 'max')] = 7.884800831951461e-10
T[(7, 35)] = 13934586.099
T[(8, 10)] = 211228318.62000006
D[(8, 10, 'most')] = 211228318.62000006
D[(8, 10, 'min')] = 211228318.62000006
T[(8, 15)] = 149236192.52999994
D[(8, 15, 'min')] = 149236192.52999994
T[(8, 32)] = 1022952895.4412675
```