

An Open Source Hydroeconomic Model for California's Water Supply System: PyVIN

H53A-1659

Mustafa S. Dogan*, Ellie White, Jon Herman, Quinn Hart, Justin Merz, Josue Medellin-Azuara, and Jay R. Lund Civil & Environmental Engineering, University of California, Davis

*Corresponding author: msdogan@ucdavis.edu

1 Open source database with existing CALVIN data: hydrology & economics

HOBBS Database

- ~1250 nodes
- ~600 conveyance links
- 49 surface reservoirs
- 38 groundwater reservoirs
- 88% of CA's irrigated acreage
- 92% of CA's urban population

<https://cwn.casil.ucdavis.edu>

2 Input: a HOBBS tool converts time-series data into matrix format

Time-series + **Demand Curves** = **Data Matrix**

Demand Curves: Price (\$/Vol) vs Q* (volume)

Data Matrix (Example):

id	cost	amplitude	lower_bound	upper_bound		
HU402A.1965-05-31	CVPM14AG.1965-05-31	0	-1760	1	0	32.83
HU402A.1965-05-31	CVPM14AG.1965-05-31	1	-1619	1	0	2.98
HU402A.1965-05-31	CVPM14AG.1965-05-31	2	-1450	1	0	2.98
HU402A.1965-05-31	CVPM14AG.1965-05-31	3	-1199	1	0	2.99
HU402A.1965-05-31	CVPM14AG.1965-05-31	4	-927.5	1	0	2.98
HU402A.1965-05-31	CVPM14AG.1965-05-31	5	-782.2	1	0	2.99
HU402A.1965-05-31	CVPM14AG.1965-05-31	6	-722.7	1	0	2.98
HU402A.1965-05-31	CVPM14AG.1965-05-31	7	-614.8	1	0	2.99
HU402A.1965-05-31	CVPM14AG.1965-05-31	8	-500.4	1	0	2.98
HU402A.1965-05-31	CVPM14AG.1965-05-31	9	-234.2	1	0	2.98
HU402A.1965-06-30	CVPM14AG.1965-06-30	0	-1764	1	0	53.70
HU402A.1965-06-30	CVPM14AG.1965-06-30	1	-1618	1	0	4.88
HU402A.1965-06-30	CVPM14AG.1965-06-30	2	-1446	1	0	4.89
HU402A.1965-06-30	CVPM14AG.1965-06-30	3	-1202	1	0	4.88
HU402A.1965-06-30	CVPM14AG.1965-06-30	4	-926.6	1	0	4.88
HU402A.1965-06-30	CVPM14AG.1965-06-30	5	-784	1	0	4.88
HU402A.1965-06-30	CVPM14AG.1965-06-30	6	-719.9	1	0	4.88
HU402A.1965-06-30	CVPM14AG.1965-06-30	7	-615	1	0	4.89
HU402A.1965-06-30	CVPM14AG.1965-06-30	8	-499.9	1	0	4.88
HU402A.1965-06-30	CVPM14AG.1965-06-30	9	-223.9	1	0	4.88
HU402A.1965-07-31	CVPM14AG.1965-07-31	0	-1763	1	0	54.33
HU402A.1965-07-31	CVPM14AG.1965-07-31	1	-1763	7	0	24.3
HU402A.1965-07-31	CVPM14AG.1965-07-31	2	-1763	7	0	24.3
HU402A.1965-07-31	CVPM14AG.1965-07-31	3	-1763	7	0	24.3
HU402A.1965-07-31	CVPM14AG.1965-07-31	4	-1763	7	0	24.3
HU402A.1965-07-31	CVPM14AG.1965-07-31	5	-1763	7	0	24.3
HU402A.1965-07-31	CVPM14AG.1965-07-31	6	-1763	7	0	24.3
HU402A.1965-07-31	CVPM14AG.1965-07-31	7	-1763	7	0	24.3
HU402A.1965-07-31	CVPM14AG.1965-07-31	8	-1763	7	0	24.3
HU402A.1965-07-31	CVPM14AG.1965-07-31	9	-1763	7	0	24.3

3 PyVIN network flow optimization model

- Linear programming network flow model with gains & losses

$$\text{minimize } Z = \sum_i \sum_j \sum_k c_{ijk} X_{ijk}$$

where
 Z: total cost
 X: flow on the arc
 c: unit cost (or penalty)
 b: external flow
 a: amplitude
 l: lower bound
 u: upper bound

subject to $\sum_i \sum_k X_{ijk} = \sum_j \sum_k a_{ijk} X_{ijk} + b_j$ for all nodes j

$$X_{ijk} \leq u_{ijk} \text{ for all arcs}$$

$$X_{ijk} \geq l_{ijk} \text{ for all arcs}$$

- An abstract model in Pyomo: model structure and data are separate

Model Formulation + Data = PyVIN

- Uses freely available state-of-the-art solvers and not solver specific

Installed Solvers:

- CPLEX
- CBC
- GUROBI
- GLPK

- The source code and data are publicly available on GitHub

4 Solver runtime comparison for different model sizes and solvers

- Runs are performed on UC Davis College of Engineering's HPC1 (High Performance Computer)
- CBC, CPLEX and GUROBI are run in parallel, and GLPK is run in serial

time (second) vs number of decision variables

Solvers: cbc, cplex, glpk, gurobi

Regression equations:

- glpk: $Y = 9.83E-10X^{2.19551}$
- cbc: $Y = 5.59E-09X^{1.83019}$
- gurobi: $Y = 3.77E-08X^{1.52444}$
- cplex: $Y = 7.84E-08X^{1.41141}$

5 Total calculation time comparison

preprocess (Pyomo builds the model) → solver (Solver solves the model) → postprocess (Results are written in .json or .yaml)

- Differences get bigger as model size increases
- cplex and gurobi have similar total calculation time

# of decision variables	output file size (json)
35,827	~7.5 MB
1,430,915	~300 MB
2,933,335	~602 MB

time (second) vs number of decision variables

Solvers: cbc, cplex, glpk, gurobi

Regression equations:

- glpk: $Y = 2.88E-08X^{1.96947}$
- cbc: $Y = 3.11E-05X^{1.27684}$
- gurobi: $Y = 4.13E-04X^{1.02861}$
- cplex: $Y = 3.46E-04X^{1.03667}$

6 Comparing preliminary results from PyVIN to existing CALVIN results

Average storage comparison of selected reservoirs (TAF)

Reservoirs: Trinity, Shasta, Folsom, New Don Pedro, Millerton

PyVIN vs CALVIN storage (TAF):

- Trinity: 1516 vs 1487
- Shasta: 3468 vs 3482
- Folsom: 606 vs 598
- New Don Pedro: 1672 vs 1683
- Millerton: 260 vs 259

- PyVIN model is still under development and minimum in-stream flow requirements are not fully represented yet
- Small differences on selected reservoirs' average storage

Monthly average flow differences between CALVIN and PyVIN (TAF/m)

Frequency vs Error (TAF/m)

7 Postprocessed preliminary results from PyVIN

Some useful results the model provides

- Agricultural & urban supply, water scarcity, cost & WTP
- Water operations & delivery reliabilities
- Surface and groundwater storage
- Hydropower generation & revenue
- Shadow prices of environmental flows
- Marginal benefits of expanding capacity

Regional urban and agricultural water supply portfolios (TAF/m)

Regions: USV, LSVD, SJSB, TB, SC

Legend: GWP, SWD, NPR, PR, DESAL, urban, ag

- Regional urban and agricultural water supply portfolios (TAF/m)

8 References & Acknowledgements

- Draper, A. J., Jenkins, M. W., Kirby, K. W., Lund, J. R., & Howitt, R. E. (2003). Economic-engineering optimization for California water management. *Journal of water resources planning and management*, 129(3), 155-164.
- Pyomo: <http://www.pyomo.org>
- PyVIN repository: <https://github.com/msdogan/pyvin>
- HOBBS repositories: <https://github.com/ucd-cws/calvin-network-tools> & <https://github.com/ucd-cws/calvin-network-data>

Authors would like to thank to Mimi Jenkins, Max Fefer, Gracie Yao, and Natalie Mall for their help and contribution.