

CALVIN & PyVIN Shortcourse

Mustafa S. Dogan ^{*},

April 1, 2017

Contents

1 Prerequisites	3
1.1 Command line	4
1.1.1 Windows	4
1.1.2 Mac OS	5
2 Introduction and background	6
3 HOBES database	8
3.1 Data flow overview	8
4 Updated version: PyVIN	9
4.1 Model architecture	9
5 Your first model run	12
5.1 Postprocessing results	12
6 References and useful links	12

^{*}Graduate Student, Dept. of Civil and Env. Eng., Univ. of California, Davis, 1 Shields Avenue, Davis, CA 95616. E-mail: msdogan@ucdavis.edu.

Agenda and Topics

- | | |
|---------------|--|
| 15 min | - Introduction and set-up |
| 1 hr | - CALVIN theory and model introduction |
| 15 min | - Break |
| 15-20 min | - HOBBES database |
| 15 min | - Data flow overview |
| 1 hr | - Break |
| 20-25 min | - PyVIN updates and model architecture |
| 15 min | - A PyVIN example |
| 15 min | - Break |
| 15-20 min | - Required software and installation |
| 20 min | - Your first PyVIN run |
| 20-25 min | - Postprocessing results |
-

1 Prerequisites

Please bring your **laptop** with below software dependencies installed if you want a hands-on PyVIN experience. Install following software in advance since some of them, such as Anaconda, takes long time to download and install. If you have any question, e-mail msdogan@ucdavis.edu

- **Python v3 with Anaconda**

Link: <https://www.continuum.io/downloads>

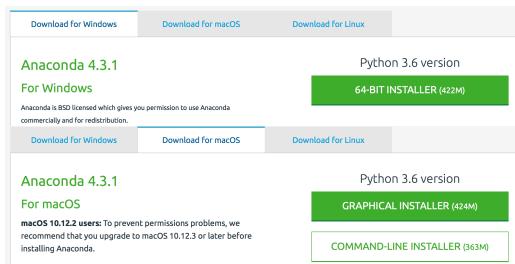


Figure 1: Anaconda package for Windows and Mac OS

Note: Download Python v3+ (not v2.7) because Pyomo command line installation works only with Python v3+

- **GitHub**

Link: <https://desktop.github.com>

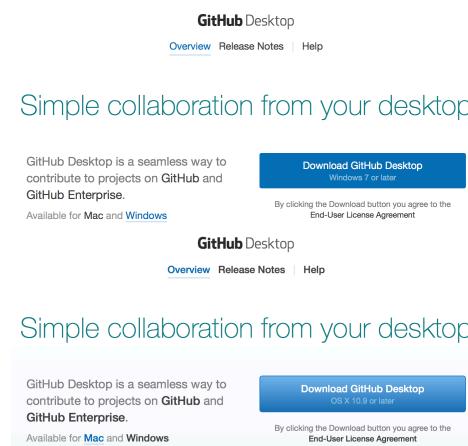


Figure 2: GitHub installation for Windows and Mac OS

- **Pyomo**

Command: `conda install -c cachemeorg pyomo pyomo.extras glpk`

Note: if you don't know how to use command line, please refer to description below and then type the command.

1.1 Command line

This shortcourse is not intended to explain command line use. It assumes some but not extensive command line knowledge. If you like to learn basics of command line, here is a nice crash course:

<https://learnpythononthehardway.org/book/appendixa.html>

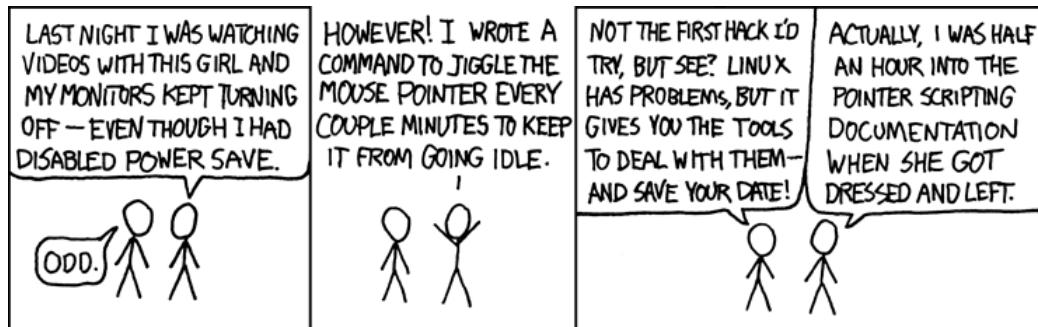


Figure 3: A command line comic from xkcd, source: <https://xkcd.com/196/>

1.1.1 Windows

After installing GitHub, double click on "Git Shell" and then type the command above to install Pyomo and GLPK solver.

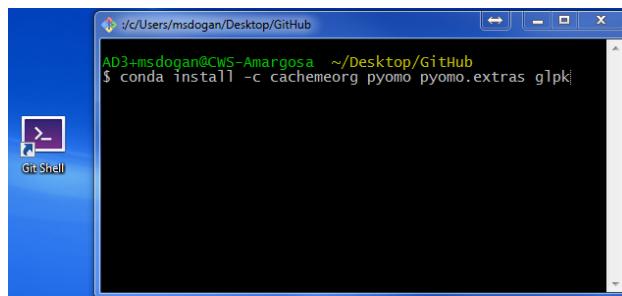


Figure 4: Command line (Git Shell) for Windows

1.1.2 Mac OS

Search "terminal" in Spotlight Search and then double click. Type the command above to install Pyomo and GLPK solver.

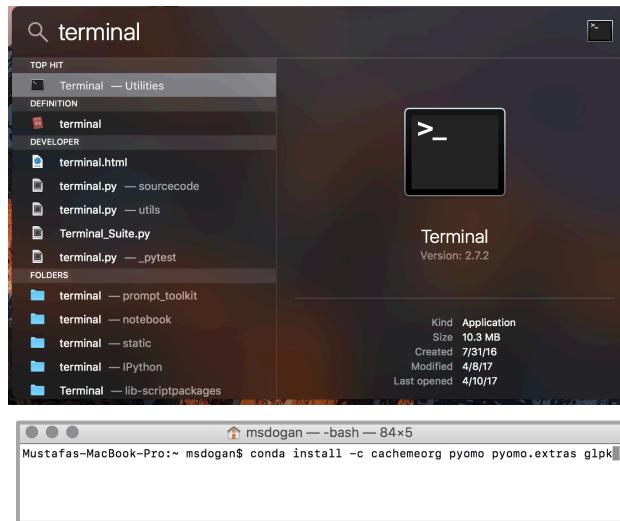


Figure 5: Command line (terminal) for Mac OS

2 Introduction and background

Developed in early 2000s, CALifornia Value INtegrated model (CALVIN) combines ideas from economics and engineering optimization with advances in software and data to suggest more integrated management of water supplies regionally and throughout California. CALVIN is an hydro-economic optimization model for California's advanced water infrastructure that integrates the operation of water facilities, resources, and demands, and it aims to optimize surface and groundwater deliveries to agricultural and urban water users. It allocates water to maximize statewide agricultural and urban economic value, considering physical and policy constraints. It replicates water market operations transferring water from users with lower willingness-to-pay (WTP) to users with higher WTP. CALVIN uses historical hydrology and 2050 water demand projections for its operations. Figure 6 regions and coverage of the model.

CALVIN forces quantitative understanding of integrated water and economic system. Motivation for the CALVIN effort include:

- making better sense of integrated system and operations
- seeking ways to improve system management
- quantifying user willingness to pay for additional water
- finding insights into changes in physical capacities and policies

With the recent updates, CALVIN is evolved to open-source PyVIN model, optimizing water resources in a short amount of time with state-of-the-art solvers and modeling platform. PyVIN has the same input data and objective as CALVIN but it is modeled on Pyomo, a Python-based high level algebraic modeling language. PyVIN also uses HOBES database, which allows better documentation, collaboration and communication between models as well as modelers.

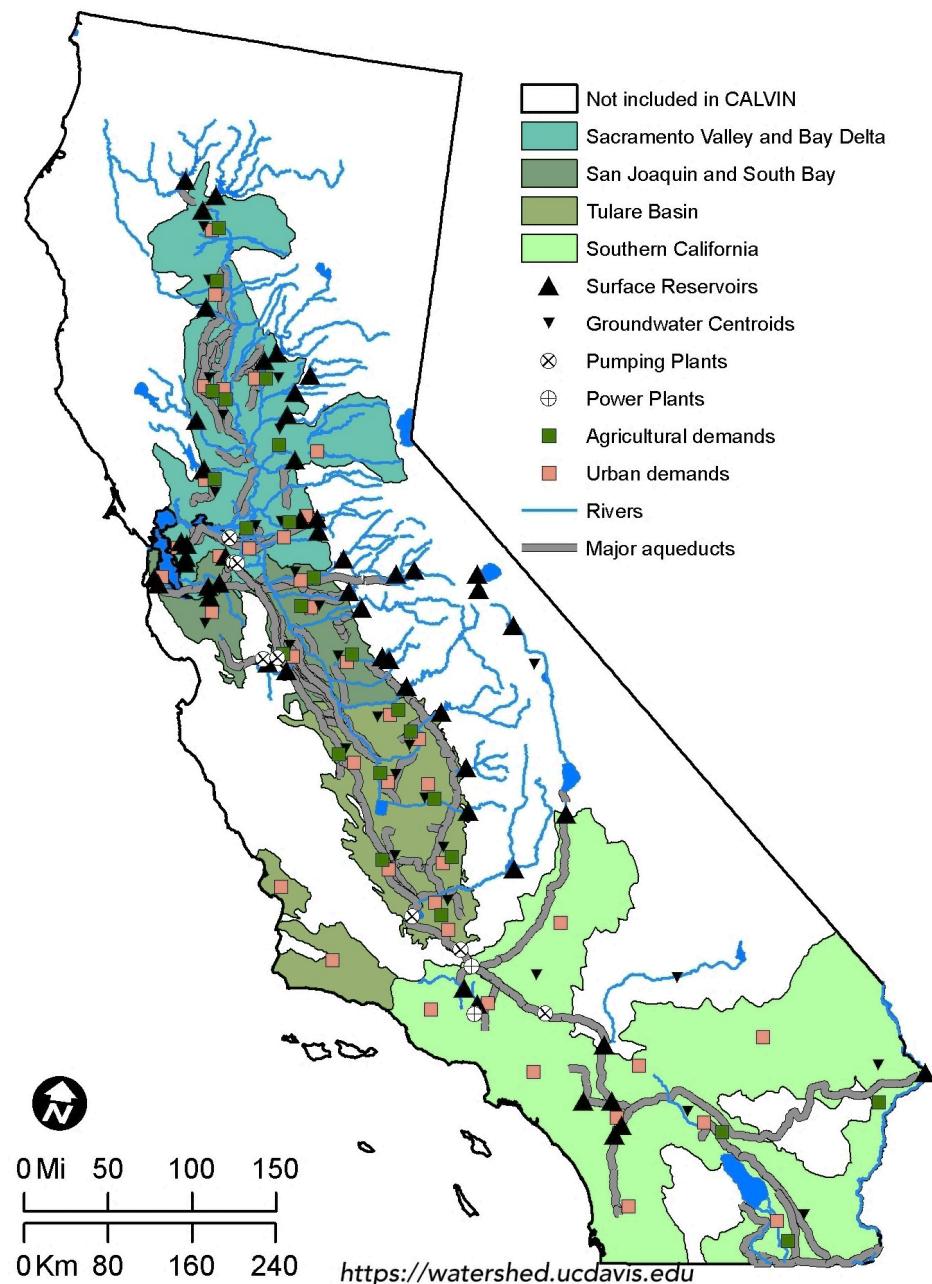


Figure 6: California's water infrastructure and CALVIN's coverage

3 HOBGES database

HOBGES serves as a cross-platform for data storage, display and documentation. It is a framework for database that aims to better organize data and makes model integration and communication easier by using common format and metadata. Classical approach in modeling is that first model is built and then required data are collected. But HOBGES reverses this order; it serves as a data hub and models are built on top of this database. HOBGES uses GitHub to keep track of changes and documentation. It also has a animation tool to display data.

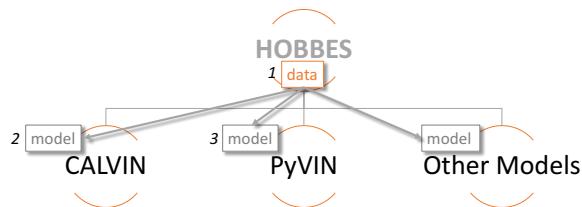


Figure 7: HOBGES database and model integration

3.1 Data flow overview

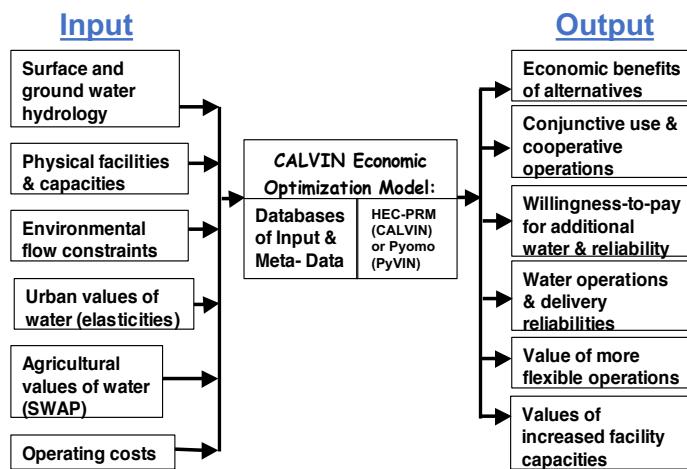


Figure 8: CALVIN (and PyVIN) data overflow with input (left column) and output (right column)

4 Updated version: PyVIN

PyVIN combines CALVIN's knowledge and extensive water infrastructure and hydrology data with a high level algebraic modeling language, Pyomo, and state-of-the-art solvers, such as CPLEX, Gurobi, CBC, and GLPK. Pyomo is an open-source Python based large-scale modeling environment. Its model representation is similar to GAMS and AMPL and it solves the problem with user defined solvers. So, PyVIN is not solver specific and users can choose and install any solver as long as it is compatible with Pyomo.

CALVIN PyVIN	
Large-scale hydroeconomic model	
Optimize water allocation to agricultural and urban users	
Minimize statewide water scarcity and operating costs	
HEC-PRM and VBA based	Pyomo and Python based
Less flexible (limited to HEC-PRM)	More flexible (full LP)
Solver runtime: ~16 hr (depending on initial solution)	Solver runtime: ~1 min (depending on solver)
Requires 32 bit Windows PC	Any computer
HEC-DSS database	Open source: data and source code

Figure 9: CALVIN (and PyVIN) data overflow with input (left column) and output (right column)

4.1 Model architecture

Modeled in different environments, both CALVIN and PyVIN solve the same objective function, subject to physical and environmental constraints. The objective is to minimize statewide water scarcity and operating costs (Equation 1). There are three constraints: upper bound (Equation 2), lower bound (Equation 3), and mass balance (Equation 4).

$$\max(z) = \sum_i \sum_j \sum_k c_{ijk} X_{ijk} \quad (1)$$

$$X_{ijk} \leq u_{ijk} \quad (2)$$

$$X_{ijk} \geq l_{ijk} \quad (3)$$

$$\sum_j \sum_i X_{ijk} = \sum_i \sum_j a_{ijk} X_{ijk} + b_j \quad (4)$$

where c represents unit cost, X is flow, u is upper bound, l is lower bound, a is amplitude to represent losses, and b is a local inflow to node j .

PyVIN is modeled as an abstract model in Pyomo, separating model structure from data. PyVIN can solve any data size and data do not have to be water resources. As long as it is a network flow problem, such as transportation or transmission, PyVIN can solve it. Also, since it is open-source and has full linear programming features, it has more flexible model representation. Any other constraints can easily be added in addition to three constraint mentioned before. The Pyomo structure code with parameters and equations is shown below.

```

1 from __future__ import division
2 from pyomo.environ import *
3 import itertools
4
5 model = AbstractModel()
6
7 # Nodes in the network
8 model.N = Set()
9
10 # Network arcs
11 model.k = Set()
12
13 model.A = Set(within=model.N*model.N*model.k)
14
15 # Source node
16 model.source = Param(within=model.N)
17 # Sink node
18 model.sink = Param(within=model.N)
19
20 # Flow capacity limits
21 model.u = Param(model.A)
22 # Flow lower bound
23 model.l = Param(model.A)
24 # Link amplitude (gain/loss)
25 model.a = Param(model.A)
26 # Link cost
27 model.c = Param(model.A)
28
29 # The flow over each arc
30 model.X = Var(model.A, within=Reals)
```

```

31
32 # Minimize total cost
33 def total_rule(model):
34     return sum(model.c[i,j,k]*model.X[i,j,k] for (i,j,k) in
35 model.A)
36 model.total = Objective(rule=total_rule, sense=minimize)
37
38 # Enforce an upper bound limit on the flow across each arc
39 def limit_rule_upper(model, i, j, k):
40     return model.X[i,j,k] <= model.u[i,j,k]
41 model.limit_upper = Constraint(model.A, rule=limit_rule_upper)
42
43 # Enforce a lower bound limit on the flow across each arc
44 def limit_rule_lower(model, i, j, k):
45     return model.X[i,j,k] >= model.l[i,j,k]
46 model.limit_lower = Constraint(model.A, rule=limit_rule_lower)
47
48 # Enforce flow through each node (mass balance)
49 def flow_rule(model, node):
50     if node in [value(model.source), value(model.sink)]:
51         return Constraint.Skip
52     outflow = sum(model.X[i,j,k]/model.a[i,j,k] for i,j,k in
53 model.A)
54     inflow = sum(model.X[i,j,k] for i,j,k in model.A)
55     return inflow == outflow
56 model.flow = Constraint(model.N, rule=flow_rule)

```

The data file `data.dat` includes list of nodes, and list of links with properties. All links have cost c , amplitude a , lower bound l , and upper bound u . Below is an example data file.

```

1 set N :=
2 INITIAL SR.SHA.1983-10-31 SR.SHA.1983-11-30 FINAL
3 INFLOW.1983-10-31 INFLOW.1983-11-30 ...;
4
5 set k := 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15;
6
7 param source := SOURCE;
8 param sink := SINK;
9
10 param: A: c a l u :=
11 SOURCE INITIAL 0 0 1 0 10000000
12 INITIAL SR.SHA.1983-10-31 0 0 1 3686.84 3686.84
13 SR.SHA.1983-10-31 SR.SHA.1983-11-30 0 -7.003 0.997 630.4 630.4
14 SR.SHA.1983-10-31 SR.SHA.1983-11-30 1 -2.974 0.997 737.4 737.4
15 SR.SHA.1983-10-31 SR.SHA.1983-11-30 2 -1.466 0.997 632.2 2032.2

```

```
16 SR.SHA.1983-11-30 SR.SHA.1983-12-31 0 -6.972 0.999 609.4 609.4
17 SR.SHA.1983-11-30 SR.SHA.1983-12-31 1 -3.056 0.999 700.8 700.8
18 SR.SHA.1983-11-30 SR.SHA.1983-12-31 2 -1.479 0.999 689.7 1941.7
19 FINAL SINK 0 0 1 0 10000000
20 SR.SHA.1984-09-30 FINAL 0 0 1 2923.297 2923.297
21 SOURCE INFLOW.1983-10-31 0 0 1 0 10000000
22 INFLOW.1983-10-31 SR.SHA.1983-10-31 0 0 1 301.765 301.765
23 SOURCE INFLOW.1983-11-30 0 0 1 0 10000000
```

5 Your first model run

5.1 Postprocessing results

6 References and useful links