[A map of a reservoir storage area

Description automatically generated](https://www.bewaterwise.com/reservoir-water-storage/2.2.4_reservoir_storage.pdf)

Figure 1 Major Reservoirs in SoCal (source: https://www.bewaterwise.com/reservoir-water-storage/2.2.4\_reservoir\_storage.pdf)

Variables:

* Gw elevation ind
* Gw pumping intensity ind
* SWDI SC
* SWDI SL
* SWDI delta
* SWDI Colorado
* SWDI MEA
* SWDI PWL

# SWP

* SWDI mead drastically influencing SWP and when we remove it, we observe a drop in performance. From r2 0.78 to 0.55
* Drastic improvement is observed when using dummy variable
  + 1 variable\*: around r2 = 0.85
  + 2 variables: around r2= 0.9

\*1 variable on top of dummy variable

* Log of dependent variable (Y) shows some improvement r2 around 0.65 compared to 0.55 before
  + No improvement is seen when using dummy variable

## With all variables

A screenshot of a computer

Description automatically generated

## Removing some variables

* Removing:
  + SWDI MEA
  + SWDI PWL
  + SWDI SL
  + SWDI Colorado
* Keeping:
  + GW elevation ind
  + Gw pumping intensity ind
  + SWDI SC
  + SWDI delta



## Dummy variable

Just be using dummy variable as input (X), the results are: r2=0.74 and adjusted r2=0.73

### 2 variables: dummy variable and one other variable

A screenshot of a computer

Description automatically generated

### 3 variables: dummy variable and 2 other variables

A screenshot of a computer

Description automatically generated

### Log dependent variable

No dummy variable

A screenshot of a computer

Description automatically generated

W dummy variable

A screenshot of a computer

Description automatically generated

# Colorado

* SWDI delta highly influences Colorado Deliveries \*
  + With Log SWDI delta only, r2 = 0.74
  + GW pumping indicator and SWDI SC also show noteworthy correlation (around r2=0.5)

\*I’m ignoring SWDI SL variable

* Log of dependent variable (Y) shows a slight decrease performance (from r2 around 0.87 compared to 0.84)

## 1 variable

A screenshot of a table

Description automatically generated

## 2 variables

A screenshot of a computer

Description automatically generated

## Log dependent variable

A screenshot of a table

Description automatically generated

# Groundwater

* GW pumping intensity is the only variable that has some correlation with groundwater deliveries \*
  + With Log GW pumping int ind only, r2 = 0.58

\*I’m ignoring SWDI SL variable

* W 2 variables, the highest correlation is r2=0.68
* W 3 variables, the highest correlation is r2=0.73
* W 4 variables, the highest correlation is r2=0.93 (this is completely overfitting)
* Log of dependent variable (Y) shows no difference in performance

## 1 variable

A screenshot of a computer

Description automatically generated

## 2 variables

A screenshot of a computer

Description automatically generated

## 3 variables

A screenshot of a computer

Description automatically generated

## 4 variables

A screenshot of a computer

Description automatically generated

## Log dependent variable

3 variables

A screenshot of a computer

Description automatically generated

2 variables

A screenshot of a computer

Description automatically generated