



Photo: The Nevada Sagebrush  
http://nevadasagebrush.com/blog/2017/11/14/pollution-becoming-problem-for-water-clarity-in-lake-tahoe/

## Introduction

### Motivation:

- California draught December 2011 to March 2017 was the driest since record keeping began in 1895<sup>[1][2]</sup>
- Water resources management is essential to optimize the water supply, hydropower generation, and flood management
- Forecasting Lake Tahoe's water levels is important as most of California's water supply originates in northern California.

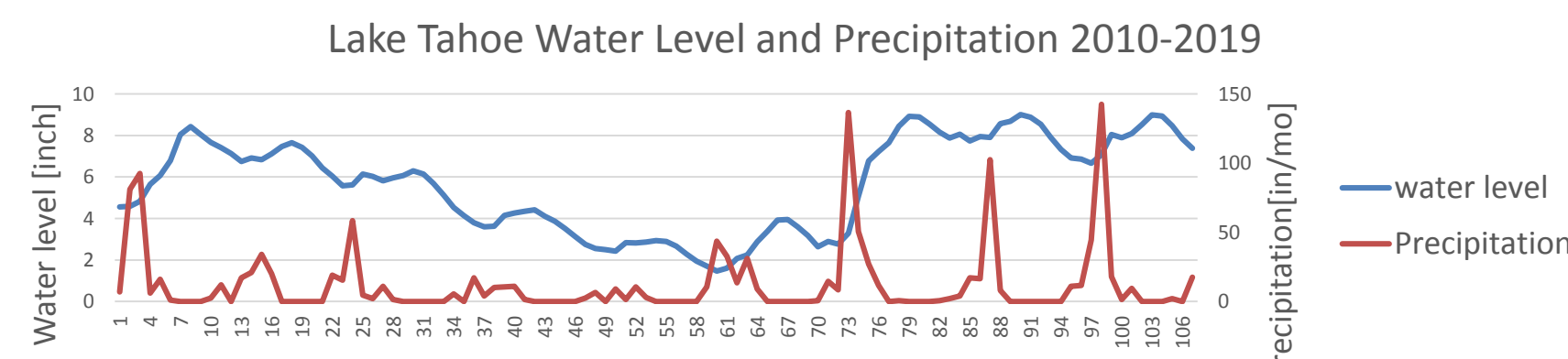
### Objective:

- Develop a model to predict Lake Tahoe's water level with metrological data

## Data and Features

### Data:

Dataset	Units	Data range
Lake Water Level	in	Monthly data: Jan 1911 -Nov 2019 <sup>[3]</sup>
Precipitation	in/month	Monthly total data: Jan 1910 - Nov 2019 <sup>[4]</sup>
Snow Fall	in/month	Monthly total data: Jan 1910 - Nov 2019 <sup>[4]</sup>
Water Discharge	ft <sup>3</sup> /sec	Monthly mean data: Jan 1910 - Nov 2019 <sup>[5]</sup>



### Features:

The mass conservation equation to model the amount of water in the lake is expressed as

$$w(t) = w(t - 1) + \Delta w_{in}(t) - \Delta w_{out}(t)$$

$\Delta w_{in}(t)$  consists of 12 months of precipitation data and 12 months of snow fall data.

$\Delta w_{out}(t)$  consists of the water discharge of the current month.

The feature vector is expressed as

$$\phi^{(i)} = [w(t - 1), d(t - 1), r(t - 1), r(t - 2), \dots, r(t - 12), s(t - 1), s(t - 2), \dots, s(t - 12)]$$

## Models

### Regression models:

- Linear Regression, SVR, Decision Tree Regression, Kernel Ridge Regression, Neural Network are compared

### Model Selection:

- k-fold cross validation (k=10)
- Choose the model with the lowest average mean squared error regression loss
- Train data: Jan 1911- Dec 2010, Test data: Jan 2011 – Nov 2019

### Feature Selection:

- the univariate linear regression tests was used

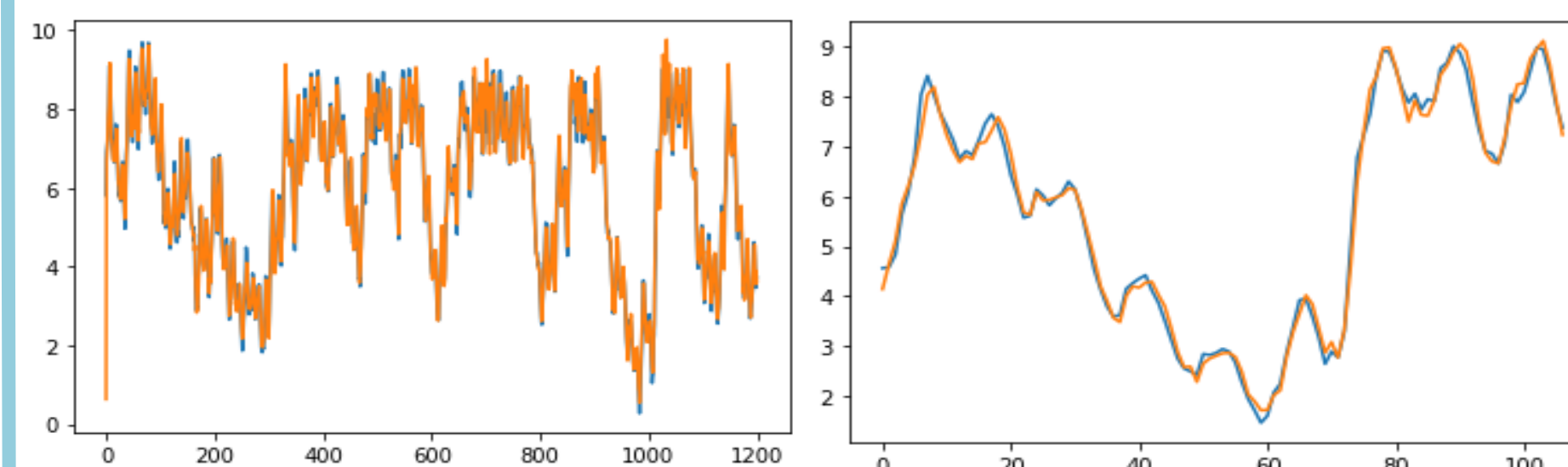
## Results

### Feature Selection:

- The precipitation and snow fall data between 5-7 months ago are the most correlated features
- It would typically take around half a year for rain to flow out to the Lake Tahoe.
- Snow pack that accumulates in winter slowly melts through summer with feeding the reservoir.

### Model Selection:

- Linear regression, SVR with linear kernel, and Kernel ridge with linear kernel models predicted the water level well.

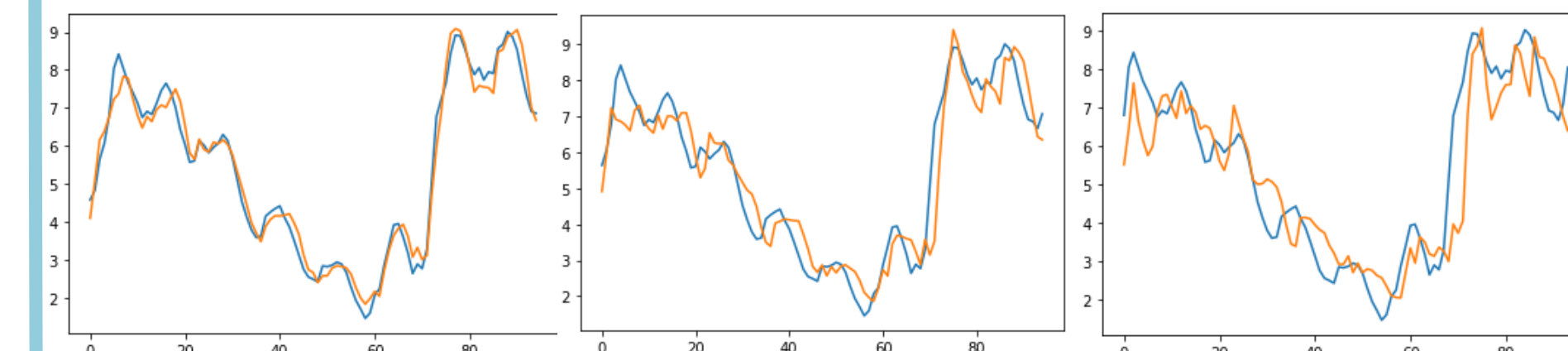


- Kernel ridge regression results. (Left) Train with error 0.08560 (blue: observed, orange: trained), (Right) Test with error 0.04540 (blue: observed, orange: predicted)
- The regularization worked well to prevent overfitting to the train set

### Predicting water level at 1-6months later:

- Kernel ridge shows good prediction
- The error increases as it predicts later month

Model	Average mean squared error regression loss of the prediction for M months ahead					
	1	2	3	4	5	6
Linear regression	0.08561	0.19315	0.32636	0.44292	0.53753	0.63852
SVR with linear kernel	<b>0.08556</b>	0.19847	0.33945	0.45790	0.54911	0.65777
Decision tree regression	0.19623	0.45903	0.69082	0.91972	1.11523	1.10458
Kernel ridge regression	0.08560	<b>0.19282</b>	<b>0.32593</b>	<b>0.44253</b>	<b>0.53727</b>	<b>0.63831</b>
Multi-layer Perceptron regressor	0.09123	0.19620	0.33405	0.46594	0.54342	0.64720



The Kernel ridge regression results on the test set predicting 2 months (left), 4 months (middle), and 6 months (right) ahead the input data (blue: observed, orange: predicted).

## Summary and Future Work

- Linear regression, SVR with linear kernel, and Kernel ridge with linear kernel models predicted the water level well.
- This would mean that the hydraulic system around the Lake Tahoe would be explained by a linear model in respect to the features
- For future development, additional data set such as temperature, humidity, sun radiation, and precipitation at other stations around the Lake Tahoe, may improve the prediction.

### Reference:

- [1] Los Angeles Times. 2017. "Gov. Brown Declares California Drought Emergency Is over - Los Angeles Times."
- [2] PPIC. 2016. "California's Latest Drought - Public Policy Institute of California."
- [3] USGS. 2019a. "USGS Current Conditions for USGS 10337000 LAKE TAHOE A TAHOE CITY CA."
- [4] Western Regional Climate Center. 2019. "TAHOE CITY, CA, Total of Precipitation (Inches), Monthly Precipitation Listings, Monthly Totals."
- [5] USGS. 2019b. "USGS Surface Water Data for USA: USGS Surface-Water Monthly Statistics."