# Minimizing Thermal Impacts of Hydropower using Reinforcement Learning

Isabel Bush
ibush@stanford.edu

Matthew Shultz
shultzm@stanford.edu

## Introduction

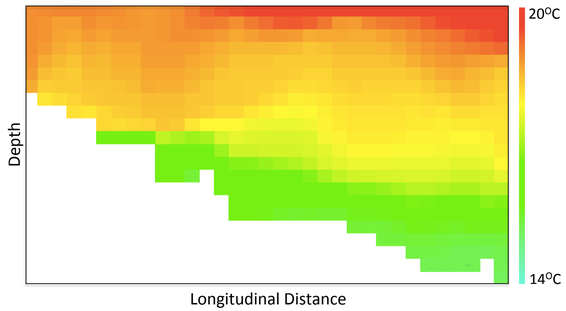
Thermal impacts of hydropower dams, combined with heat waves under a warming climate, have led to high die-off rates for salmon on the Columbia River, who cannot tolerate high water temperatures.

Dams often have two or three outlets from top to bottom, with the lower outlets passing cooler water and the upper outlets passing warmer water due to stratification in reservoirs. We apply Q-learning to identifying a policy for managing both reservoir elevations and water temperatures by controlling flow through a combination of the outlets.

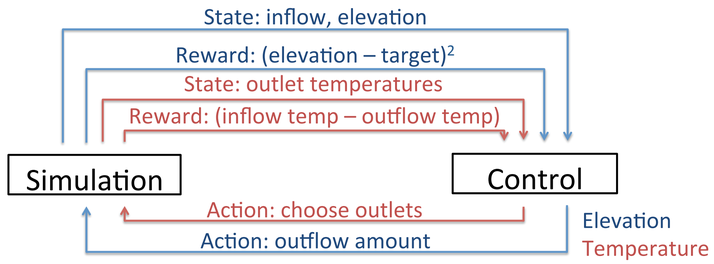
## Simulation

We use the 2D hydrodynamical model CE-QUAL-W2[1] to simulate the system, along with a bathymetry model of the Lower Granite Dam on the Snake River[2], and hydrological and meteorological time-series data from the area[3,4].

Temperature Stratification in the Simulated Reservoir




## System Model




## Q-Learning Algorithms

At each step of Q-learning, we update the expected utility of taking action a at state s, given that it resulted in a next state s' with reward r.

**Discretized Lookup Table**

We use a lookup table for stored Q-values, where the state and actions are discretized[5].

$$Q(s,a) \leftarrow Q(s,a) + \alpha(r + \gamma \max_{a' \in A} Q(s',a') - Q(s,a))$$

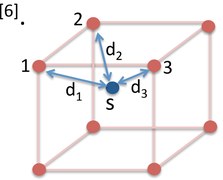
**Weighted K Nearest Neighbors**

We calculate continuous-state Q-values as weighted sums of Q-values for the KNN of points spanning the state space[6].

$$w_i = \frac{1}{1+d_i^2} \quad \forall i \in knn$$

$$V(a) = \sum_{i=1}^{knn} Q(i,a) \frac{w_i}{\sum w_i}$$

$$Q(i,a) \leftarrow Q(i,a) + \alpha(r + \gamma \max_{a' \in A} V'(a') - V(a))p(i) i \in knn$$



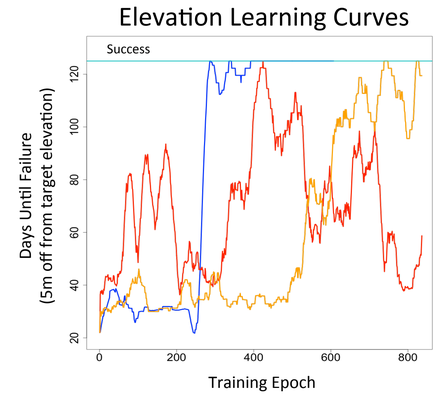

**Linear Function Approximation**

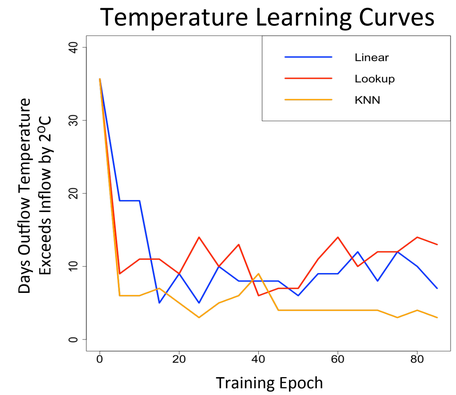
We approximate the Q-value as a weighted sum of state-action features, given by discretized state and an indicator for the action.

$$Q(s,a;\theta) = \theta \cdot \phi(s,a)$$

$$\theta \leftarrow \theta - \alpha[Q(s,a;\theta) - (r + \gamma \max_{a' \in A} Q(s',a';\theta))]\phi(s,a)$$

## Results




## Discussion

Our work applies Q-learning to an atypical control target, a hydropower dam, and demonstrates that it can be used to find a policy for compliance with mandated thermal loading targets. The greatest challenge we faced was the substantial size of the continuous state space and the large computational expense to repeatedly run hydrodynamic simulations. The KNN algorithm addressed this challenge somewhat by compactly representing the full continuous space, while Lookup and Linear algorithms had to rely on more coarsely discretized features.

## Future Work

Further work should combine both elevation and temperature objectives during training, which may give more flexibility in reaching water temperature goals but will require handling even larger state spaces. Future work implementing a continuous linear model may improve both convergence and generalization by lifting the requirement for discretization and incorporating fewer state variables over all.

[1] T. Cole and S. Wells. *CE-QUAL-W2: A Two-Dimensional, Laterally Averaged, Hydrodynamic and Water Quality Model*. U.S. Army Corps of Engineers and Portland State University, 4.0 edition, 2016
[2] B. Cope. Temperature simulation of the snake river above lower granite dam using transect measurements and the ce-qual-w2 model. *EPA*, 2002.
[3] NCDC. Quality controlled local climatological data, 2016.
[4] USGS. Water data for the nation, 2016.
[5] C. J. C. H. Watkins. Learning from delayed rewards. *PhD thesis, University of Cambridge, England*, 1989.
[6] J.Martin, J.deLope, and D.Maravall The knn-td reinforcement learning algorithm. June 2009.