International Congress on Environmental Modelling and Software

9th International Congress on Environmental Modelling and Software - Ft. Collins, Colorado, USA - June 2018

Jun 26th, 3:40 PM - 5:20 PM

Deep Reinforcement Learning for Optimal Operation of Multipurpose Reservoir Systems

Matthew E. Peacock
Colorado State University - Fort Collins, mpeacock86@gmail.com

John W. Labadie Dr.
Colorado State University - Fort Collins, john.labadie@colostate.edu

Follow this and additional works at: https://scholarsarchive.byu.edu/iemssconference

Deep Reinforcement Learning for Optimal Operation of Multipurpose Reservoir Systems

Matthew E. Peacock[a], John W. Labadie[a]
[a] Department of Civil and Environmental Engineering, Colorado State University
mpeacock@rams.colostate.edu, labadie@engr.colostate.edu

Abstract: Dynamic programming (DP) is considered the ideal optimization method for solving multipurpose reservoir system operational problems since it realistically addresses their complex nonlinear, dynamic, and stochastic characteristics. The only drawback to DP is the so-called "curse of dimensionality" that has plagued the method since its inception by Richard Bellman in the 1950's. Dimensionality issues arise from the need to discretize the state-action space and random variates which leads to an explosion in computational and memory requirements with increased state-space dimensionality. DP also requires development of spatial-temporal stochastic hydrologic models for reservoir system operations, which may be difficult under complex climatic and meteorological conditions. A deep reinforcement learning algorithm is applied to solving DP problems for reservoir system operations which effectively overcomes dimensionality issues without requiring any model simplifications, or sacrificing any of the unique advantages of DP. The algorithm uses an iterative learning process which considers delayed rewards without requiring an explicit probabilistic model of the hydrologic processes. The algorithm is executed in a model-free stochastic environment whereby the algorithm implicitly learns the underlying stochastic behavior of the system for developing dynamic, optimal feedback operating policies. Dimensionality issues are addressed through use of accurate function approximators for the state-value and policy functions based on deep neural networks. The deep reinforcement learning algorithm is applied to developing optimal reservoir operational strategies in the Upper Russian River basin of Northern California in the presence of multiple noncommensurate objectives, including flood control, domestic and agricultural water supply, and environmental flow requirements.

Keywords: approximate dynamic programming, reinforcement learning, stochastic optimization, machine learning, curse of dimensionality, multipurpose reservoir operations