

12/5/2025

Questions for the authors of the Winter Run Chinook Machine Learning Model from the review panel (Rich Zabel, Anna Sturock, and Nancy Monsen)

Is there more documentation available? We have some specific questions about which predictive parameters were included in the full model, how lags and windows were determined, and which model selection routine was used to eliminate parameters, among other questions

How did the authors choose the particular approach to use (i.e., the XGB-DART approach)?

Can you please clarify this statement: “We used an 80%/20% training/testing data split and divided the dataset into slices, resulting in five, 1,225-day slices comprised of 980 training days followed by 245 testing days.” What are the day slices composed of? Did you split the data into training/testing for each model run?

Is there a github site where the data and code are stored?

Why did the model perform more poorly when applied to the genetically determined Winter run Chinook?

How specifically is the model used in management? How would you recommend that it be used?

Was the logistic regression model presented for illustrative purposes, or is it used in management? How?

Is the Tillotson et al. model also used or was it abandoned? If it was abandoned, why was it?