Mathematical Model Descriptions of the Inseason Projection for Canadian-origin Yukon Chinook Salmon

Aaron Lambert

6/20/2022

**Preseason Forecast (Prior to be updated)**

Where  is a draw of for the Canadian-origin runsize in year y (the year of interest) from a normal distribution with a mean equal to the natural log of the preseason point-estimate for year , and a standard deviation equal to the empirical performance of the preseason forecast from years that the preseason forecast is available (2007-2010,2013-2021).

The lognormal standard deviation is calculated as

where is the lognormal forecast error in each year (2007-2010,2013-2021) and   is the geometric mean of the annual forecast errors.

**Model Version 1.0**

In model version 1.0, updating the projection based on the preseason forecast   using inseason information is performed by estimating the relationship between Pilot Station Sonar (PSS) daily cumulative passage and the end of season (EOS) reconstructed Canadian-origin runsize in previous years. This relationship is then used to predict the current years EOS Canadian-origin runsize.

In model 1.0, cumualtive historic PSS passage on day is regressed against historic EOS Canadian-origin runsize in normal space where

where is the intercept coefficient, is the slope coefficient, and is the cumulative PSS passage up to day in year .

Slope and intercept parameters are estimated by minimizing the difference between projected Canadian-origin runsize  and the true, observed Canadian-origin runsize where,

and is the residual standard deviation for predictions based on PSS passage through day and is directly estimated.

Priors on parameters are to be uninformative and bounded at zero where,

The final projection is then calculated by updating the preseason forecast estimate   with the PSS prediction as

**Model Version 2.0**

Version 2.0 is similar to version 1.0, though PSS passage is regressed against the natural log of EOS Canadian-origin passage as

and,

and the priors on the regression parameters remain uninformative and are

**Model Version 2.1**

Version 2.1 is the same as 2.0 in all aspects of the calculations except that a lognormal correction is subtracted from the PSS prediction where,

**Model Version 2.2**

Version 2.2 is the same as 2.0 in all aspects of the calculations except that a lognormal correction is added from the PSS prediction where,

**Model Version 3.0**

In version 3.0, GSI information is incorporated in the projection during data preprocessing prior to running the Stan model as,

where   is the mean proportion of Canadian-origin Chinook salmon passing PSS calculated from the mean over each strata and mean strata dates as determined by GSI sampling in prior years. This serves as a comparison for further models that incorporate GSI information in the Stan model.

The adjusted PSS passage () is used in place of as in model version 2.0 with no lognormal correction.

**Model Version 3.1**

In version 3.1 GSI is directly incorporated into the Stan model using a Beta distribution. This version relies on heavily restricted GSI parameter restrictions to reparametrize the mean and sd GSI proportions to Beta distribution shape parameters.

To incorporate GSI data, the EOS Canadian origin passage is estimated as

Where is the proportion of Chinook passing PSS in year and on day . In this version, The proportion   is calculated as the mean proportion over all years with GSI information in each strata (pulse) where each strata is defined by mean end and start day and the corresponding standard deviation  . The proportion is multiplied by , the number of Chinook salmon passing PSS on day in year

The likelihood for the GSI proportions is

where the shape parameters are reparametrized as the mean and standard deviation where

To simulate a smooth transition throughout the season from a high proportion of Canadian-origin Chinook at the start of the run, to a relatively low proportion at the end of the run, a random walk is generated using a beta distribution where the first day of the run is drawn as

and the following days are drawn with the previous days () proportion as the mean and a standard deviation drawn as

The GSI proportion for each day up to the current day is drawn from a beta distribution that is

and reparametrized in terms of and where

**Model Version 3.2**

In version 3.2, the likelihood is calculated the same as before using the reparametrized beta distribution in terms of mean GSI proportions calculated from mean observed GSI proportions from each strata defined by mean start and end days.

However, in version 3.2, the random walk is generated using a normal distribution and using an inverse logit transformation to coerce the value into beta space [0,1] where

and,

where,

Next,

**Model Version 3.3**

In version 3.3, everything is the same as version 3.2 except that the likelihood is from a normal distribution where

**Model Version 3.4**

This model is identical to version 3.2 except that the empirical GSI mean   and standard deviation   are calculated as the mean and standard deviation of GSI proportions across all years for each day , ignoring strata.

**Model Version 3.5**

This model is identical to version 3.3 except that the empirical GSI mean   and standard deviation   are calculated as the mean and standard deviation of GSI proportions across all years for each day , ignoring strata.