Forecast Code “Library” Model Documentation

Brooke Davis

June 2021

I created a [repository](https://github.com/brookemdavis/Forecast-Development) with the intent to store a library of consistent stock-recruitment forecast models coded in JAGS, stan, and TMB that can be drawn from for future forecasting (and potentially status assessment) work. The JAGS versions of the models were derived from the current forecast versions found in “FC\_SUB\_SR\_Models.R”, but made more flexible by removing any hardwired numeric values, and changing some notation to try and create consistent notation across models and coding platforms.

**Basic Ricker Model**

The JAGS version of this code can be found in the “Functions” file, rather than in a standalone file (function is called “Ricker.model.MCMC”). The Stan and TMB versions can be found in their respective folders named “Single\_Stock\_Ricker.stan” and “Single\_Stock\_Ricker.cpp”. Note that the stan and TMB versions include the estimation of SMSY, with the stan version using the approximation found in Hilborn (1985), while the TMB version includes the exact solution outlined in Scheuerell (2016). I was unable to figure out how to implement the exact solution, using the Lambert W function in stan. This functionality may become useful for future benchmark/LRP work. The TMB version also includes some code for estimating Sgen, that is currently commented out, but could also be useful in future benchmark/LRP work. A script testing the three versions of this model can be found in the file “Testing\_Ricker\_Mods.R”.

The model is based on the linearized form of the Ricker model where is used rather than because is generally more intuitive to put a prior around, based on observed spawner numbers.

Where

Linearized by taking the natural logarithm of either side:

This model is fit using a lognormal likelihood, implemented by placing a normal likelihood on .

Priors are place on , , and as follows.

Lognormal prior on , implemented as normal prior on :

Lognormal prior on , implemented as normal prior on

Inverse gamma prior on variance, , implemented as a gamma prior on precision, in TMB and JAGS, since they don’t have built in inverse gamma distributions. Since stan has a built in inverse gamma distribution, we are able to put the prior directly on variance. Note that in TMB the parameterization of the gamma distribution is different (parameterized using shape and scale where scale is 1/rate) than it is in stan, JAGS, and R (shape and rate, where rate = 1/scale) so we actually provide as the scale parameter.

When implementing a Bayesian model in TMB or stan, a Jacobian adjustment is required if a prior is put on a “transformed” variable. In our case, when using TMB, since we are putting our gamma prior on precision, not standard deviation (), which is the parameter used in the model, we require this adjustment. The same applies to the stan version, except we are putting the prior on , so the Jacobian is slightly different. We have avoided needing this for our other parameters, since we have put priors on them in the same form that they appear in the likelihood. For more information see an example file and vignette I wrote with Cole Monnahan [here](https://github.com/brookemdavis/TMB-Priors).

in the TMB version of the model is calculated using the exact solution provided in Scheuerell (2016):

Where is the Lambert W function, which satisfies the following:

The TMB code to execute the Lambert W function was sourced from [here](https://kaskr.github.io/adcomp/lambert_8cpp_source.html).

In the stan version, I simply used the commonly used Hilborn (1985) approximation of :

The explicit solution for could be implemented in stan as well, I just didn’t prioritize figuring out how to do this.

For all models, in order to report estimated recruitment with prediction intervals, rather than confidence or credible intervals, a new variable is reported that has the likelihood equation applied to it.

Notations used in code.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Documentation | Meaning | TMB | stan | JAGS |
|  | Observed recruits, year i | logR = | R\_Obs | R\_Obs |
|  | Fitted recruits, year i | R\_Fit | R\_Fit | R\_Fit |
|  | Predicted recruits, year i | R\_Pred | R\_Pred | R\_Pred |
|  | Spawners, year i | S | S | S |
|  | Productivity parameter | logA = | logA = | logA = |
|  | Density-dependence parameter | B, =logSmax | = Smax | beta |
|  | SD on recruitment likelihood | logSigma = | sigma | = tau |
|  | Mean of normal prior on | logA\_mean | logA\_mean | logA\_mean |
|  | SD of normal prior on | logA\_sig | logA\_sig | = logA\_tau |
|  | Spawners at maximum productivity | = logSmax | Smax | Smax |
|  | Mean of normal prior on | logSmax\_mean | logSmax\_mean | logSmax\_mean |
|  | SD of normal prior on | logSmax\_sig | logSmax\_sig | = logSmax\_tau |
|  | gamma/inverse gamma distribution shape and rate parameter (assumed to be the same) | Sig\_Gam\_Dist | Sig\_Gam\_Dist | Sig\_Gam\_Dist |

**Ricker Single Environmental Covariate Model**

Currently in the forecast model, the only way to incorporate an environmental covariate is by adding a single covariate to either a Ricker or Power model. I had been working on replicating this simple JAGS model (RickerCov.model.MCMC in “Functions.R” file) in TMB (Single\_Stock\_Ricker\_Covariate.cpp) and stan (Single\_Stock\_Ricker\_Env.stan) but had not gotten to the point of running and testing it. Rather, I switched directions and was focusing on creating a more flexible version of the model that could incorporate more than one environmental covariate. The simple covariate model has the exact same setup as the model above, except with the addition of a single covariate, :

Where

We use a normal prior on the environmental covariate parameter, :

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Documentation | Meaning | TMB | stan | JAGS |
|  | Observed environmental covariate in year i | Env | env | env |
| g | Parameter estimating effect of covariate | g | g | g |
|  | Mean of prior on g | g\_mean | g\_mean | g\_mean |
|  | SD of prior on g | g\_sig | g\_sig | = g\_tau |

**Ricker Multiple Environmental Covariate Model**

This model is not currently in the code repository, and has so far only been written in TMB for use in the Chilliwack and Pink forecasts for 2021. The version in the Chilliwack analysis (Sockeye-Forecast-2021\Chilliwack-2021\Code\TMB\Single\_Stock\_Ricker\_Multi\_Env.cpp) should be the one used for future development, since it will provide forecast values for both age cohorts (by setting N\_Pred = 2). This model is set up exactly the same as the one above, but now is a matrix, with each column giving a different covariate, and each row representing the set of observed covariate in each year. In this model g is then a vector with a length representing the number of environmental covariates being used in the model. Similarly, and are vectors in this model.

Linearized likelihood equation:

With a normal prior on each :

|  |  |  |
| --- | --- | --- |
| Documentation | Meaning | TMB |
|  | Observed environmental covariate h, in year i | Env |
|  | Parameter estimating effect of covariate | g |
|  | Mean of prior on | g\_mean |
|  | SD of prior on | g\_sig |

**Basic Power Model**

A basic power model is also coded up in each platform, following the same format as the original forecast model. A script testing the three versions of this model can be found in the file “Testing\_Power\_Models.R”.

where

Normal priors are put on both A and B:

As with the Ricker model forms, an inverse gamma prior is put on variance, either directly (in stan) or by putting a gamma distribution on precision (). In both stan and TMB these transformations require a Jacobian adjustment.

In TMB it is common to fit strictly positive parameters in log-space, to help model stability. In the TMB version both A and B are limited to be strictly positive (since the model is parameterized using and ). These didn’t come up as a problem in model testing (ie. there didn’t seem to be discrepancies between models fit in TMB, stan, JAGS), but should be revisited as an assumption, since it is inconsistent between the three versions of the code.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Documentation | Meaning | TMB | stan | JAGS |
|  | Observed recruits, year i | logR = | R\_Obs | R\_Obs |
|  | Fitted recruits, year i | R\_Fit | R\_Fit | R\_Fit |
|  | Predicted recruits, year i | R\_Pred | R\_Pred | R\_Pred |
|  | Spawners, year i | S | S | S |
|  | Power parameter 1 | logA = | A | A |
|  | Power parameter 2 | logB = |  | B |
|  | SD on recruitment likelihood | logSigma = | sigma | = tau |
|  | Mean of normal prior on | A\_mean | A\_mean | A\_mean |
|  | SD of normal prior on | A\_sig | A\_sig | = A\_tau |
|  | Mean of normal prior on | B\_mean | B\_mean | B\_mean |
|  | SD of normal prior on | B\_sig | B\_sig | = B\_tau |
|  | gamma/inverse gamma distribution shape and rate parameter (assumed to be the same) | Sig\_Gam\_Dist | Sig\_Gam\_Dist | Sig\_Gam\_Dist |

**Power Model with Multiple Covariates**

Like the Ricker model, for the Chilliwack and Pink forecast for 2021, a flexible covariate model was developed in TMB. This model can be found in the Forecast repository in the folder: Fraser-Sockeye-Forecast-2021\Chilliwack-2021\Code\TMB, file name “Single\_Stock\_Power\_Multi\_Env.cpp”.

The linearized form of the likelihood equation is:

and the priors on ’s are exactly the same as the Ricker version outlined above:

|  |  |  |
| --- | --- | --- |
| Documentation | Meaning | TMB |
|  | Observed environmental covariate h, in year i | Env |
|  | Parameter estimating effect of covariate | g |
|  | Mean of prior on | g\_mean |
|  | SD of prior on | g\_sig |

**Larkin Model**

A Larkin model, assuming density-dependent effects for cohorts up to 3 years prior, was fit in all platforms as well, and tested using the R script “Testing\_Larkin\_Models.R”.

Linearized form:

A lognormal prior is put on , by putting a normal distribution on

Normal priors are put on each parameter, following the format used in the current forecast WinBUGS/JAGS model. This is one model choice that should be revisited, as these distributions may be quite informative, depending on the parameterization.

This prior form requires a Jacobian adjustment in the TMB version of this model, since the model is parameterized around .

An inverse gamma prior was put on variance:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Documentation | Meaning | TMB | stan | JAGS |
|  | Observed recruits, year i | logR = | R\_Obs | R\_Obs |
|  | Fitted recruits, year i | R\_Fit | R\_Fit | R\_Fit |
|  | Predicted recruits, year i | R\_Pred | R\_Pred | R\_Pred |
|  | Spawners, year i | S | S | S |
|  | Productivity parameter | logA = | logA = | logA = |
|  | Density-dependence parameters of lag h=0:3 | B, =logSmax | = Smax | beta0, beta1, beta2, beta3 |
|  | SD on recruitment likelihood | logSigma = | sigma | sigma, = tau |
|  | Mean of normal prior on | logA\_mean | logA\_mean | logA\_mean |
|  | SD of normal prior on | logA\_sig | logA\_sig | = logA\_tau |
|  | Mean of normal prior on | B\_means[h] | B\_means[h] | B\_means[h] |
|  | SD of normal prior on | B\_sigs[h] | B\_sigs[h] | = B\_taus[h] |
|  | gamma/inverse gamma distribution shape and rate parameter (assumed to be the same) | Sig\_Gam\_Dist | Sig\_Gam\_Dist | Sig\_Gam\_Dist |

**Models to Consider next**

For the recent Chinook RPA, several model forms were considered for Harrison, and fit in TMB. The code used for this process was written by myself, Catarina Wor, Kendra Holt, and Lauren Weir. It is held in a [repository](https://github.com/lhweir/Simple_Simulator) owned by Lauren Weir. You may need to ask her permission to access this repository, I’m unsure if it is public. There are several TMB models in here that would be useful starting points for additional models to be used for Sockeye Forecasting. Documentation for each of these models will be included in an Appendix of the Chinook RPA, which should be published in the coming months. You can also contact Lauren Weir for the final drafts of this documentation.

**Ricker Model with Time-varying Productivity**

There is a simple version of this model already coded in the forecast, but has not been used in formal forecasts before. A version of this model written in TMB, which has a better prior structure than the current JAGS/WinBUGS version can be found in the Chinook RPA (Simple\_Simulator) repository in “Harrison\_SR\_Analysis/R/Rickerkf\_ratiovar.cpp”. These models can be incredibly sensitive to priors, as it can dictate whether the model loads all of the variance either on changes in productivity, or process/observation error. I will send you a memo I wrote on this a few years ago.

**Ricker Model with Autocorrelated Residuals**

A version of this model written in TMB, can be found in the Chinook RPA (Simple\_Simulator) repository in “Harrison\_SR\_Analysis/R/Ricker\_autocorr\_ch.cpp”.

**Lognormal Correction Factor**

In the recent Chinook RPA, the issue of whether or not to use a lognormal correction factor came up. This made me realize that none of our stock-recruitment models use this. I think it is not cut-and-dry whether or not this should be used, but since we are using a lognormal likelihood, it is something that should be considered and explored in the future. Versions of the Basis Ricker model, plus versions with time-varying productivity, and autocorrelated residuals can be found in the Chinook RPA repository (versions with “\_LN\_Correct” added to the end of them).

**Works cited**

Scheuerell MD. 2016. An explicit solution for calculating optimum spawning stock size from Ricker’s stock recruitment model. PeerJ 4:e1623 <https://doi.org/10.7717/peerj.1623>

Hilborn R. 1985. [Simplified calculation of optimum spawning stock size from ricker stock recruitment curve](https://doi.org/10.1139%2Ff85-230). Canadian Journal of Fisheries and Aquatic Sciences 42(11):1833-1834