

# futR()

Kirstin Holsman

## Contents

Repo maintained by: Kirstin Holsman  
Alaska Fisheries Science Center  
NOAA Fisheries, Seattle WA  
**kirstin.holsman@noaa.gov**

---

## Overview

futR() is a generic Rpackage for fitting recruitment models to indices of spawning biomass (or numbers) and recruitment with or without climate covariates. The recruitment model is based on Template Model Builder (TMB) and standard stock-recruitment models modified to include the effects of environmental covariates (Deriso 1980; Schnute 1985). This includes Ricker (logistic), Beverton Holt, log-linear, and log-linear with biomass lagged by year ‘y-1’. The model can be fit with and with out random effects on spawning stock biomass (SSB) and recruitment (R) (i.e., measurement error on SSB and rec) using the methods of Porch and Lauretta (2016) and with the optional unbiased estimate of sigma (sensu Ludwid and Walters 1981, Porch and Lauretta 2016). Environmental covariates are optional and can be included to influence pre-spawning and post-spawning success. In all cases, parameters are estimated by minimizing the negative log-likelihood function:

$$-LL = 0.5 \left( n \ln(2 \frac{2}{R}) + \sum_{i=1}^{n_y} \left( \frac{\ln(\frac{R_i}{\bar{R}})}{\bar{R}} \right)^2 \right)$$

For more information see Holsman et al. 2020 Climate and trophic controls on groundfish recruitment in Alaska.

---

## futR() options

### Recruitment models (rectype)

There are four base model formulations available for estimating the relationship between recruitment and spawners (and/ or environmental factors). The simplest includes linear models, as well as Ricker and Beverton-Holt spawner recruitment relationships. These are invoked using the input of “rectype” in the model function.

1. Log-linear
2. Log-linear with spawners
3. Beverton Holt
4. Ricker

In addition, the main effect of environmental covariates is optional through estimating  $\beta$  (for pre-larval or spawning habitat effects) and  $\lambda$  (for post-spawning success and survival to recruitment age (e.g., age 0+ survival)):

- $\beta$  = vector  $(1, n_{cov})$  of covariate effects on pre-larval/ effective number of spawners
- $\lambda$  = vector  $(1, n_{cov})$  of covariate effects on post-spawning success (e.g., age 0+ survival)
- $\beta_0$  = binary  $(0, 1)$  vector  $(1, n_{cov})$  of inclusion in pre-larval effects
- $\lambda_0$  = binary  $(0, 1)$  vector  $(1, n_{cov})$  of inclusion in pre-larval effects post-spawning success

### rectype = 1. Log-linear

When rectype is set to 1, the model fits a log-normal linear model without an index of spawners (i.e., the mean recruitment as a function of annual changes in environmental covariates,  $X$ ). This model is best for evaluating recruitment variation when an index of spawners or spawning biomass is not available, or when recruitment is more strongly driven by environmental conditions than by spawning biomass (e.g., hatchery production). The general formulation is as follows:

$$\log(\hat{R}_i) = a + \sum_{k=1}^{n_k} \theta_i^\beta \beta_k X_{k,i} + \varepsilon_i$$

where  $\varepsilon_i$  represents a normally distributed independent random variable with mean 0 and variance  $\sigma_R^2$  (i.e.,  $\varepsilon_i \sim N(0, \sigma_R^2)$ ),  $a$  is the estimated intercept,  $X$  is a matrix of environmental covariates,  $\beta$  is a vector of estimated parameters representing covariate effects on mean recruitment, and  $\theta^\beta$  is a binary vector of length  $n_k$  of covariate inclusion. (i.e.,  $\theta^\beta \in \{0, 1\}$ ).

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & \dots & b_{1,n_y} \\ \vdots & \ddots & \vdots \\ x_{n_k,1} & \dots & x_{n_k,n_y} \end{bmatrix}, \quad = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{n_k} \end{bmatrix}$$

### rectype = 2. Log-linear with spawners

When rectype is set to 2, the model fits a log-normal linear model as a function of the index of spawners  $\hat{S}_i$  in the year  $i$  and annual changes in environmental covariates,  $X$ ):

$$\log(\hat{R}_i) = a + \sum_{k=1}^{n_k} \theta_i^\beta \beta_k X_{k,i} + (b + \sum_{k=1}^{n_k} \theta_i^\lambda \lambda_k X_{k,i}) * \hat{S}_i + \varepsilon_i$$

where  $\varepsilon_i$  represents a normally distributed independent random variable with mean 0 and variance  $\sigma_R^2$  (i.e.,  $\varepsilon_i \sim N(0, \sigma_R^2)$ ),  $a$  is the estimated intercept,  $b$  is the estimated productivity or slope,  $\mathbf{X}$  is a matrix of environmental covariates, where  $\beta$  and  $\lambda$  are vectors of estimated parameters representing covariate effects on pre- and post-spawning success, respectively, and  $\theta^\beta$  and  $\theta^\lambda$  are binary vectors of length  $n_k$  of covariate inclusion (i.e.,  $\theta^\beta \in \{0, 1\}$  and  $\theta^\lambda \in \{0, 1\}$ ). Note that in most cases, setting  $a = 0$  is recommended  $\hat{R}_i \sim 0$  as  $\hat{S}_i \rightarrow 0$

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & \dots & b_{1,n_y} \\ \vdots & \ddots & \vdots \\ x_{n_k,1} & \dots & x_{n_k,n_y} \end{bmatrix}, \quad = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{n_k} \end{bmatrix}, \quad = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_{n_k} \end{bmatrix}$$

### rectype = 3. Beverton Holt

The Beverton-Holt model is appropriate “if there is a maximum abundance imposed by food availability or space, or if the predator can adjust its predatory activity immediately to changes in prey abundance” (Wootton 1990, p. 264).

When rectype is set to 3, the model fits a Beverton-Holt (Beverton and Holt, 1957) model as a function of the index of spawners  $\hat{S}_i$  in the year  $i$  and annual changes in environmental covariates,  $X$ ). This model is most appropriate when there is a limit to the capacity of the system to support recruitment (e.g., limits to spawning habitat or limits to available food resources). Inclusion of environmental covariates on each of the terms in the model can therefore be used to model annual variation in spawning habitat and carrying capacity (covariates on the  $a$  term below) or environmental changes in other density dependent processes like predation or disease (covariates on the  $b$  term below):

$$\hat{R}_i = \frac{a\hat{S}_i e^{(\sum_{k=1}^{n_k} \theta_i^\beta \beta_k X_{k,i})}}{1 + b\hat{S}_i e^{(\sum_{k=1}^{n_k} \theta_i^\lambda \lambda_k X_{k,i})}} + e^{\varepsilon_i}$$

where  $\varepsilon_i$  represents a normally distributed independent random variable with mean 0 and variance  $\sigma_R^2$  (i.e.,  $\varepsilon_i \sim N(0, \sigma_R^2)$ ),  $a$  is the estimated productivity at lower spawning numbers and  $b$  is the density-dependent parameter; the asymptote of the slope is approximately  $a/b$  (Quinn II and Deriso 1999). As in other formulations,  $X$  is a matrix of environmental covariates,  $\beta$  and  $\lambda$  are vectors of estimated parameters representing covariate effects on pre- and post-spawning success, respectively, and  $\theta^\beta$  and  $\theta^\lambda$  are binary vectors of length  $n_k$  of covariate inclusion (i.e.,  $\theta^\beta \in \{0, 1\}$  and  $\theta^\lambda \in \{0, 1\}$ ).

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & \dots & b_{1,n_y} \\ \vdots & \ddots & \vdots \\ x_{n_k,1} & \dots & x_{n_k,n_y} \end{bmatrix}, \quad = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{n_k} \end{bmatrix}, \quad = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_{n_k} \end{bmatrix}$$

### rectype = 4. Ricker

When rectype is set to 4 the model fits a Ricker S/R curve. The Ricker model (Ricker, 1954) is a dome shaped recruitment curve best used to describe populations with strong density dependence (e.g., cannibalism). The model estimates recruitment in year  $i$  and a function of the index of spawners contributing to that year’s recruitment  $\hat{S}_i$  (e.g., spawning biomass in year  $i - 1$ ). If specified, the model also includes the additional effects of annual changes in environmental conditions  $X$ ; inclusion of environmental covariates on each of the terms in the model can therefore be used to model annual variation in spawning habitat and carrying capacity (covariates on the  $a$  term below) or environmental changes in other density dependent processes like cannibalism (covariates on the  $b$  term below):

$$\hat{R}_i = a\hat{S}_i e^{a-b\hat{S}_i} e^{\varepsilon_i}$$

the general Ricker formulation can be rewritten in linear form and expanded to include covariates (modified from Mueter et al. 2011):

$$\log \hat{R}_i = a \sum_{k=1}^{n_k} \theta_i^\beta \beta_k X_{k,i} - b\hat{S}_i \sum_{k=1}^{n_k} \theta_i^\lambda \lambda_k X_{k,i} + \ln \hat{S}_i + \varepsilon_i$$

where  $\varepsilon_i$  represents a normally distributed independent random variable with mean 0 and variance  $\sigma_R^2$  (i.e.,  $\varepsilon_i \sim N(0, \sigma_R^2)$ ),  $a$  is the estimated slope of the model near  $\hat{S}_i \sim 0$  and  $b$  is the density-dependence shape parameter.  $X$  is a matrix of environmental covariates,  $\beta$  and  $\lambda$  are vectors of estimated parameters representing covariate effects on pre- and post-spawning success, respectively, and  $\theta^\beta$  and  $\theta^\lambda$  are binary vectors of length  $n_k$  of covariate inclusion (i.e.,  $\theta^\beta \in \{0, 1\}$  and  $\theta^\lambda \in \{0, 1\}$ ).

$$\mathbf{X} = \begin{bmatrix} x_{1,1} & \dots & b_{1,n_y} \\ \vdots & \ddots & \vdots \\ x_{n_k,1} & \dots & x_{n_k,n_y} \end{bmatrix}, \quad = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{n_k} \end{bmatrix}, \quad = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_{n_k} \end{bmatrix}$$


---

## Obs. error options (sigMethod)

Finally, for more advanced model fitting, various options for observation error on recruitment or spawning estimates is available through alternative “sigMethod” selections:

1. No observation error ( $\tau = 0$ )
2. estimate sigma, random effects on SSB if  $\tau > 0$ ,  $\tau$  input
3. as in 1 but with unbiased sigma estimate,  $\tau$  input
4. as in 1 but with defined measurement error for rec (indep of random effects on Spawners/SSB)
5. as in 1 but with defined measurement error for rec and Spawners/SSB)

### sigMethod = 1

The base model fits the lognormally distributed process error for recruitment ( $\sigma$ ) using maximum likelihood (i.e., ordinary least squares). When sigMethod = 0, no observation error ( $\tau = 0$ ) is estimated for either spawners (S) or recruitment (R) estimates:

$$\log \hat{R}_i \sim N(0, \sigma_R) \text{ where } \sigma_R = \sigma$$

Similarly, spawner indices ( $\hat{S}_{i-1}$ ) include log-normally distributed observation error as:  $\log \hat{S}_{i-1} \sim N(0, \sigma_S)$  where  $\sigma_S = 0$

### sigMethod = 2

The base model fits the lognormally distributed process error for recruitment ( $\sigma$ ) using maximum likelihood (i.e., ordinary least squares). Additionally, when sigMethod = 1 and  $\tau$  is input ( $0 < \tau < 1$ ), observation errors are estimated for recruitment and spawners. Observation errors are statistically independent, random normal variables with similar variance but because of lack of other information  $\sigma_R$  is modeled as a function of process error via the  $\tau$  scalar (see eq 4 in Porch, C. E., and M. V. Laretta. 2016). Observation error for recruitment is modeled with log-normally distributed observation error as a function of  $\sigma$  as:  $\log \hat{R}_i \sim N(0, \sigma_R)$  where  $\sigma_R = (1 + \tau) * \sigma$

Similarly, spawner indices ( $\hat{S}_{i-1}$ ) include log-normally distributed observation error as:  $\log \hat{S}_{i-1} \sim N(0, \sigma_S)$  where  $\sigma_S = \tau * \sigma$

### sigMethod = 3

As in sigMethod = 2, except when sigMethod = 3  $\sigma$  is estimated using the unbiased sigma estimate (sensu Ludwig and Walters 1982):

$$\sigma_i = \frac{1}{(n_y - k)} * \left( \frac{\epsilon_{Ri}^2}{1 + \tau} + \frac{\epsilon_{S,i}^2}{\tau} \right)$$

where