Bayesian biomass dynamic model with kobe advice: the **bdm** and **kobe** packages

Charles T T Edwards & Laurie T Kell

June 2014

1   The **bdm** package

The package fits a state space biomass dynamic model using Bayesian methods, specifically the Hamiltonian MCMC implemented in the package **rstan**. To a large extent, the package is an external wrapper for **rstan**, providing functionality relevant to the intended application. The package is generalisable, meaning that any number of model formulations can be specified by the user. The default implements the Fletcher-Schaefer hybrid model [1], formulated in terms of the depletion x, which is the biomass relative to the carrying capcity K (equation 1b). The hybrid model is generic, in the sense that it allows a range of values for the depletion at MSY, or shape of the production curve relative to the biomass depletion. This is specified by the parameter ϕ, which is given as an input value with a default of $\phi = 0.5$ (i.e. $n = 2$), making it equivalent to the logistic model.

$$x_{t+1} = x_t + g(x_t) - \frac{C_t}{K} \tag{1a}$$

$$g(x_t) = \begin{cases} r.x_t.\left(1 - \dfrac{x_t}{2\phi}\right) & \text{if } x \leq \phi \\[2ex] \gamma.m.\left(x_t - x_t^n\right) & \text{if } x > \phi \end{cases} \tag{1b}$$

$$\phi = \left(\frac{1}{n}\right)^{(1/(n-1))} \tag{2a}$$

$$\gamma = \frac{n^{n/(n-1)}}{n-1} \tag{2b}$$

$$m = \frac{r\phi}{4} \tag{2c}$$

1.1 Estimation framework

Parameters are estimated within a Bayesian state-space framework. This re-formulates the process equation to include a time-dependent error term (the process error, ϵ^p) and a parallel observation process that relates an abundance index I to the unobserved biomass state with some degree of error (the observation error, ϵ^o), according to an estimated catchability scalar q.

$$x_{t+1} = [x_t + g(x_t) - H_t].\epsilon_t^p \tag{3a}$$

$$I_{it} = [q_i x_t].\epsilon_{it}^o \tag{3b}$$

The advantage of this class of models is that they allow both process and observation error to be represented simultaneously, which is important for effective precautionary or risk based management [2].

Parameters to be estimated in the model are therefore: r, K, q for each index, and the error terms. The large number of parameters necessitates a Bayesian approach with appropriate priors. Parametric distributional assumptions for ϵ^p and ϵ^o are required, which unfortunately cannot be estimated in an hierarchical manner with fisheries data. We assume them to follow a log-normal distribution with an expectation of one, and therefore fix values of σ_p and σ_o on input, based on a subjective measure of model fit to the data, with minimum bounds $\sigma_p \geq 0.05$ and $\sigma_o \geq 0.15$.

$$r \sim LN(\mu_r, \sigma_r^2) \tag{4a}$$

$$ln(K) \sim U(.,.) \tag{4b}$$

$$\epsilon_{..}^p \sim LN(-\sigma_p^2/2, \sigma_p^2) \tag{4c}$$

$$\epsilon_{..}^o \sim LN(-\sigma_o^2/2, \sigma_o^2) \tag{4d}$$

The r and K parameters of the logistic model are highly correlated, and their estimation is helped through the use of an informative prior or priors. We assumed an uninformative log-uniform prior for K, but an informative log-normal prior for r. The expectation and variance for the prior on intrinsic growth, with $E[r] = \exp(\mu_r + \sigma_r^2/2)$ can be constructed from available life-history data using the function `bdm::rcalc`, which implements methods described by McAllister et al. [3].

Log-normal prior distributions for the error terms are specified to have an expection of one, which gives an intuitive interpretation of the expected quantities.

$$E[x_{t+1}] = x_t + g(x_t) - H_t \tag{5a}$$

$$E[I_{it}] = q_i x_t \tag{5b}$$

The catchability q is estimated analytically from its maximum posterior density estimate assuming an uninformatinve uniform prior (i.e. $q \sim U(.,.)$).

$$\hat{q}_i = \exp\left[\frac{1}{n_t} \sum_t \{ln(I_{it}) - ln(x_t)\} + \frac{\sigma_o^2}{2}\right] \tag{6}$$
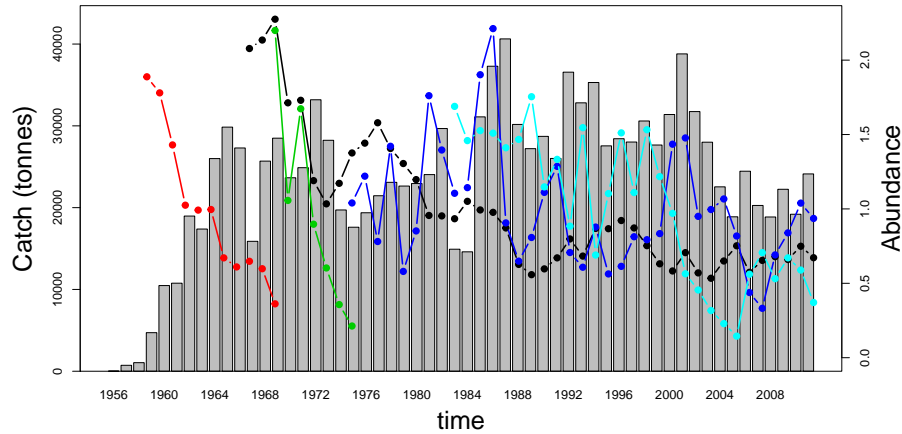
2

Figure 1: Data intputs for South Atlantic albacore

2 Application to albacore

2.1 Bayesian fit

We fit the default (logistic) model to albacore data (figure 1) from the South Atlantic. Trace and histogram plots for the posterior samples are shown in figures 2 and 3, and fits to the data in figure 5.

```
# load data into empirical data object
data(albsa)
dat <- edat(harvest = albsa$catch_tonnes, index = cbind(fleet1 = albsa$fleet1_cpue,
    fleet2 = albsa$fleet2_cpue, fleet3 = albsa$fleet3_cpue, fleet4 = albsa$fleet4_cpue,
    fleet8 = albsa$fleet8_cpue), time = rownames(albsa))

# initialise model
mdl <- bdm()

# update error terms and priors using values from Meyer and
# Millar (CJFAS 1999)
sigmao(dat) <- sqrt(0.0086/(1.71 - 1))
sigmap(dat) <- sqrt(0.0102/(3.79 - 1))
# intrinsic growth rate
mdl <- update_bdm(mdl, list(a = -1.38, b = 0.51, par = "r"))

# compile model
mdl <- compile_bdm(mdl)

# fit model
mdl <- fit(mdl, dat, iter = 5000, run = "run01")

traceplot(mdl, par = c("r", "logK"), inc_warmup = FALSE)
```
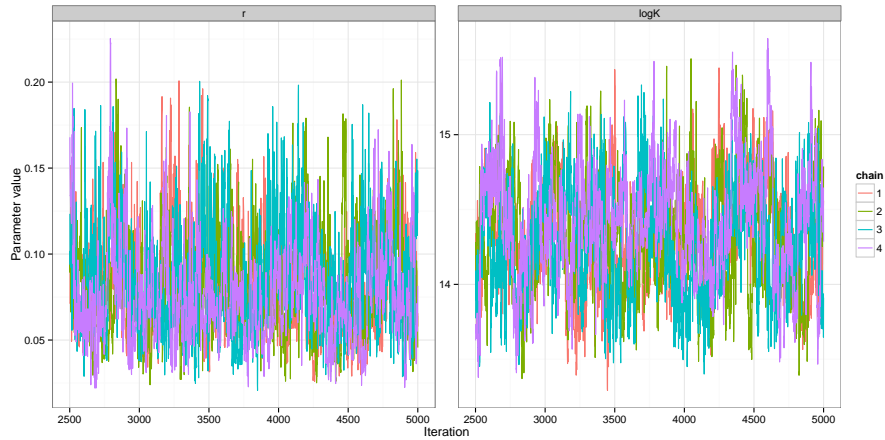
Figure 2: Trace plots for r and $ln(K)$

```
histplot(mdl, par = c("r", "logK"))
```

```
histplot(mdl, par = c("q"))
```

2.2 Kobe advice

It is possible to produce kobe plots by using the `bdm::as.kobe` function, which creates a `data.frame` in the appropriate format for the **kobe** package. The kobe phase plot for albacore is shown in figure 6. We also perform some simple projections using `bdm::project` and plot the distribution of values from the final year in figure 7. Predicted biomass dynamics are shown in figure 8.

```
# projections
harvest.time <- 30
harvest.scenarios <- c(0.01, 0.05, 0.1)
mdl.projections <- project(mdl, harvest.scenarios, harvest.time,
    harvest_rate = TRUE)
```
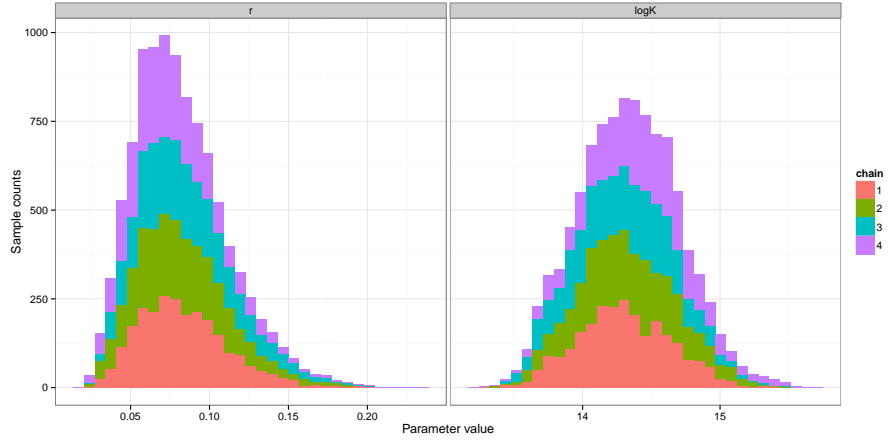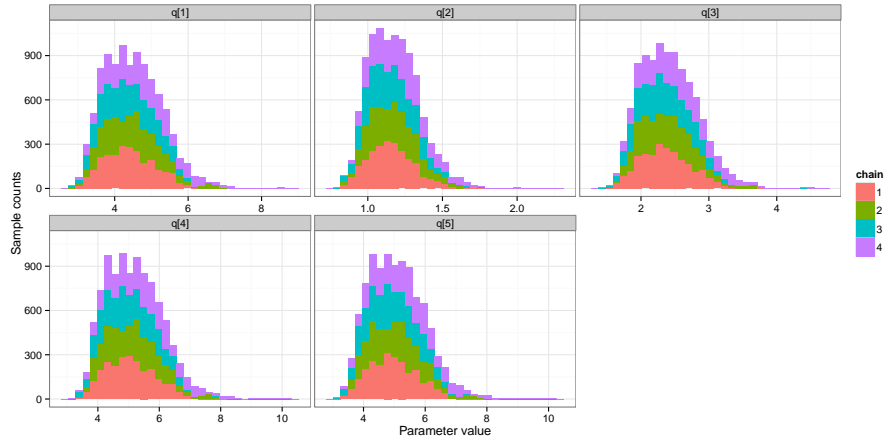
4

Figure 3: Posterior histograms for r and $ln(K)$

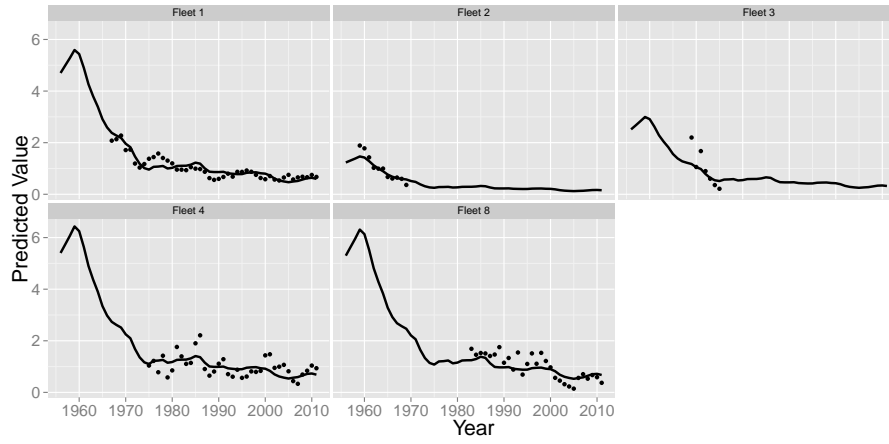**Figure 4:** Posterior histograms for catchability q

5

Figure 5: observed and predicted abundance indices for each fleet
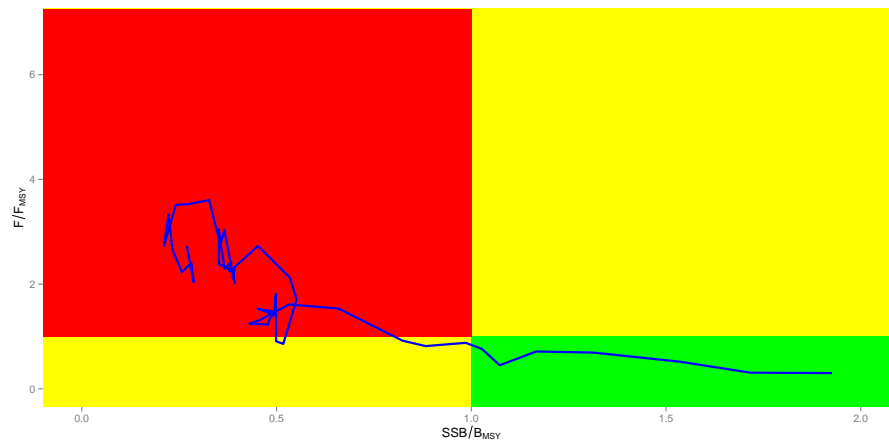

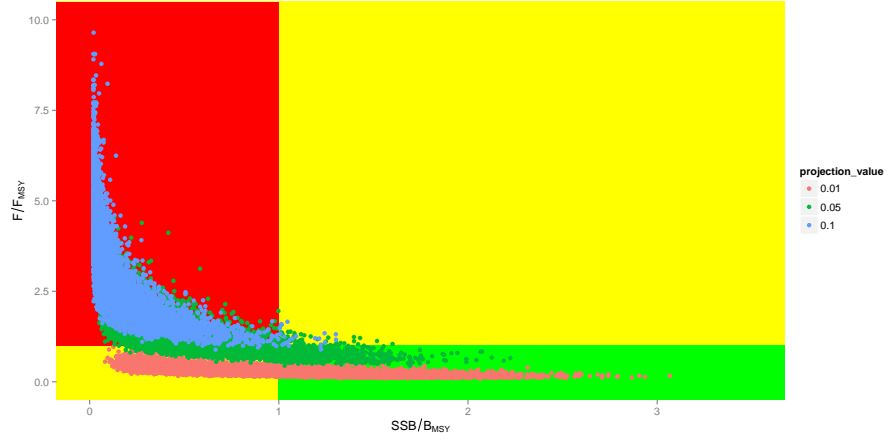**Figure 6:** Kobe trace of stock trajectory

6

Figure 7: Distribution of posterior biomass samples in the final year of the projection
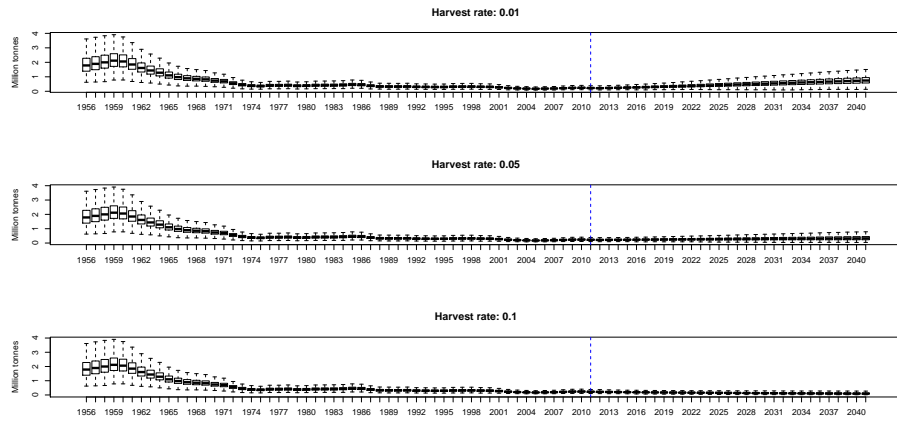

Figure 8: Simple projection box-plots, illustrating predicted dynamics under different harvest rate scenarios.

References

[1] M. K. McAllister, E. A. Babcock, E. K. Pikitch, and M. H. Prager. Application of a non-equilibrium generalized production model to south and north atlantic swordfish: combining bayesian and demographic methods for parameter estimation. *Collected Volume of Scientific Papers ICCAT*, 51(5):1253–1550, 2000.

[2] J. Harwood and K. Stokes. Coping with uncertainty in ecological advice: lessons from fisheries. *Trends in Ecology & Evolution*, 18(12):617–622, 2003.

[3] M. K. McAllister, E. K. Pikitch, and E. A. Babcock. Using demographic methods to construct bayesian priors for the intrinsic rate of increase in the schaefer model and implications for stock rebuilding. *Canadian Journal of Fisheries and Aquatic Sciences*, 58(9):1871–1890, 2001.