SCRS/2017/055 Collect. Vol. Sci. Pap. ICCAT, 74(4): 1628-1638 (2017)

BAYESIAN SURPLUS PRODUCTION MODELS FOR MAKO SHARK, USING ALTERNATIVE INTEGRATION ALGORITHMS

Elizabeth A. Babcock[1], Enric Cortes[2]

SUMMARY

The Bayesian Surplus Production (BSP) software, which uses the Sampling-Importance-Resampling (SIR) method to integrate posterior distributions, was used for the ICCAT mako assessments through 2012. The 2014 assessment of blue shark used both the BSP software and the Markov Chain Monte Carlo (MCMC) algorithm, implemented in the JAGS software, and found that the JAGS and BSP model results were not always consistent. We applied both the BSP1 software (without process error) and the BSP2 software (with process error), and two independent MCMC software packages, JAGS and Stan, to the data from the 2012 mako shark assessment for the North Atlantic to determine whether the same problem exists. We also used the SIR and MCMC algorithms from LearnBayes to fit the same function with both algorithms. Although all the modeling approaches give fairly consistent posteriors for r, the posteriors of K were somewhat different. This may be because there is a long period of catches with no CPUE data, or because the catch and CPUE data are not consistent with each other. The lack of information in the data may cause the model to be sensitive to minor differences in how the model is configured.

RÉSUMÉ

Le logiciel de production excédentaire bayésien (BSP), qui utilise la méthode d'échantillonnage-importance-rééchantillonnage (SIR) pour intégrer les distributions a posteriori, a été employé pour réaliser l'évaluation de requin-taupe bleu jusqu'en 2012. L'évaluation de 2014 du requin peau bleue utilisait à la fois le logiciel BSP et l'algorithme Markov Chain Monte Carlo (MCMC), mis en œuvre dans le logiciel JAGS, et révélait que les résultats des modèles JAGS et BSP n'étaient pas toujours cohérents. Le logiciel BSP1 (sans erreur de processus) et le logiciel BSP2 (avec erreur de processus) et deux progiciels indépendants MCMC, JAGS et Stan, ont été appliqués aux données de l'évaluation du requin-taupe bleu de 2012 de l'Atlantique Nord pour déterminer si les mêmes problèmes se posent. Les algorithmes de SIR et de MCMC de LearnBayes ont été utilisés afin d'ajuster la même fonction avec les deux algorithmes. Même si toutes les approches de modélisation donnent des distributions a posteriori relativement cohérentes pour r, les distributions a posteriori de K étaient légèrement différentes, ce qui pourrait s'expliquer par le fait qu'une longue période de captures soit dépourvue de données de CPUE ou bien que les données de capture et de CPUE ne coïncident pas. L'absence d'informations dans les données pourrait causer la sensibilité du modèle aux légères différences de configuration de modèles.

RESUMEN

El programa de producción excedente bayesiano (BSP), que utiliza el método Sampling-Importance-Resampling (SIR) para integrar distribuciones posteriores, se ha utilizado en las evaluaciones de marrajo dientuso de ICCAT hasta 2012. En la evaluación de 2014 de tintorera se utilizaron tanto el programa BSP como el algoritmo Markov Chain-Monte Carlo (MCMC), implementado en el programa JAGS, y se halló que los resultados de los modelos JAGS y BSP no siempre eran coherentes. En este documento, se aplicaron el programa BSP1 (sin error de proceso) y BSP2 (con error de proceso), así como dos paquetes informáticos MCMC

[1] Department of Marine Biology and Ecology, Rosenstiel School of Marine and Atmospheric Science, University of Miami, 4600 Rickenbacker Cswy., Miami, FL 33149. USA. ebabcock@rsmas.miami.edu
[2] National Marine Fisheries Service, Southeast Fisheries Science Center, Panama City Laboratory, 3500 Delwood Beach Road, Panama City, FL 32408 USA

independientes, JAGS y Stan, a los datos de la evaluación de marrajo dientuso de 2012 para el Atlántico norte con el fin de determinar si se producía el mismo problema. Hemos utilizado también algoritmos SIR y MCMC de LearnBayes para ajustar la misma función con ambos algoritmos. Aunque todos los enfoques de modelación proporcionaron distribuciones posteriores bastante coherentes para r, las distribuciones posteriores de K diferían en cierto modo. Esto podría deberse a que hay un largo periodo de capturas sin datos de CPUE o a que los datos de captura y CPUE no son coherentes unos con otros. La falta de información en los datos puede hacer que el modelo sea sensible a pequeñas diferencias en el modo en que se configura el modelo.

KEYWORDS

Catch/effort, mathematical models,
stochastic models, stock assessment, population dynamics

1. Introduction

The Bayesian Surplus Production (BSP) software, which uses the Sampling-Importance-Resampling (SIR) method to integrate posterior distributions, was used for the mako shark assessments through 2012 (Anonymous 2013, Babcock and Cortes 2009). The 2015 assessment of blue shark used both the BSP software and the Markov Chain Monte Carlo (MCMC) algorithm, implemented in the JAGS software (Plummer 2015), and found that the JAGS and BSP model results were not always consistent (Anonymous 2016). The difference in outputs from different software packages may have been caused by slight differences in how the models were set up, such as how they handled process error. These differences would not matter if the data were highly informative. However, in the case of blue shark, the data were not informative. The catch data were poorly estimated and were not consistent with the CPUE data, in that catches increased at the same time as CPUE increased. This made it difficult for the models to fit the data, and magnified the effect of small differences between software packages.

The data for shortfin mako shark have many of the same problems as the blue shark data, in that the CPUE series are short, and not very consistent with the catch data. To determine whether Bayesian production models applied to the mako shark data would give different results from different software packages, we re-ran one of the models from the 2012 north Atlantic assessment (Model N1) in BSP and JAGS. We also used a newer MCMC integration package called Stan, which uses a different variation of the MCMC (Stan Development Team 2016a,b). Finally, we used the SIR and MCMC algorithms from the LearnBayes R package to fit an identical model with both SIR and MCMC integration methods (Albert 2009, Albert 2014). A preliminary version of this analysis conducted for the Mako shark data meeting in 2017 found large differences in results between algorithms; however, the current analysis finds that the algorithms are more consistent, except in the estimated variance for the unfished biomass (K).

2. Methods

The Bayesian surplus production model (BSP) software available in the ICCAT catalog of methods (McAllister and Babcock 2003) was modified for the shark assessments to allow catches to be estimated based on historical longline fishing effort (Babcock and Cortes 2015). The discrete-time version of the model was used, so that:

(1) $$B_{t+1} = rB_t - \frac{r}{K} B_t^2 - C_t$$

where B_t= biomass at the beginning of year *t*, *r* is the intrinsic rate of increase, *K* is carrying capacity and C_t is the catch in year *t*. The log-likelihood was:

(2) $$\ln L = \sum_j \sum_y \frac{-\left[\ln(I_{j,y}) - \ln(\hat{q}_j \hat{B}_y)\right]^2}{2\sigma_{j,y}^{\ 2}} - \ln\left(I_{j,y}\right) - 0.5\ln\left(2\pi\sigma_{j,y}^{\ 2}\right)$$

were $I_{j,y}$ is the CPUE in year y for series j, $\hat{q}_j$ is the constant of proportionality for series j, $\hat{B}_y$ is the estimated biomass in year y, and $\sigma_{j,y}{}^2$ is the observation error variance (=1/weight) applied to series j in year y. The weights were either assumed to be constant across all series and data points and σ^2 was estimated, or σ^2 was set equal to a constant value.

The BSP software can estimate the catchability q as a parameter with a uniform prior, or it can calculate the MLE estimate of q, which is:

$$(3)\quad \hat{q}_j = \exp\left(\frac{1}{n}\sum_y \ln(I_{j,y}) - \ln(\hat{B}_y)\right)$$

BSP1 and BSP2 can also estimate the observation error variance for each series using an MLE method (Walters and Ludwig, 1994):

$$(4)\quad \hat{\sigma}_j^2 = \frac{1}{n}\sum_y \left(\ln(I_{j,y}) - \ln(\hat{q}_j\hat{B}_y)\right)^2$$

Using the MLE estimates speeds convergence, but may influence the posterior distributions of other parameters.

The original BSP model (BSP1, McAllister and Babcock 2003) does not include process error, but the version called BSP2 (McAllister 2014) can include process error, so that the biomass in each time period is drawn from a lognormal distribution with the median equal to the predicted value and a specified process error standard deviation.

Similar models were coded in JAGS (Lunn et al. 2013, Plummer 2015, Su and Yajima 2014) and Stan (Stan Development Team 2016a, 2016b). These two packages both use MCMC algorithms, and both can be run from R, but they use different MCMC algorithms and have slightly different syntax. A similar model was also coded in R, and the SIR and MCMC algorithms from LearnBayes were used to calculate the posteriors (Albert 2009, 2014). In the JAGS and Stan versions, the catchability (q) was estimated as a free parameter, and the error standard deviation was either input or estimated as a free parameter. It is not possible to use the MLE shortcuts for q and the error standard deviation in JAGS, because the code is object-oriented rather than sequential. However, in the R version fitted with LearmBayes, it was possible to use the MLE estimates of q and sigma, so that the code is nearly identical to the code in BSP1. The JAGS and Stan version included process error, but no process error was included in the LearnBayes version.

All models were applied to the catch, CPUE and effort data available from the 2012 mako shark assessment for the North Atlantic (Anonymous 2015). Priors were set up as in the 2012 assessment. The starting biomass ratio (Bo/K) was lognormal with a mean of 1.0 and CV of 0.2, bounded between 0.2 and 1.1. The base case prior for K was uniform on log(K), bounded between 0.01 and 5 million. The prior for r was lognormal with a mean of 0.058 and a log standard deviation of 0.12, bounded between 0.01 and 2. The prior for q for each index was uniform if it was estimated. If the observation error standard deviation was estimated, it was given an uninformative uniform prior. Post-model pre-data (PDPM) runs were used to see if the model setup and priors favored any particular solution (McAllister 2014). PMPD runs were set up in JAGS by making all the CPUE data equal to NA, and in Stan, BSP1 and BSP2 by having only one CPUE data point. In both cases, the data contain no information so model output will be entirely determined by the priors.

There were three model runs in addition to the post-model pre-data run that varied in the set up of observation and process error (**Table 1**). For some model runs, the observation error standard deviation was set to 1, as was done in some runs in previous assessments, and in some runs the value was set to 0.2, which is close to the MLE estimate. Production models in JAGS do not fit well without at least some process error. Thus, we used a low process error (CV=0.001) to get as close as possible to the case with no process error. In some models, process error CV was set to 0.05.

In BSP1 and BSP2, either the multivariate t-distribution or the priors of the parameters was used for the importance function, and the importance sampling was assumed to have converged when the highest weight was less than 0.5% of the total, and the CV of the weights was similar to the CV of the likelihood times priors. For JAGS and Stan, the MCMC was considered converged if the Gelman-Rubin diagnostic was near 1, and the effective number of parameters was more than 300. For LeanBayes, the multivariate t distribution was used as an importance function.

3. Results

The catch and CPUE data for mako sharks are not consistent with each other, in that both CPUE and catch decline from the mid-1980s through 2000 and then increase (**Figure 1**). Nevertheless, all of the models in BSP1, BSP2, Stan, and LearnBayes were able to converge on a solution. The JAGS models converged eventually, but it was necessary to run more than 10 million MCMC iterations (**Table 1**).

In a preliminary version of this analysis, the JAGS and Stan models estimated very high values of r, so that the model was able to fit the U-shaped trend in the indices. However, those fits had used an incorrect, larger value of the prior standard deviation for r, allowing the model to estimate a value of r very different from the prior. When that error was corrected, all versions of the model estimated posteriors of r that were quite similar to the prior, as expected when the data are uninformative (**Table 2, Table 3, Figure 2, Figure 3, Figure 4, Figure 5**). However, the BSP models generally estimated a higher posterior CV of K than did JAGS and Stan. The BSP runs had CVs of 0.9 to 1.0 for K, and the posterior distribution had non-zero probabilities associated with K values up to the upper boundary of 5 million. The BSP posteriors for r and K looked very similar to the post model pre-data run, even when they were fitted to data. In contrast, the Jags and STAN runs estimated a value of K with a fairly similar mode but with a smaller CV, and a more bell-shaped posterior distribution. The post model pre data runs were quite similar in all three cases.

The difference in the posteriors of K between the software packages is not caused by the differences in the algorithms used to fit the model. When SIR and MCMC were used to estimate the posteriors of the exact same model code, the results were identical as expected (**Table 3, Figure 5**). The LearnBayes fits also estimated a posterior of K with a relatively low CV.

The wider posterior distribution of K in the BSP runs makes these models more optimistic about current status than the MCMC runs (**Figure 6**). The mean current biomass tends to be higher in the BSP models than in the JAGS and Stan runs.

4. Discussion

The fact that the SIR and MCMC programs give somewhat different results is surprising, since they have very similar model structures. Also, when we ran identical models with both SIR and MCMC in LearnBayes, the results were nearly identical as expected. The slight differences in how each software package estimates the catchability and observation error variance may explain some of the differences in model results. There may be other subtleties in the software packages that we have not identified. None of these minor differences would influence the results substantially if the data were informative enough to give a more precise posterior estimate of K. The lack of agreement between similar models in estimating population parameters and status may be considered a diagnostic that the data are poor, and perhaps not suitable for assessment.

Considering that there is no particular reason to choose one software package over another, we recommend that the working group continue to use multiple software packages and treat them as valid alternative model runs. Also, since the models have slightly different input configurations it is important to closely compare the priors and data input for each software package to ensure that the models really are using the same data and priors. We also suggest running sensitivity analysis with different standard deviations for the prior for r, since the results are highly sensitive to this assumption.

5. Acknowledgements

E. Babcock’s work was supported in part by NOAA through a grant to CIMAS.

References

Albert, J. 2009. Bayesian Computation with R, 2nd Edition. Springer. New York. 298 p.

Albert, J. 2014. LearnBayes: Functions for learning Bayesian inference. R package version 2.15.

Anonymous. 2005. Report of the 2004 inter-sessional meeting of the ICCAT subcommittee on by-catches: shark stock assessment. ICCAT Col.Vol.Sci.Pap. 58(3): 799-890.

Anonymous. 2008. Report of the 2007 data preparatory meeting of the shark species group. ICCAT Col. Vol. Sci. Pap. 62(1):1325-1404.

Anonymous. 2009. Report of the 2008 shark stock assessment meeting. ICCAT Col. Vol. Sci. Pap. 64(5): 1343-1491.

Anonymous. 2013. 2012 shortfin mako stock assessment and ecological risk assessment meeting (Olhão, Portugal - June 11 to 18, 2012). ICCAT Col. Vol. Sci. Pap. 69(4): 1427-1570.

Anonymous. 2016. Report of the 2015 ICCAT blue shark stock assessment session (Oceanário de Lisboa, Lisboa, Portugal - July 27 to 31, 2015) ICCAT Col. Vol. Sci. Pap.

Babcock, E. A. and E. Cortes. 2015. Bayesian surplus production model applied to blue shark catch, CPUE and effort data. ICCAT Col. Vol. Sci. Pap. 72(6): 1440-1448.

Babcock, E. A. and E. Cortes. 2009. Surplus production model applied to the data for blue and mako sharks available at the 2001 ICCAT Bycatch working group and other published data. ICCAT Col. Vol. Sci. Pap. 64(5): 1568-1577.

Lunn, D., Jackson, C., Best, N., Thomas, A., and Spiegelhalter, D. 2013. The BUGS Book: A Practical Introduction to Bayesian Analysis. CRC Press.

MCallister, M.K. and E. A. Babcock. 2003. Bayesian surplus production model with the Sampling Importance Resampling algorithm (BSP): a user's guide. Available from www.iccat.int.

MCallister, M.K. 2014. A generalized Bayesian surplus production model software (BSP2). Collect. Vol. Sci. Pap. ICCAT 70(4): 1725-1757.

Meyer R. and R. B Millar. 1999. BUGS in Bayesian stock assessments. Can. J. Fish. Aquat. Sci. 56: 1078–1086

Plummer, M. 2015. Package rjags: Bayesian graphical models using MCMC. R package version 4-4.

Stan Development Team. 2016a. RStan: the R interface to Stan. R package version 2.14.1.

Stan Development Team. 2016b. Stan Modeling Language Users Guide and Reference Manual, Version 2.14.0.

SU, Y.-S., and Yajima, M. 2014. R2jags: A Package for Running jags from R.

Walters, C.J. and Ludwig, D. 1994. Calculation of Bayes posterior probability distributions for key population parameters: a simplified approach. Can. J. Fish. Aquat. Sci. 51: 713-722.

Table 1. Model specifications for the four model runs applied in BSP, JAGS and Stan.

Model	Estimation error standard deviation	Process error
1	Estimated with uniform prior in BSP1, JAGS and STAN, sigma=0.2 in BSP2	None in BSP1, error =0.001 in BSP2, JAGS, STAN
2	sigma=1	None in BSP1, error =0.001 in BSP2, JAGS, STAN
3	Estimated with uniform prior in BSP1, JAGS and STAN, sigma=0.2 in BSP2	Not run in BSP1, error=0.05 in BSP2, JAGS, STAN
PMPD	Estimated with uniform prior in BSP1, JAGS and STAN, sigma=0.2 in BSP2	None in BSP1, error =0.001 in BSP2, JAGS, STAN

Table 2. Posterior means and CVs for each model run. Note that BSP results for K and Bo are in thousands, while JAGS and STAN results are not.

Software	Variable	Model 1		Model 2		Model 3		PMPD	
BSP1	K (1000)	1373.63	0.9	1412.1	0.9			1319.27	1.0
	r	0.06	0.1	0.06	0.1			0.06	0.1
	Bo (1000)	1225.18	0.9	1271.04	0.9			1189.5	1.0
	Bcur/Bmsy	1.73	0.1	1.72	0.2			1.62	0.3
	Fcur/Fmsy	0.29	0.9	0.33	1.2			0.96	4.4
BSP2	K (1000)	914.29	0.9	1401.28	0.8	1207.22	0.9	1310.06	0.8
	r	0.06	0.1	0.06	0.1	0.06	0.1	0.06	0.1
	Bo(1000)	793.86	0.9	1258.55	0.8	1107.48	0.9	1177.67	0.8
	Bcur/Bmsy	1.62	0.2	1.71	0.2	1.68	0.2	1.61	0.3
	Fcur/Fmsy	0.4	0.6	0.35	1.0	0.48	0.8	0.78	1.4
JAGS	Rhat	1.02		1.31		1.00	NA	1.73	
	n.eff	150		13		2800	NA	5	
	K(1000)	249.41	0.22	159.46	0.24	185.69	0.28	1448.36	0.88
	r	0.06	0.12	0.06	0.12	0.06	0.12	0.06	0.12
	Bo/K	0.81	0.20	0.99	0.17	0.98	0.20	1.12	0.13
	Bcur/Bmsy	1.27	0.14	0.80	0.32	1.21	0.21	1.72	0.25
	Fcur/Fmsy	0.88	0.28	2.56	0.46	1.31	0.30	31.18	6.78
STAN	Rhat	1.00		1.00		1.00		1.01	
	n.eff	3441		1747		42533		526	
	K (1000)	177.09	0.10	149.14	0.14	165.98	0.17	1219.41	1.02
	r	0.07	0.12	0.06	0.12	0.06	0.12	0.06	0.12
	Bo/K	0.78	0.20	1.00	0.20	0.96	0.20	1.02	0.20
	Bcur/Bmsy	0.98	0.10	0.74	0.27	1.11	0.18	1.62	0.25
	Fcur/HRmsy	1.39	0.16	2.80	0.43	1.50	0.21	0.66	1.67

Table 3. Posterior means and CVs from LearnBayes SIR and MCMC algorithm applied to the same function.

Variable	SIR mean	SIR CV	MCMC mean	MCMC.CV
K (1000)	459.97	0.99	460.13	0.99
r	0.06	0.13	0.06	0.13
B0/K	0.79	0.17	0.79	0.17

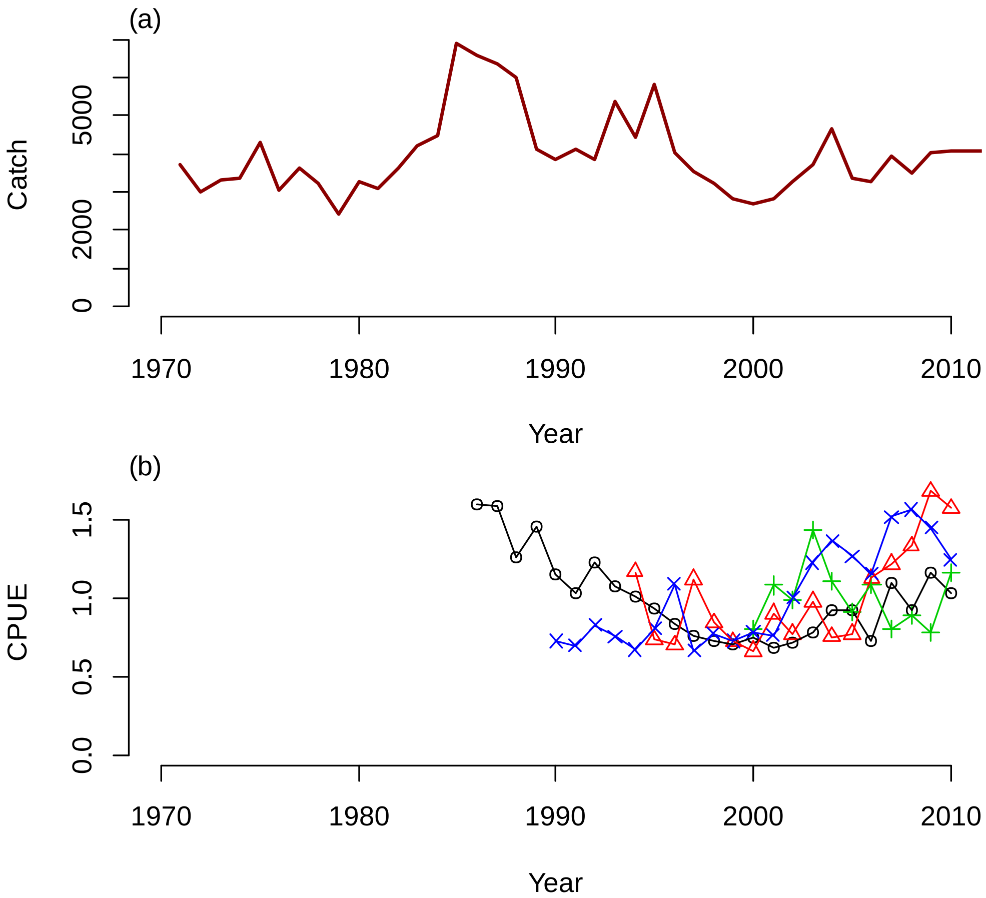

Figure 1. Mako shark catch (a) and CPUE (b) from the 2012 assessment.

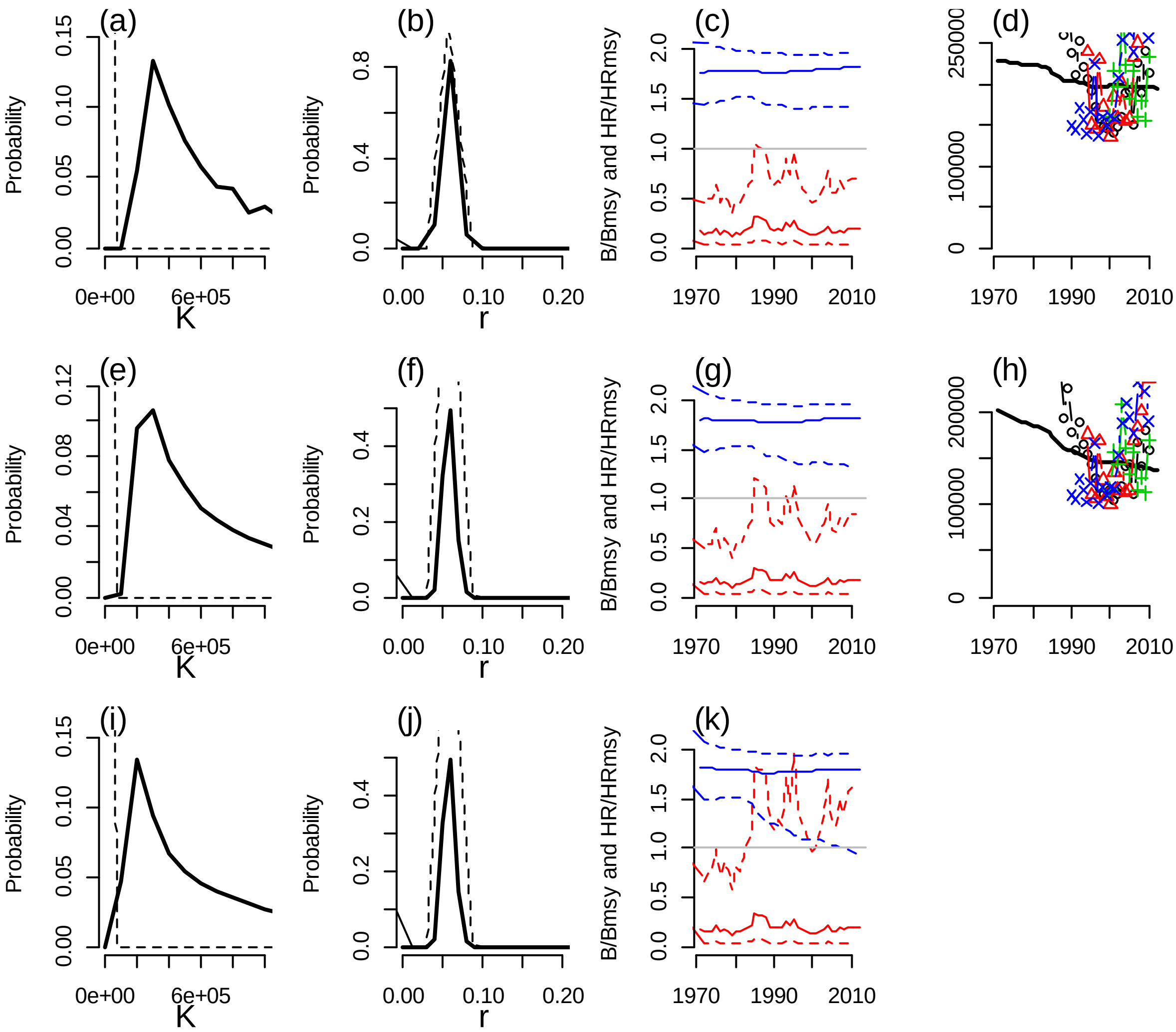

Figure 2. BSP1 models with an estimated error standard deviation (a-d), error standard deviation of one (e-h), and a post-model pre-data run (i-k). The left column shows the prior (dashed line) and posterior (solid line) for K, second column shows the prior and posterior for r, third column shows the trend in biomass ratio (blue) and harvest rate ratio (red) with 80% credible intervals, and last column shows the biomass trend fitted to the CPUE series.

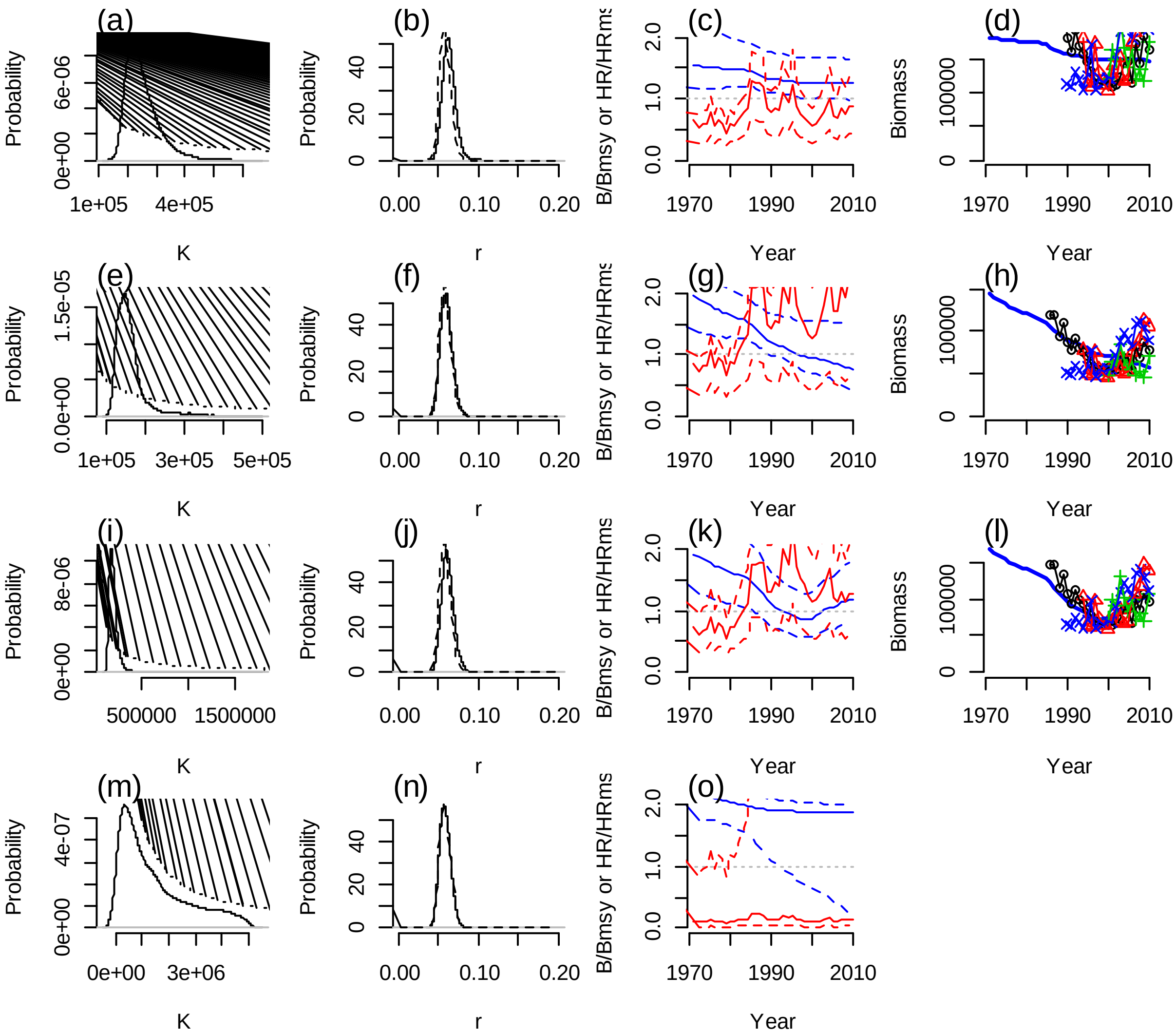

Figure 3. JAGS models with an estimated error standard deviation and process error of 0.001 (a-d), error standard deviation of one and process error of 0.001 (e-h), error standard deviation estimated and process error equal to 0.05 (i-l), and a post-model pre-data run (m-o). The left column shows the prior (dashed line) and posterior (solid line) for K, second column shows the prior and posterior for r, third column shows the trend in biomass ratio (blue) and harvest rate ratio (red) with 95% credible intervals, and last column shows the biomass trend fitted to the CPUE series.

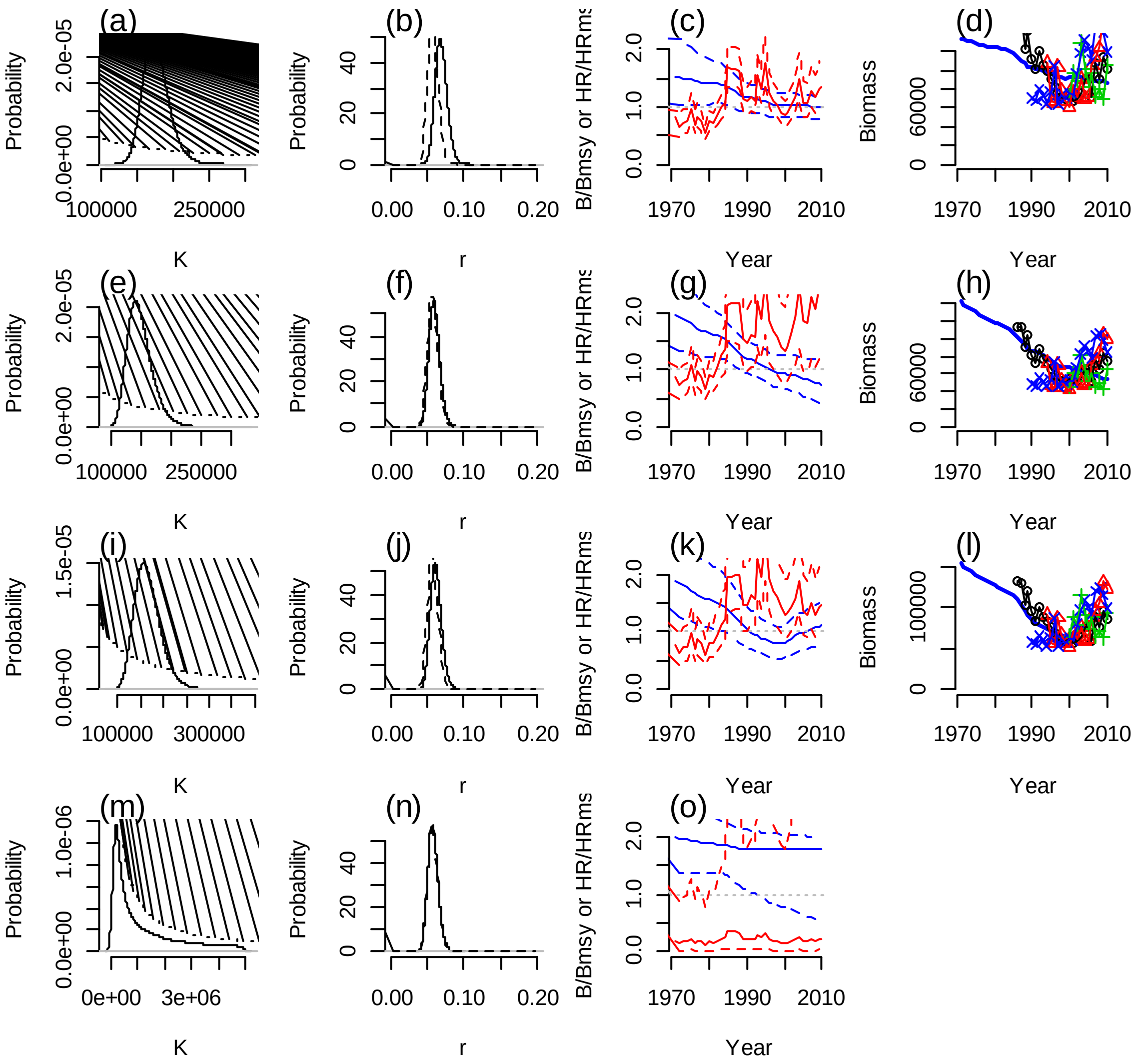

Figure 4. STAN models with an estimated error standard deviation and process error of 0.001 (a-d), error standard deviation of one and process error of 0.001 (e-h), error standard deviation estimated and process error equal to 0.05 (i-l), and a post-model pre-data run (m-p). The left column shows the prior (dashed line) and posterior (solid line) for K, second column shows the prior and posterior for r, third column shows the trend in biomass ratio (blue) and harvest rate ratio (red) with 95% credible intervals, and last column shows the biomass trend fitted to the CPUE series.

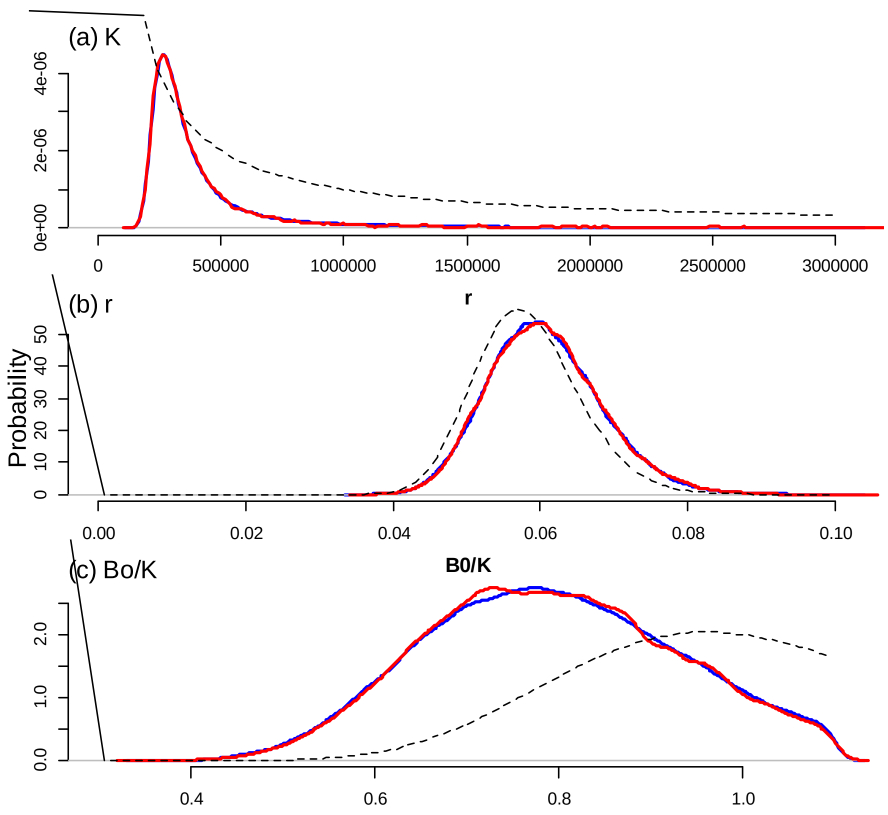

Figure 5. Posteriors calculated from SIR and MCMC using the LearnBayes functions. Red is SIR, blue is MCMC, and the dashed line is the prior. If only red is visible, the SIR and MCMC results are identical.

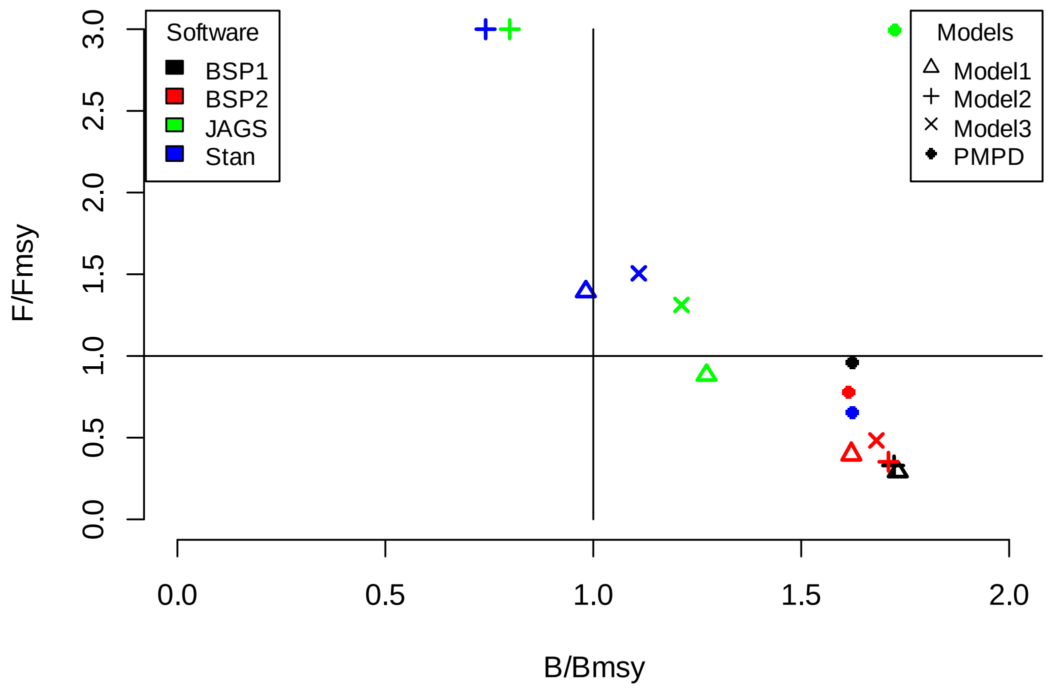

Figure 6. Mean current status. Values of F/Fmsy greater than 3 were plotted at 3. Colors indicate software and plotting characters indicate model set up.