 Collect. Vol. Sci. Pap. ICCAT, 80(4): 656-664 (2023)

BAYESIAN SURPLUS PRODUCTION MODELS FOR BLUE SHARKS USING THE LEGACY BSP SOFTWARE

Elizabeth A. Babcock[1]

SUMMARY

The 2015 blue sharks stock assessment included Bayesian surplus production models conducted with an old software called BSP that used the Sampling-Importance-Resampling algorithm rather than MCMC for numerical integration, along with some JAGS code that was similar to the JABBA R package that is currently used. The legacy BSP software and the old JAGS code were used with the new catch and CPUE data, but the same settings as were used for the 2015 assessment to verify that the choice of software does not influence the assessment results. The BSP software has some features that are not available in JABBA and have been used for blue sharks, such as the ability to estimate catches in the early years of the fishery from effort, and then use catches for the rest of the years. Conversely, BSP does not have JABBA's ability to model catches as uncertain, and JABBA provides many useful diagnostics. Because the differences in software are minor and JABBA is more convenient and reproducible, JABBA should be the preferred Bayesian state-space production models for future assessments.

RÉSUMÉ

L'évaluation du stock de requin peau bleue de 2015 comprenait des modèles bayésiens de production excédentaire réalisés avec un ancien logiciel appelé BSP qui utilisait l'algorithme Sampling-Importance-Resampling plutôt que MCMC pour l'intégration numérique, ainsi qu'un certain code JAGS qui était similaire au paquet JABBA R actuellement utilisé. Le logiciel BSP classique et l'ancien code JAGS ont été utilisés avec les nouvelles données de capture et de CPUE, mais les mêmes configurations ont été utilisées que pour l'évaluation de 2015 afin de vérifier que le choix du logiciel n'influence pas les résultats de l'évaluation. Le logiciel BSP présente certaines caractéristiques qui ne sont pas disponibles dans JABBA et qui ont été utilisées pour le requin peau bleue, comme la possibilité d'estimer les captures au cours des premières années de la pêcherie à partir de l'effort, puis d'utiliser les captures pour le reste des années. Inversement, BSP n'a pas la capacité de JABBA de modéliser les captures comme incertaines, et JABBA fournit de nombreux diagnostics utiles. Étant donné que les différences entre les logiciels sont mineures et que JABBA est plus pratique et reproductible, JABBA devrait être le modèle de production bayésien état-espace préféré pour les évaluations futures.

RESUMEN

La evaluación de stock de tiburón azul de 2015 incluía modelos de producción excedente bayesianos realizados con un software antiguo llamado BSP que utilizaba el algoritmo de Sampling-Importance-Resampling en lugar del MCMC para la integración numérica, junto con algún código JAGS que era similar al paquete JABBA R que se utiliza actualmente. El software BSP heredado y el antiguo código JAGS se utilizaron con los nuevos datos de capturas y CPUE, pero con los mismos ajustes que se utilizaron en la evaluación de 2015 para verificar que la elección del software no influye en los resultados de la evaluación. El software BSP tiene algunas características que no están disponibles en JABBA y que se han utilizado para el tiburón azul, como la posibilidad de estimar las capturas en los primeros años de la pesquería a partir del esfuerzo, y luego utilizar las capturas para el resto de los años. Por el contrario, BSP no tiene la capacidad de JABBA de modelar las capturas como inciertas, y JABBA proporciona muchos diagnósticos útiles. Dado que las diferencias en el software son menores y JABBA es más cómodo y reproducible, JABBA debería ser el modelo de producción bayesiano estado-espacio preferido para futuras evaluaciones. -

[1]Department of Marine Biology and Ecology, Rosenstiel School of Marine, Atmospheric, and Earth Science, University of Miami, 4600 Rickenbacker Cswy., Miami, Florida, 33149, USA. ebabcock@miami.edu

KEYWORDS

Bayesian, stock assessment, blue shark

1. Introduction

In the last ICCAT stock assessment for blue sharks (Anon., 2016), Bayesian state-space production models were applied to both the North and South Atlantic stocks, using a legacy Visual Basic software package (BSP and BSP2) that uses the Sampling-Importance-Resampling (SIR) algorithm rather than the Markov Chain Monte Carlo (MCMC) to integrate the posterior distribution (McAllister *et al*. 2001, Babcock and Cortes 2005, 2009, 2015). In addition, similar models were implemented in the MCMC package JAGS (Plummer 2023, Su and Yajima 2023, R Core Team 2023), with slightly different code for the North vs. South Atlantic stocks. At that time, the JABBA R library (Winker *et al*. 2018) was not yet available, although the JAGS code used for the South Atlantic runs was a precursor of JABBA. The JABBA library facilitates Bayesian production modeling by providing useful diagnostics and plots, suggesting default settings, and incorporating modern methods, for example, to weight the CPUE indices. However, some of its settings and methods may be different from the model runs that were used for advice in the previous assessment, in ways that might influence the estimates of status.

The goal of this analysis is to re-run some of the Bayesian production models that were run in 2015 with the same software and settings, but using the current data to evaluate whether changes in software might be part of the reason for any differences between the 2015 and 2023 assessments. These model runs are not intended to be included in the assessment, but rather to inform discussion of how the surplus production models have evolved since the last assessment.

2. Methods

The software used for this analysis include the original BSP software in Visual BASIC (BSP1), a version of the same code which also includes process error (BSP2), JAGS code written by the BSP authors for the 2015 assessment that replicates the models structure of BSP2, and JABBA, run with settings designed to mimic the BSP settings as closely as possible (**Tables 1** and **2**). Except for using the updated CPUE and catch data (**Figures 1** and **2**), the models were set up using the same inputs as some of the 2015 runs, including the priors for the population growth rate parameter r, unfished biomass K, and starting biomass ratio. The first year of the fishery was assumed to be 1957 in the north and 1971 in the south. Although BSP1 is able to estimate catches before the first year of complete catch data from effort, BSP2 is not. In the 2015 assessment, catches in the north before 1971 were estimated from a regression of catch and effort. The same estimated catches were used for those years in this exercise for BSP2 and JAGS (**Figure 1**). The CPUE series available from this assessment are quite different from those in 2015 (**Figure 2**).

The primary difference between the many runs done with BSP, BSP2 and JAGS in 2015 was the method used to weight the abundance indices. Methods included iterative re-weighting, which is a method that estimates an error variance for each series, and then inputs the modal value of each variance as a fixed constant, effort weighting, in which the weight given to each series was proportional to the relative effort in the fishery (thus giving more weight to commercial than recreational fleets), and a method in which all points in all series were given the same error variance, estimated as a free parameter (equal weighting). For this exercise, equal weighting was used for BSP1, iterative re-weighting for BSP2 and JAGS, and an estimated variance per series in JABBA. The runs in 2015 with BSP1 with different weighting were quite similar to each other, so the equal weighting run was chosen as representative. The BSP2, JAGS and JABBA runs are using a single estimate per series to make them as similar as possible.

3. Results

For the North Atlantic stock, the four model configurations gave surprisingly different results (**Table 3**, **Figure 3**, **Figure 4**). The model with no process error (BSP1), similar to many of the runs in the 2015 assessment, was unable to find a trend in the CPUE series, and simply estimated a flat trajectory with a very high (and uncertain) K. The results with BSP2 and the JAGS implementation were similar in the estimated trajectory for both biomass and harvest rate, as expected. The JABBA run was more pessimistic. This may be because it is not possible in JABBA to use any prior other than the lognormal. Thus, the JABBA prior for K, even with a large CV was more informative than the other priors. However, this is not the only explanation for the difference, as changing the K prior did not greatly change the estimated trajectory (not shown).

For the South Atlantic, BSP1 fit a flat trend with a large K, the same as for the north, due to the lack of process error. The other three models were more consistent and found similar trends (**Table 3, Figure 5**).

4. Discussion

The lack of data from the early years of the time series is one reason that all the model fits are highly uncertain. It may be worth revisiting some of the older series, such as the early Japanese longline series, and the US-Obs-Cru series. Also, considering the uncertainty in the catch inputs, some catch sensitivities may be needed. Finally, given the highly uncertain CPUE series, it might be worth trying more than one weighting. In 2015, runs that were weighted by the relative effort in the corresponding fisheries were considered among the best, and this has not yet been tried with the 2023 data.

Although there are differences between the estimated trends for the different software packages, they can mainly be explained by the differences in priors and model structures, not anything inherent to the software applications. The legacy BSP1 and BSP2 software is written in Visual BASIC, which is no longer maintained, and it is very difficult to learn and to run on a modern computer. Also, the SIR application is not very efficient for these kinds of models compared to modern MCMC methods, and the convergence diagnostics are not as well developed as for MCMC. Thus, the legacy software should not be used for assessment. If the assessment team wishes to explore the implications modeling approaches that are not available in JABBA, they can easily be coded in JAGS, STAN or PyMC to take advantage of modern advances in MCMC algorithms and data science (Salvatier *et al.* 2016, Stan Development Team 2023).

Acknowledgements

E. Babcock's work on the BSP software has been supported over the years by NOAA via CIMAS.

References

Anonymous. 2016. Report of the 2015 ICCAT blue shark stock assessment session. Collect. Vol. Sci. Pap. ICCAT, 72(4): 866-1019

Babcock E.A. and Cortes E. 2015. Bayesian surplus production model applied to blue shark catch, CPUE and effort data. SCRS/2015/150

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.

Babcock, E. A. and E. Cortes. 2005. Updated Bayesian surplus production model applied to blue and mako shark catch, CPUE and effort data. ICCAT Col. Vol. Sci. Pap. 58(3):1044-1053.

McAllister, M. K., E. K. Pikitch and E. A. Babcock. 2001. Using demographic methods to construct Bayesian priors for the intrinsic rate of increase in the Schaefer model and implications for stock rebuilding. Can. J. Fish. Aquat. Sci. 58: 1871-1890.

Plummer M (2023). rjags: Bayesian Graphical Models using MCMC. R package version 4-14, CRAN.R-project.org/package=rjags>.

R Core Team. 2023. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Salvatier J., T.V. Wiecki, and C. Fonnesbeck. 2016. Probabilistic programming in Python using PyMC3.

Stan Development Team. 2023. RStan: the R interface to Stan. R package version 2.21.8. mc-stan.org/.

Su Y., and M. Yajima. 2021. R2jags: Using R to Run 'JAGS'. R package version 0.7-1. CRAN.R-project.org/package=R2jags.

Winker, H., Carvalho, F., Kapur, M. 2018. JABBA: Just Another Bayesian Biomass Assessment. Fish. Res. 204: 275-288.

Table 1. Software available for comparison.

Software	Language	Numerical integration	Features used for BSH	Process error
BSP1	Visual BASIC	Sampling Importance Resampling	Estimating catch from effort in early years only	No
BSP2	Visual BASIC	Sampling Importance Resampling	BSP with process error	Yes
JAGS	JAGS from R	MCMC with Gibbs Sampler	Written to replicate BSP2	Yes
JABBA	JAGS from R	MCMC with Gibbs Sampler	Used inputs as similar to BSP2 as possible	Yes

Table 2. Model runs with 2023 data.

Area	Software	Process error	Index weighting
North	BSP1	0	equal, estimated one sigma
	BSP2	0.05	iterative re-weighting
	JAGS	0.05	used BSP2 values
	JABBA	0.05	estimated per index
South	BSP1	0	equal, estimated one sigma
	BSP2	0.05	iterative re-weighting
	JAGS	0.05	used BSP2 values
	JABBA	0.05	estimated per index

Table 3. Comparison of model results. Unfished biomass K and biomass in the first year B_0 are in 1000s. H is harvest rate.

	BSP1			BSP2			JAGS					JABBA		
North	Variable	Mean	SD	Mean	Median	sd	Mean	sd	LCI	Median	UCI	mu	lci	uci
	r	0.39	0.05	0.35	0.35	0.05	0.37	0.04	0.29	0.36	0.46	0.40	0.30	0.51
	K	4736	6322	2256	1412	2813	682	733	333	498	2375	392	296	554
	B_0 or B_0/K	3876	5664	1963	1239	2413	1.81	0.24	1.30	1.84	2.18	0.97	0.68	1.30
	B/Bmsy	1.93	0.05	1.74	1.71	0.32	1.61	0.22	1.18	1.61	2.05	1.29	0.91	1.64
	H/Hmsy	0.06	0.05	0.22	0.17	0.18	0.30	0.12	0.06	0.30	0.54	0.44	0.26	0.72
South	r	0.24	0.09	0.15	0.14	0.03	0.18	0.05	0.10	0.17	0.28	0.23	0.14	0.40
	K	33285	28544	4416	4180	1505	2804	2241	553	1913	8110	524	284	1012
	B_0 or B_0/K	27940	24870	3504	3349	1078	1.80	0.24	1.29	1.83	2.18	0.99	0.69	1.31
	B/Bmsy	1.97	0.05	1.94	1.94	0.13	1.92	0.23	1.46	1.93	2.35	1.56	1.24	1.95
	H/Hmsy	0.03	0.05	0.14	0.12	0.08	0.29	0.22	0.05	0.23	0.78	0.73	0.37	1.10

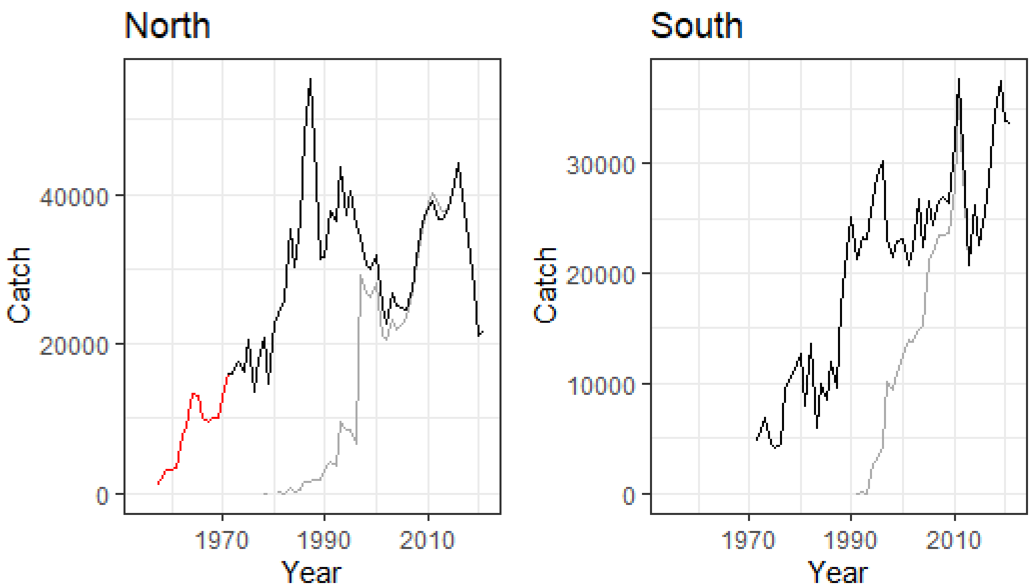

Figure 1. Catches used in the model runs (black and red, where red was estimated from effort for the 2015 assessment), with Task I shown in grey for comparison. Data before 2013 was taken from the 2015 assessment.

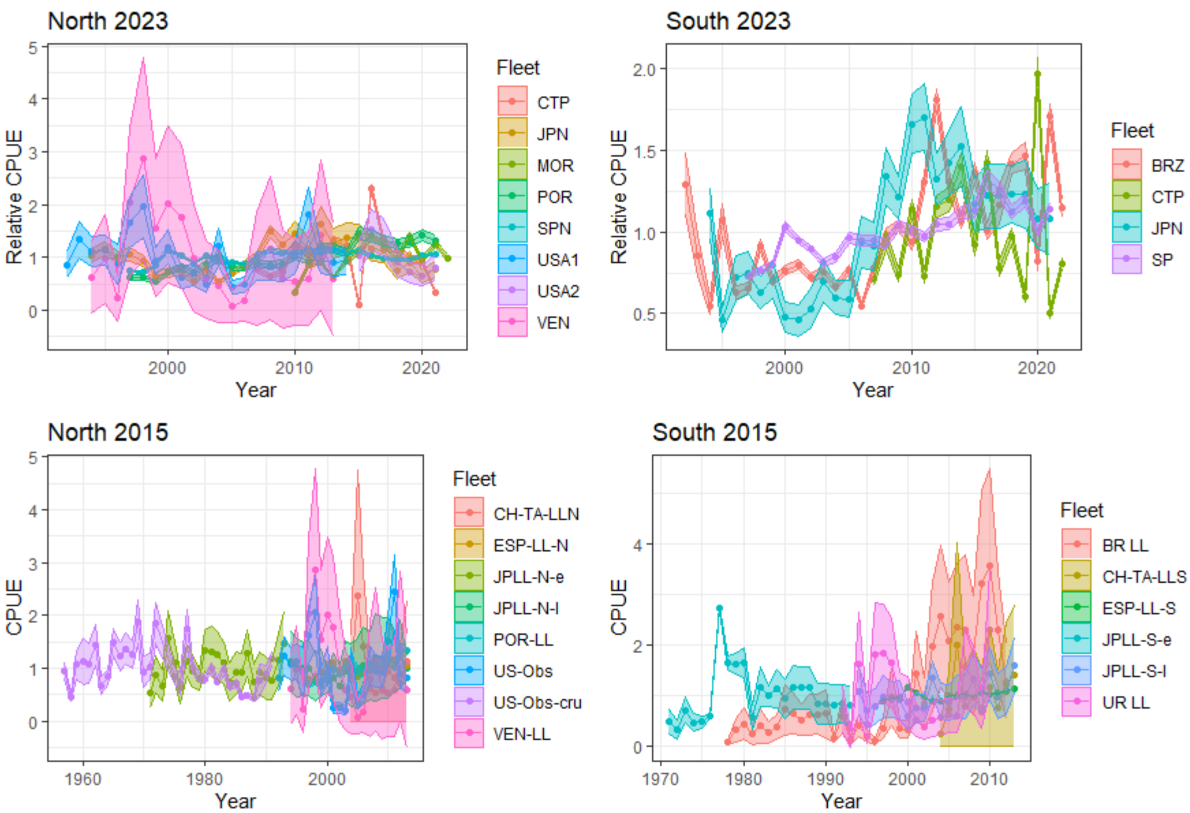

Figure 2. Comparison of indices between the current data (2023) and the past assessment (2015), plus or minus one standard error. The 2023 series have been divided by their means. The 2015 series were divided by the means during periods of overlap between series.

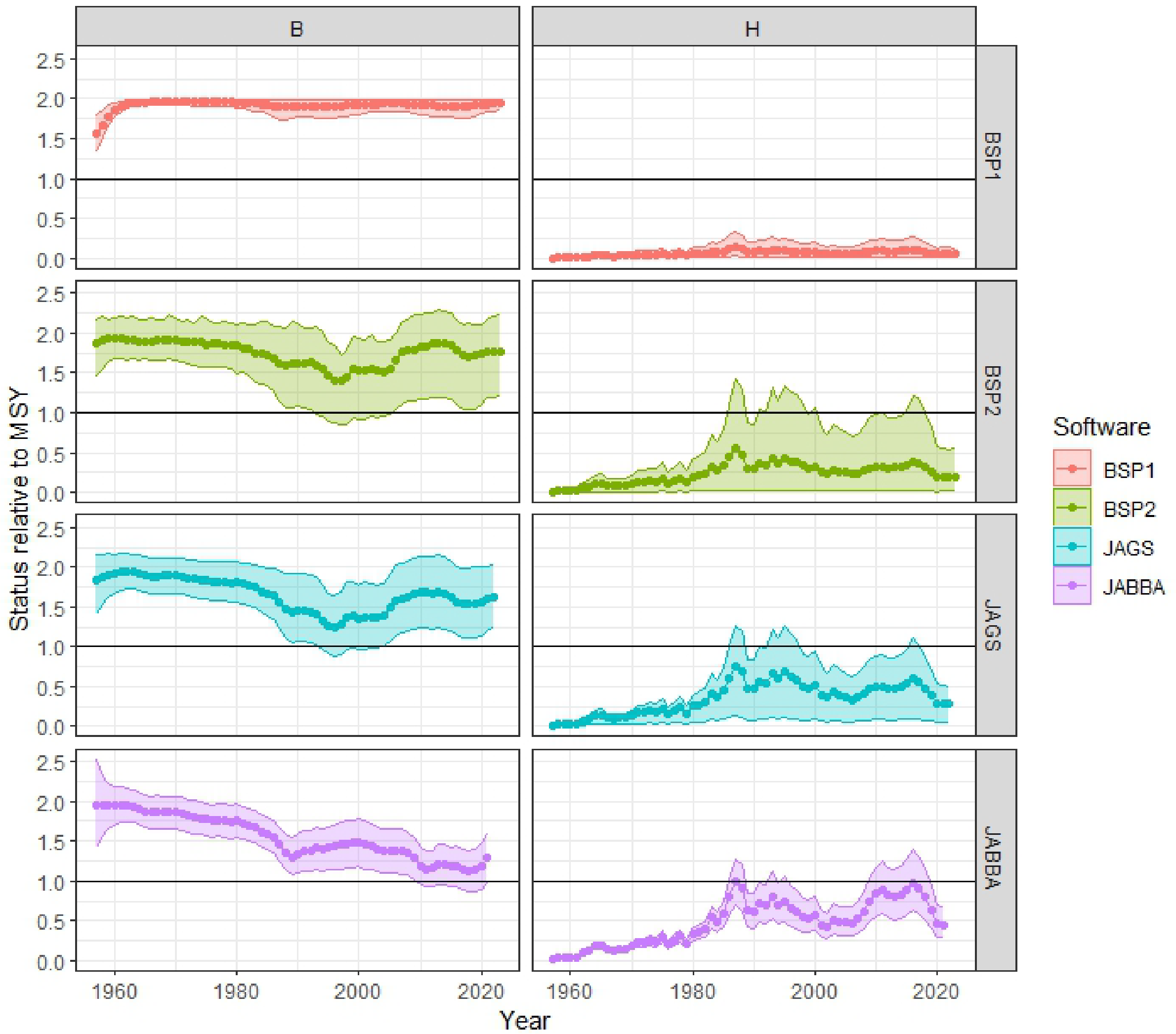

Figure 3. North Atlantic Bayesian production model results with four software packages. See **Table 2** for model descriptions.

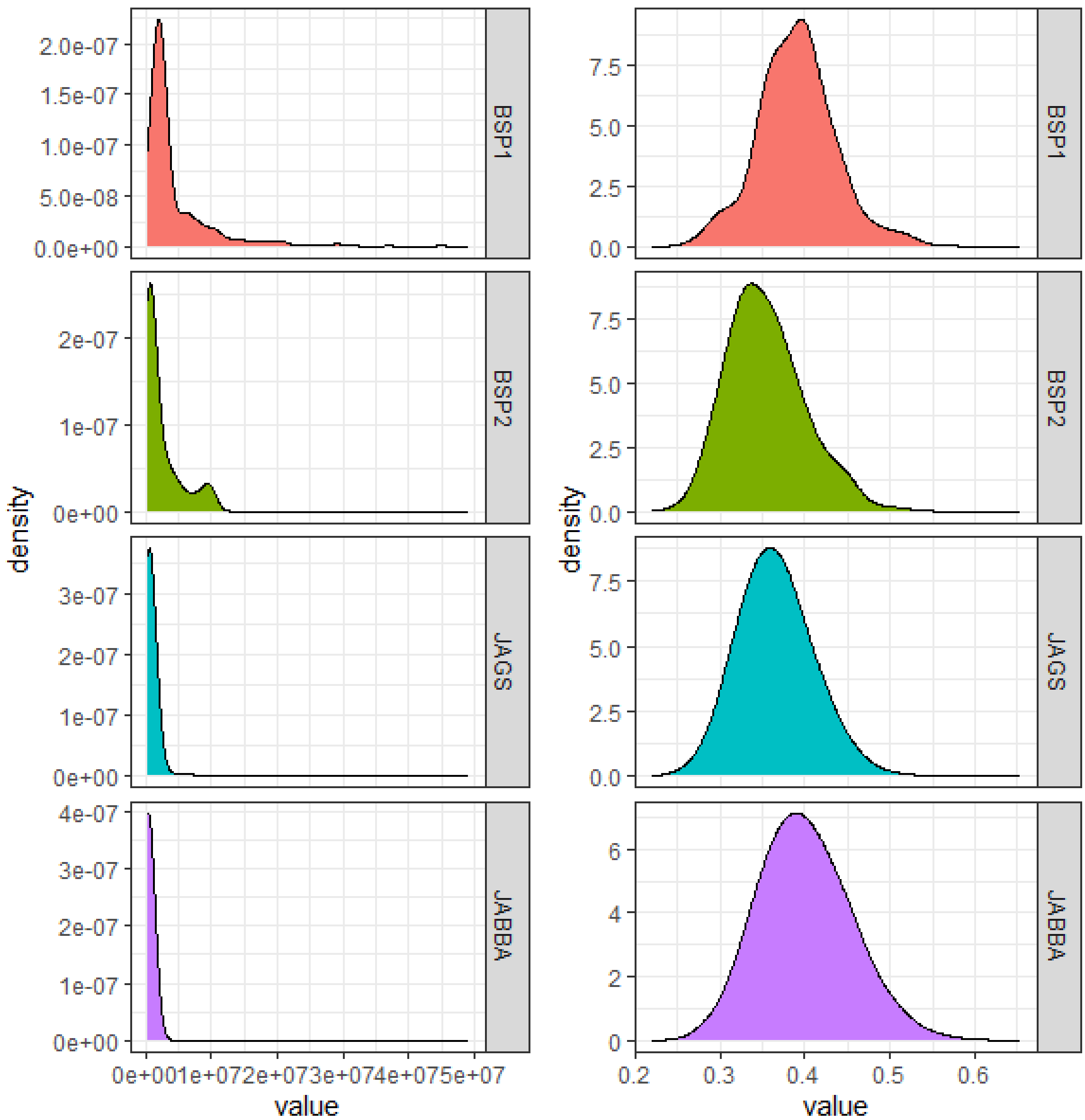

Figure 4. North Atlantic Bayesian production model posterior distributions for r and K for all four model runs.

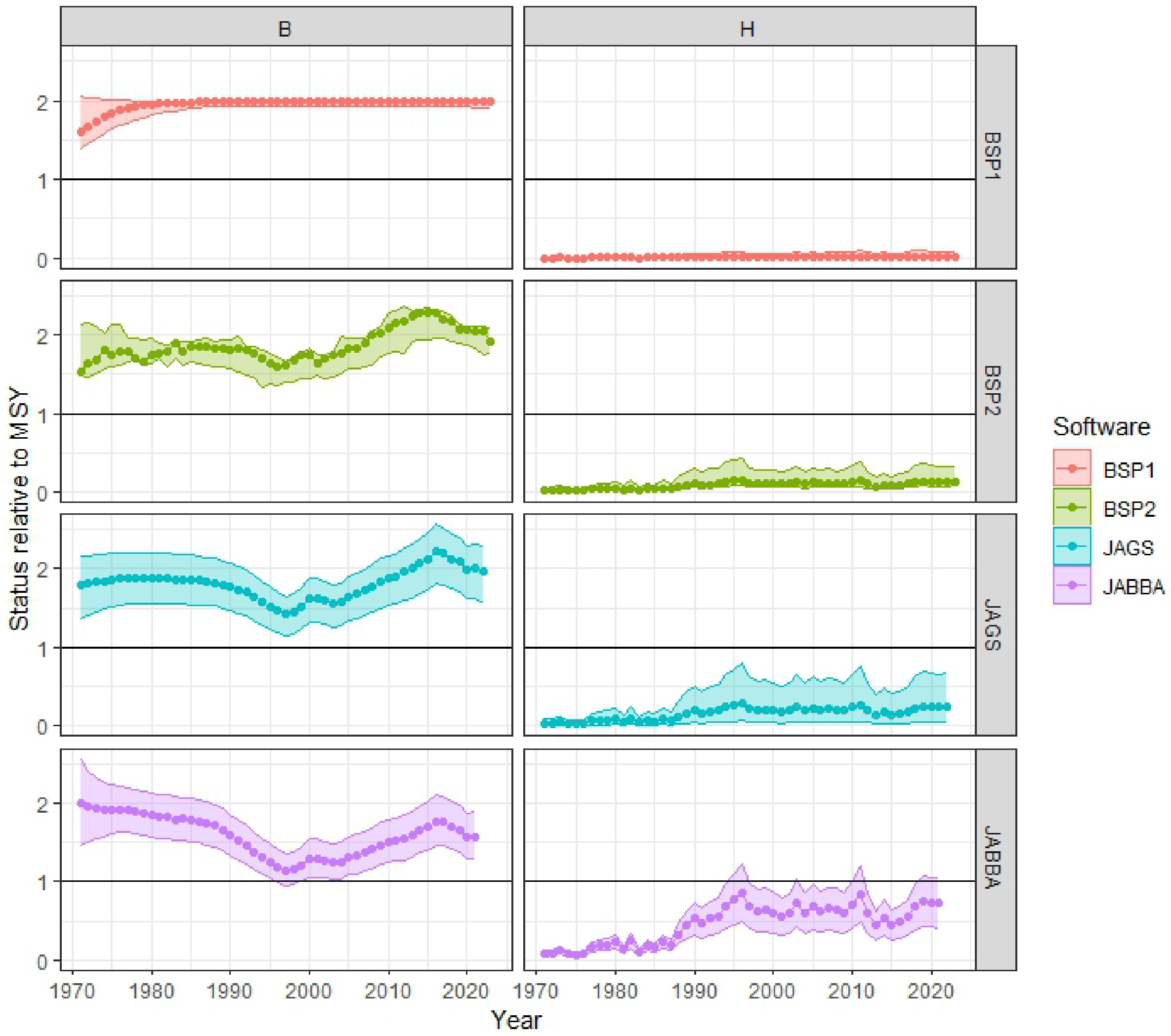

Figure 5. South Atlantic Bayesian production model results with four software packages. See **Table 2 f**or model descriptions.