A Draft Framework for Assigning Probabilities to Alternative Assumptions concerning the Contribution of Sardine Spawning Biomass on the South Coast to Recruitment on the West Coast

Doug S. Butterworth

MARAM (Marine Resource Assessment and Management Group)
Department of Mathematics and Applied Mathematics
University of Cape Town, Rondebosch 7701, South Africa

October 2016

Abstract

A draft framework is put forward, by way of giving examples, to provide a quantitative basis for the combination of information from the hydrodynamics model and fits to stock-recruitment relationships to provide weightings for alternative hypotheses for the proportion of South Coast sardine spawning biomass which contributes to recruitment on the West Coast, for use in revising the pelagic OMP. The intent is that this draft be refined prior to full deliberation, including the International Panel, at the November 28—December 2 workshop. Some suggestions are provided for further work desirably conducted before that event to facilitate discussions there.

Introduction

Probably the question of the extent to which sardine spawning biomass on the South Coast contributes to subsequent recruitment on the West Coast will be a, if not the, dominant consideration in revising the pelagic OMP to address whether and how the (directed) sardine TAC might need to be allocated spatially. The only approach that could in principle provide a clear-cut answer to this question would seem to be an experiment using close-kin genetics to assess the proportions of West Coast recruits originating from parents on each coast, but that would be very expensive and in any case could not (for practical reasons) provide an answer in the very short term. Thus indirect inferences are necessary. It seems that the only two sources of pertinent information available at present are:

a) the hydrodynamics model indicating the proportion of eggs spawned on each coast likely to reach the West Coast recruitment area; and

b) stock-recruitment estimates, in the context of their relative compatibility in terms of fitted stock-recruitment relationships with different proportions (p) of South Coast spawning biomass contributing to recruitment on the West Coast.

To advance the OMP development further, these two sources of information need to be quantified in some way to provide relative probabilities (weightings) for different proportions p. This document sets out a draft framework to address this, by way of giving examples, with the intent that it be refined by the PWG in initial discussions, and then taken forward in discussion with the International Panel at the IWS over November 28—December 2 later this year.

Framework

Details of the framework suggested are set out in the form of examples in the Appendix. First three different stock-recruitment relationships are fit to the outputs from one example assessment for various proportions (p) of South Coast spawning biomass contributing to West Coast recruitment. The results are then combined using Akaike weighting as a measure of the relative support in this information for different values of p, with results as shown in Figure A.4 of the Appendix.

These would be the final output for weightings if the hydrodynamics model were considered purely as an hypothesis generator, with an associated uniform (uninformative) prior for p. However conceivably a prior for p might be associated with this hydrodynamics model. An example of the outcome from such an approach, in combination with the Akaike weighting of the stock-recruitment fit information, is provided in Figure A.5.

Take care not to confuse annual variability in p with uncertainty in the value of p. In reality p varies from year to year, and its annual values provide a distribution, though that distribution is, of course, unknown. The symbol p as used in this document, however, refers to a central statistic (e.g. the median) of that distribution. Thus the distributions/weightings referenced in this document relate to the extent of uncertainty about the value for that central statistic, and have nothing to do with the extent to which the actual value of the proportion contributing to west coast recruitment may vary from year to year.

Discussion

Clearly some refinements are needed in taking this approach forward. Of the stock-recruitment curves considered, the Ricker provides clearly the worst fit to the data (see lnL values in Table A.2), and therefore makes little contribution to the final weighting shown in Figure A.4. The role for the Hockey-stick form in such a computation is also questionable because it forces recruitment to be independent of spawning biomass over most of the range of that biomass, and therefore has less potential to distinguish between different values of p, although the behaviour of the log likelihood below values of p of about 0.06 is interesting and merits more investigation to identify the reason. However, further work should perhaps rather focus on a single generalised relationship with an extra parameter to allow for different shapes. It is also important to note that while overall this particular analysis of stock-recruitment function fits suggests more of the South Coast spawning biomass contributing to West Coast recruitment to be the more likely hypotheses (aside from the mini-peak at very low p for the Hockey-stick form), the information content of this approach is not high, and the possibility that there is no contribution from the South Coast spawning biomass is certainly not excluded (Figure A.4).

Providing a prior for the results of the hydrodynamics model (which provides a point estimate of 0.083 for p) is more difficult, because this needs to take account a number of aspects:

a) sensitivity of the point estimate of $p = 0.083$ to variations of the input parameters to the model over their ranges of uncertainties (unfortunately it is not immediately possible to obtain such information);

b) consideration of the impact of possible alternative structures for the model; and

c) the confidence to be placed on inferences from such models, given that fundamentally this constitutes a model of the impact of some environmental effects on recruitment, and such models have not proven to have great reliability elsewhere when their predictions have been tested as new data become available.

Future work

The following areas of future work towards an improved basis for further discussion of this with the Panel at the IWS might be valuable:

1) The stock-recruitment values in Table A.1 arise from an example assessment which included the assumption of a particular stock-recruitment function (hockey-stick in this case). It would be preferable to use results from an assessment which does not add any strong assumption about the form of a stock recruitment function to its estimation process.

2) A more general stock-recruitment form be fit to the values from the assessment for the reasons explained above, e.g.:

$$R_y^{mod} = \frac{\alpha S_y}{1 + \frac{\alpha}{\beta}(S_y)^\gamma}$$

with results integrated over a range of values of $\gamma > 0$. (Note that $\gamma = 1$ gives the Beverton-Holt form, while values of $\gamma > 1$ produce a Ricker-like dome.)

3) A document be developed motivating the choice of a prior for p based on the hydrodynamics model.

4) The literature (e.g. Myers, RAM: When do environment-recruitment correlations work? Reviews in Fish Biology and Fisheries 8, 285-305, 1998) be searched for evidence of the predictive success or otherwise of environment-recruitment relationships based on oceanographic relationships which have been advanced.

5) Developing a similar framework to this for assigning weights to different hypotheses for future west-to-east movement should be considered.

Acknowledgment

Andrea Ross-Gillespie assisted with the computations and development of the plots.

Appendix: Methodology

Fitting the recruitment curves

Estimates for recruitment and spawning biomass are available for the West and South Coast for the years 1984-2015 (Table A.1 and Figure A.1). Three curves are fitted to this spawning biomass and West Coast recruitment information:

1. Hockey stick:

$$R_y^{mod} = \begin{cases} \alpha S_y & \text{for } S_y \leq \beta/\alpha \\ \beta & \text{for } S_y > \beta/\alpha \end{cases} \tag{1}$$

2. Beverton-Holt

$$R_y^{mod} = \frac{\alpha S_y}{1 + \frac{\alpha}{\beta} S_y} \tag{2}$$

3. Ricker

$$R_y^{mod} = \alpha S_y e^{-\beta S_y} \tag{3}$$

where α and β are estimable parameters and S_y is a measure of the spawning biomass contributing to the West Coast recruitment given by:

$$S_y = S_y^{west} + p S_y^{south} \tag{4}$$

where S_y^{west} and S_y^{south} are the respective West and South Coast spawning biomass estimates, and p is an estimable parameter that can take on values between 0 and 1.

Assuming that residuals are log-normally distributed, the negative log-likelihood is given by:

$$-lnL = \sum_y \left(ln\sigma_y + \left(lnR_y^{west} - lnR_y^{mod} \right)^2 / (2\sigma_y^2) \right) \tag{5}$$

where R_y^{west} is the recruitment estimate for the West Coast in year y and σ_y is given by:

$$\sigma_y = \begin{cases} \sigma_2 & \text{for } 2000 \leq y \leq 2004 \\ \sigma_1 & \text{otherwise} \end{cases} \tag{6}$$

(It has been conventional to assume a higher variance for these five "peak-related" years in previous analyses.)

Closed form equations are available for the maximum likelihood estimates for the σ_1 and σ_2 parameters:

$$\sigma_i = \sqrt{\frac{1}{n} \sum_{y \in Y_i} \left(lnR_y^{west} - lnR_y^{mod} \right)^2} \tag{7}$$

where the set $Y_2 = \{2000; 2004\}$ and Y_1 contains the remaining years.

Results from these fits are given in Table A.2 and Figures A.2 and A.3.

Probability histogram for p (Akaike weights)

The following steps were taken to obtain a probability histogram for p, weighted across the three different stock-recruitment models:

1. Assume that the three models (model$_i$ $\in$\{Hockey stick, Beverton-Holt, Ricker\}) and 11 p values ($p_j \in$\{0.0, 0.1; ..., 1.0\}) are equally likely each with a consequent prior weight of $W_{ij}^{prior} = 1/33$ (i.e. a uniform prior distribution[1]).

2. Multiply this prior weight by the Akaike weight for each (model$_i$, p_j) combination:

$$W_{ij} = W_{ij}^{prior} e^{-\Delta lnL_{ij}} \tag{8}$$

   where ΔlnL_{ij} is the difference between the negative log likelihood for (model$_i$, p_j) and the lowest negative log likelihood value for the 33 (model$_i$, p_j) combinations.

3. Normalise the resulting 33 W_{ij} values to sum to 1.

4. For each p_j sum over the models$_i$ to calculate $W_j = \sum_i W_{ij}$, where W_j provides the value for the probability histogram at p_j.

Results from these steps are shown in Figures A.4 and A.5.

[1] This document also reports on the probability histogram when a normal prior distribution with mean 0.083 and standard deviation 0.3 is assumed. In this case, the prior weight of (1/33) is replaced by $W_{ij}^{prior} = e^{(p_j-0.2)^2/(2(0.3)^2)}/(\sqrt{2\pi(0.3)^2})$, normalised so that $\sum_{ij} W_{ij}^{prior} = 1$.

Table A.1: Estimates of sardine recruitment (in billions) and spawning biomass (in thousand tons) available for the analysis from an example assessment (from FISHERIES/2016/JUL/SWG-PEL/22REV2).

Year	Recruitment		Spawning biomass		Year	Recruitment		Spawning biomass	
	West	South	West	South	..cont.	West	South	West	South
1983	0.80	0.04	1.06	0.01	1999	41.43	4.23	132.98	197.63
1984	9.27	1.58	14.98	0.66	2000	167.27	2.45	110.05	203.66
1985	12.46	2.08	14.98	7.92	2001	132.72	2.70	76.21	277.32
1986	18.70	1.66	49.74	29.67	2002	114.29	2.25	164.94	644.60
1987	13.38	1.84	115.60	38.66	2003	22.49	2.40	179.03	719.10
1988	14.58	2.48	123.28	15.71	2004	15.50	2.75	146.33	1209.63
1989	13.94	2.43	232.48	26.76	2005	25.27	3.40	37.93	673.96
1990	14.93	2.56	233.32	29.96	2006	14.28	2.72	25.10	251.11
1991	24.26	2.52	131.81	35.22	2007	18.44	2.59	23.60	155.06
1992	32.57	2.56	105.58	53.60	2008	18.11	4.25	28.51	98.08
1993	14.98	2.27	162.05	53.58	2009	53.82	1.97	97.19	141.95
1994	33.51	1.85	264.11	95.38	2010	13.39	2.46	180.89	74.64
1995	25.09	2.29	172.47	65.59	2011	27.03	2.66	347.02	196.07
1996	43.13	2.57	180.84	62.68	2012	19.58	1.65	218.57	76.17
1997	39.06	2.05	197.03	76.89	2013	17.31	4.12	126.61	140.67
1998	37.41	2.96	172.73	75.52	2014	28.12	3.18	147.46	176.59

Table A.2: Parameter estimates for the three different forms of stock-recruitment relationship and three different values for p for each form.

Form	p	α	β	σ_1	σ_2	-lnL
(a) Hockey-stick	0.0	0.692	24.381	0.384	1.355	-8.32
	0.5	0.670	23.303	0.397	1.386	-7.30
	1.0	0.623	23.306	0.399	1.386	-7.20
(b) Beverton-Holt	0.0	1.163	29.074	0.403	1.385	-6.91
	0.5	0.840	27.981	0.386	1.351	-8.22
	1.0	0.792	27.127	0.385	1.343	-8.31
(c) Ricker	0.0	0.693	0.008	0.432	1.248	-5.55
	0.5	0.382	0.005	0.465	1.449	-2.82
	1.0	0.262	0.003	0.521	1.487	0.36

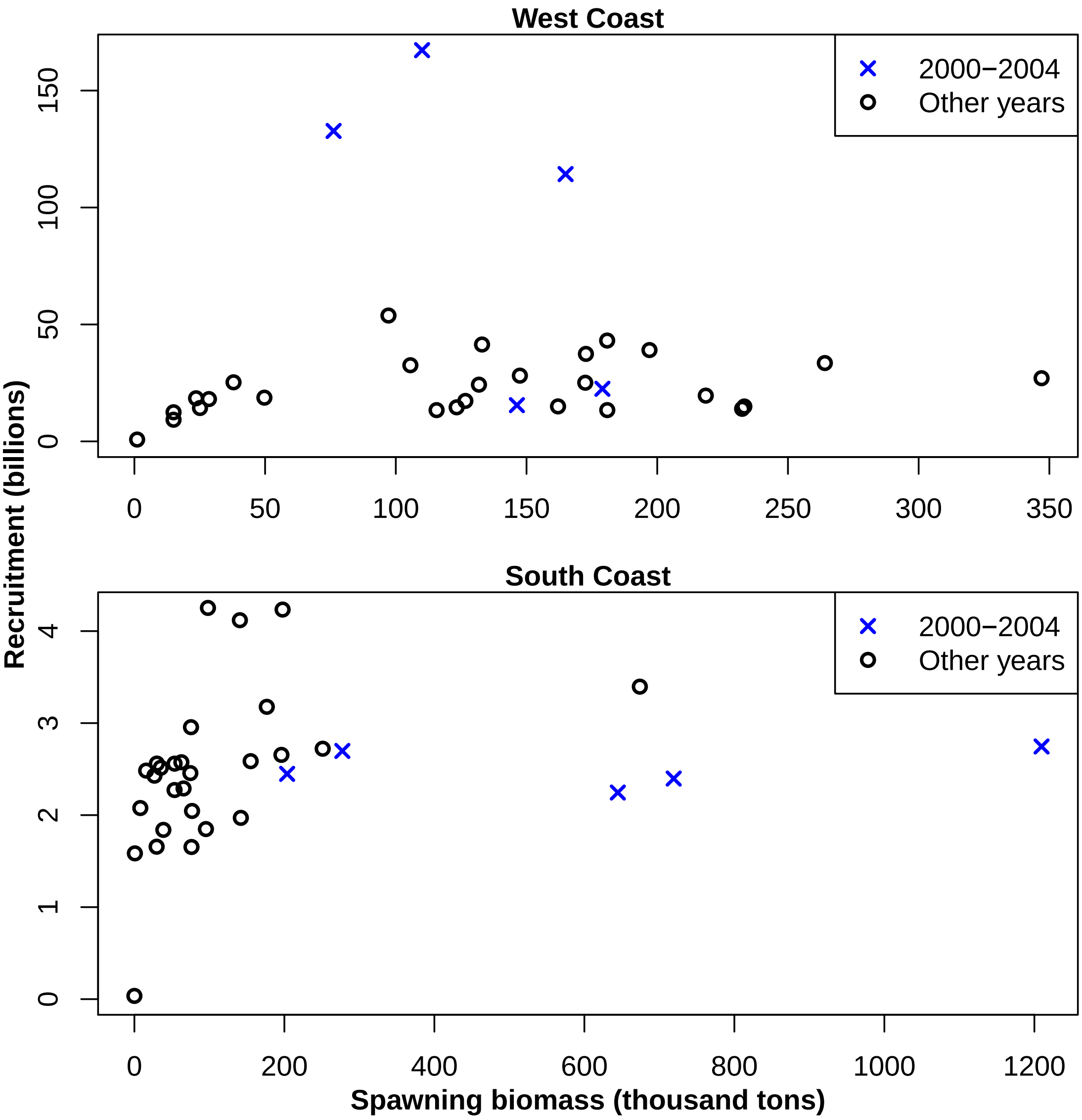

Figure A.1: Plots of recruitment vs spawning biomass for the example assessment. Blue crosses indicate point corresponding to the years 2000-2004, which have conventionally been taken to reflect a higher variance about the stock-recruitment relationship.

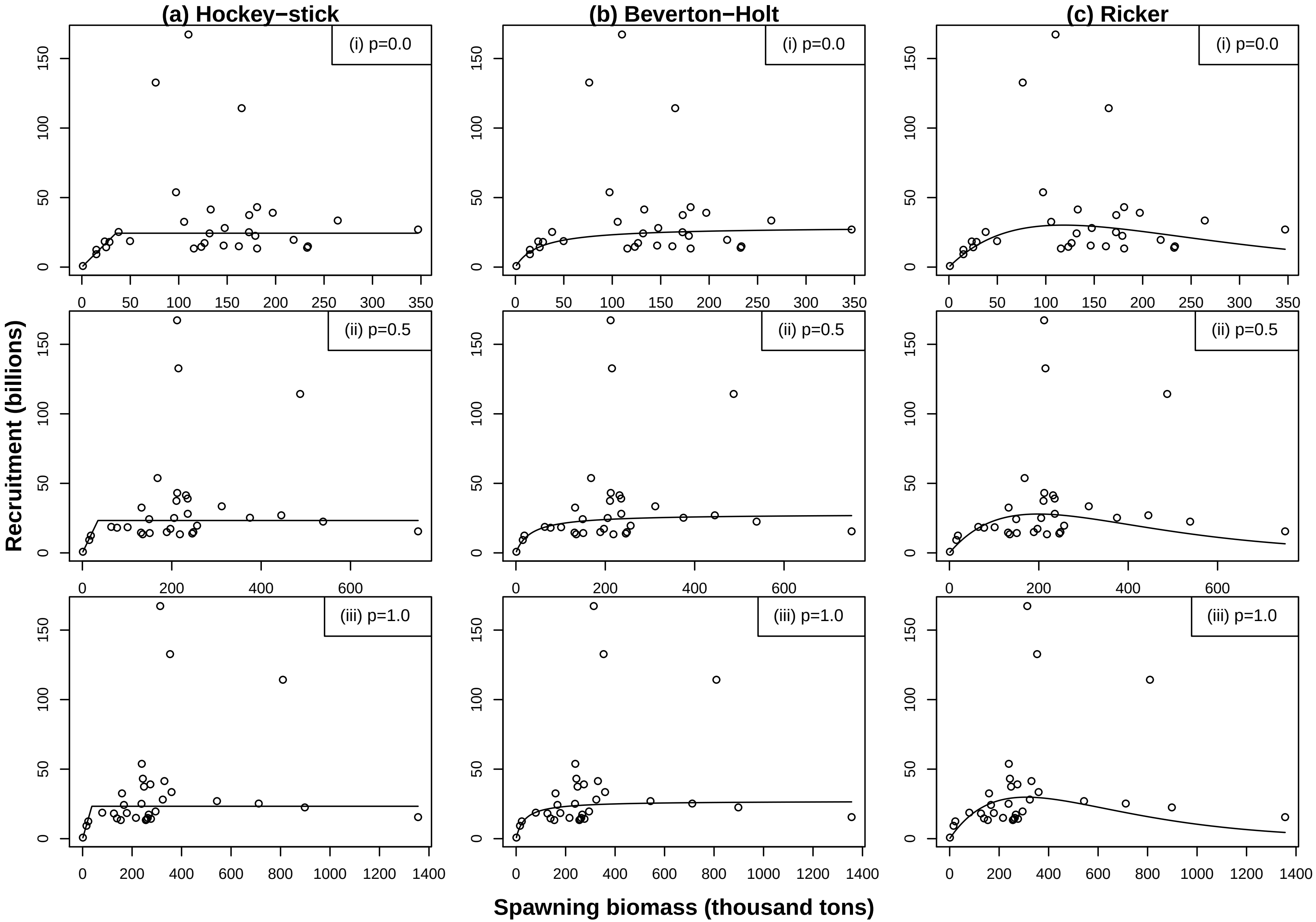

Figure A.2: Plots of the fits of the estimated stock-recruitment relationships to the spawning biomass and recruitment estimates from the example assessment.

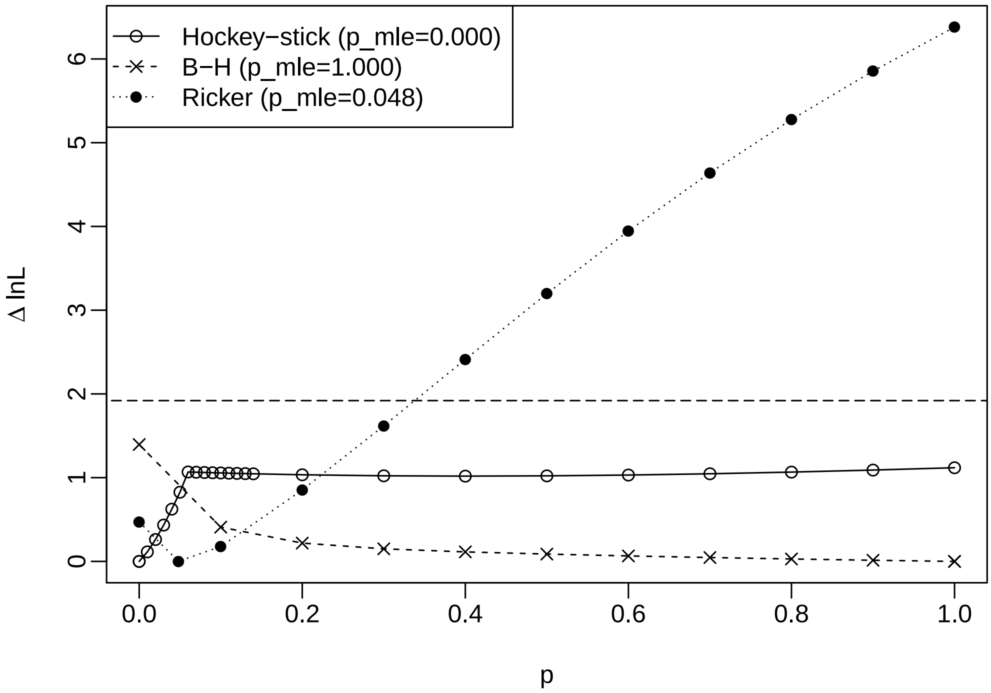

Figure A.3: Negative log-likelihood profiles for the p parameter for the three different stock-recruitment relationships, where each is shown relative to its lowest value over the range of [0, 1] for p. The legend includes MLEs for p over the [0, 1] range considered for each relationship; points above the horizontal dashed line correspond to values of p which are outside the associated 95% CI.

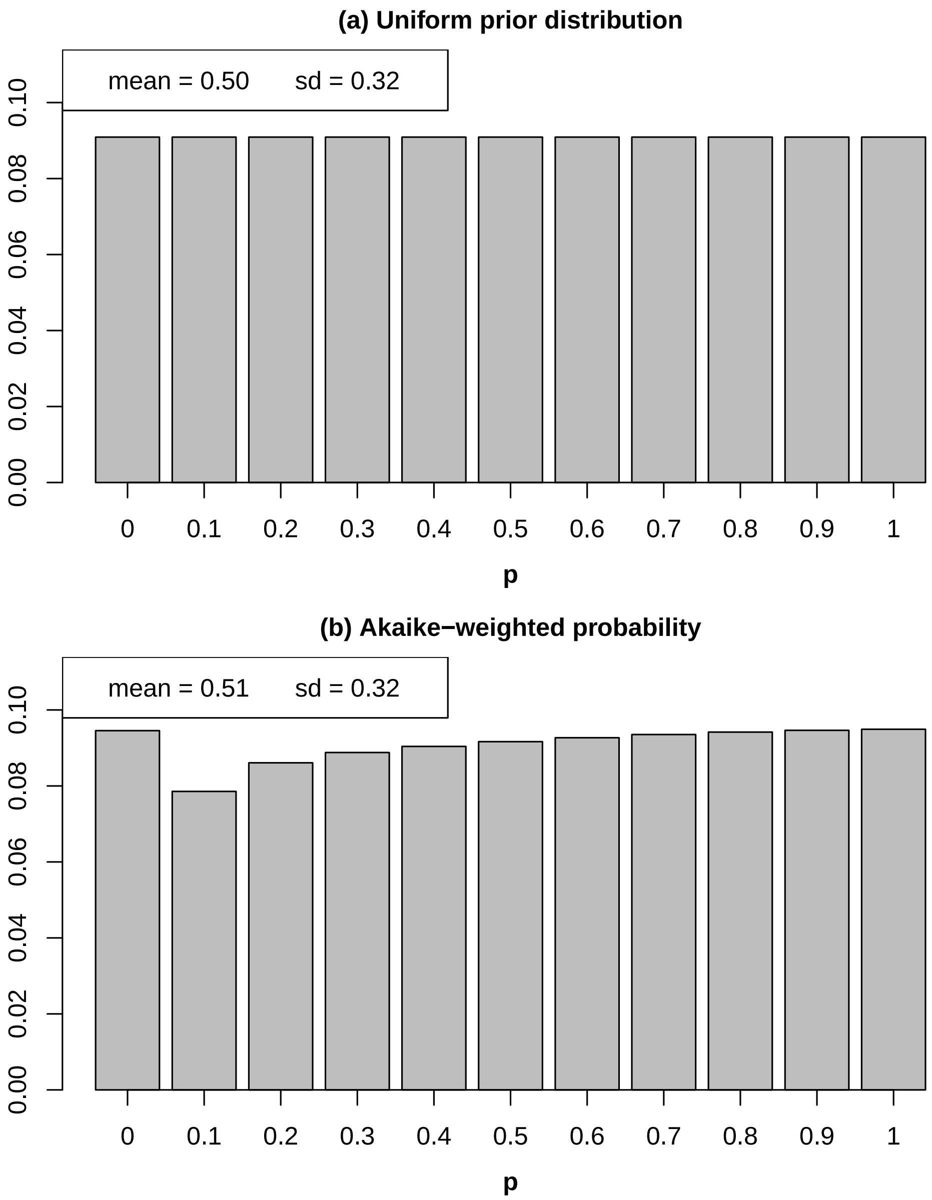

Figure A.4: Probability histogram for p calculated using Akaike weights as described in the Appendix under the assumption of a uniform prior distribution. The top panel shows the uniform prior distribution and the bottom panel shows the distribution when this prior is multiplied by the Akaike weights. The mean and standard deviation for each distribution are reported in the top left legend.

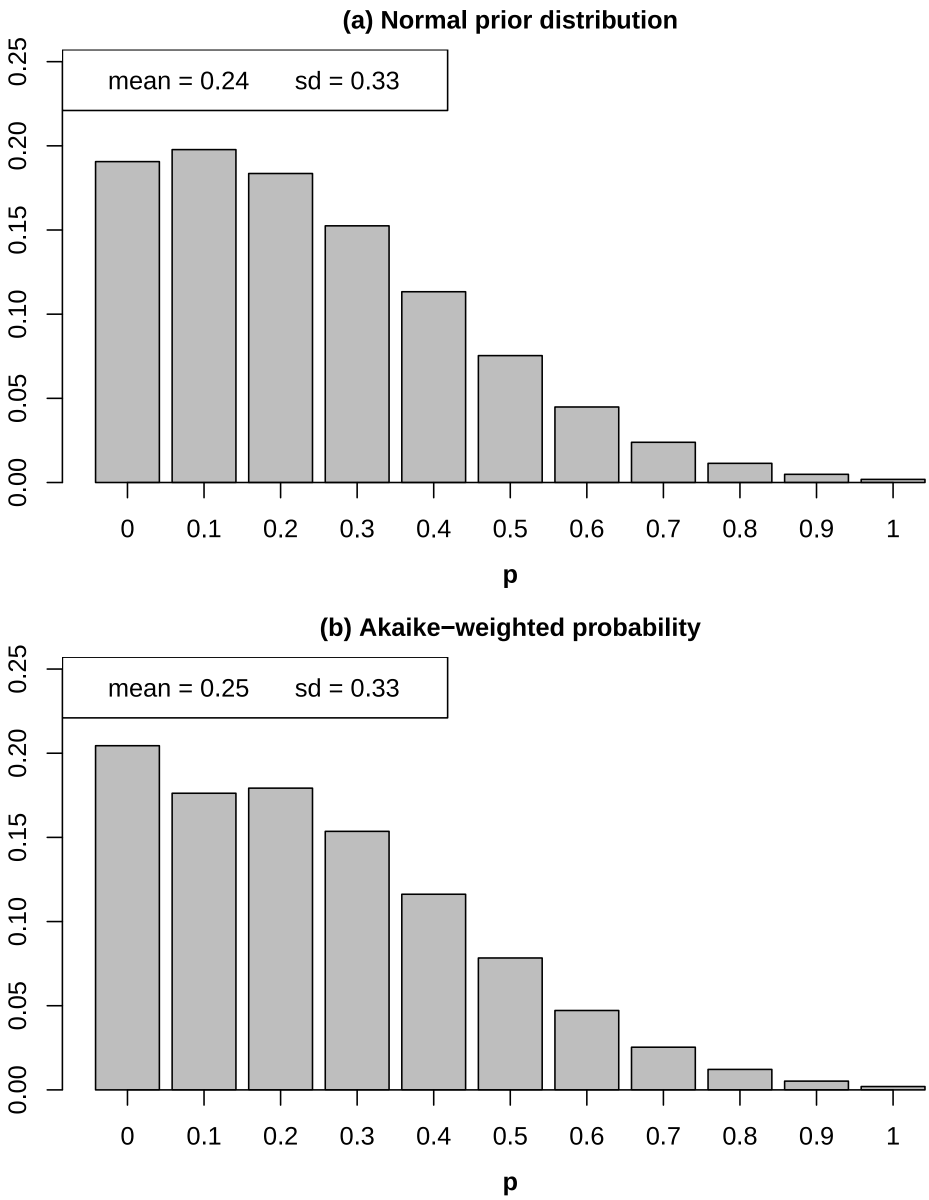

Figure A.5: Probability histogram for p calculated using Akaike weights from the fits to the stock recruitment relationships as described in the Appendix, but here under a normal prior distribution with a mean of 0.083 and a standard deviation of 0.3. Note that the mean and standard deviation reported for the prior distribution in the top panel are different to those used to generate the histogram since the whole range for the normal distribution is not included in the range for p between 0 and 1.